

В. В. ПРАСОЛОВ

# **ЗАДАЧИ И ТЕОРЕМЫ ЛИНЕЙНОЙ АЛГЕБРЫ**

Москва  
2008

**Прасолов В. В.**

П70      Задачи и теоремы линейной алгебры. — 2-е изд. — М., 2008.

© Прасолов В. В., 2008

# Оглавление

<b>1. Определители</b>	<b>17</b>
1.1. Вычисление определителей	19
1.1.1. Определение	19
1.1.2. Некоторые свойства определителя	22
1.1.3. Определитель произведения матриц	23
1.1.4. Правило Крамера	24
1.1.5. Определитель Вандермонда	24
1.1.6. Интерполяционный многочлен Лагранжа	25
1.1.7. Определитель, сводящийся к определителю Вандермонда	26
1.1.8. Определитель Коши	27
1.1.9. Матрица Фробениуса	29
1.1.10. Циркулянт	29
1.1.11. Матрица Якоби	32
1.1.12. Определитель Смита	33
1.1.13. Определитель преобразованной матрицы	34
1.1.14. Степени степенного ряда	35
1.1.15. Биномиальные определители	36
1.2. Миноры и алгебраические дополнения	42
1.2.1. Определение минора	42
1.2.2. Ранг матрицы	42
1.2.3. Формула Бине—Коши	43
1.2.4. Алгебраическое дополнение минора	44
1.2.5. Присоединённая матрица	45
1.2.6. Миноры присоединённой матрицы. Тождество Якоби	46
1.2.7. Тождество Льюиса Кэрролла	48
1.2.8. Ассоциированная матрица	50
1.2.9. Обобщённое тождество Сильвестра	50

1.2.10. Миноры матрицы Якоби . . . . .	51
1.2.11. Миноры матрицы Грина . . . . .	52
1.2.12. Вполне положительные матрицы . . . . .	52
1.2.13. Теорема Митчелла . . . . .	54
1.2.14. Теорема Чеботарёва . . . . .	55
1.3. Дополнение по Шуру . . . . .	58
1.3.1. Определитель блочной матрицы . . . . .	58
1.3.2. Обращение блочной матрицы . . . . .	61
1.3.3. Теорема Хэйнсворт . . . . .	61
1.3.4. Пфаффиан . . . . .	62
1.4. Симметрические функции. Суммы степеней. Числа Бер- нулли . . . . .	63
1.4.1. Примеры симметрических многочленов . . . . .	64
1.4.2. Основная теорема о симметрических многочленах . . . . .	66
1.4.3. Сумма степеней . . . . .	68
1.4.4. Теорема Фаульгабера—Якоби . . . . .	68
1.4.5. Числа Бернулли . . . . .	71
1.4.6. Функции Шура . . . . .	72
1.4.7. Формула Джамбелли . . . . .	75
1.4.8. Скалярное произведение в пространстве симмет- рических многочленов . . . . .	77
1.4.9. Комбинаторное определение функций Шура . . . . .	79
Решения . . . . .	80
<b>2. Линейные пространства</b>	<b>97</b>
2.1. Двойственное пространство. Ортогональное дополнение . . . . .	101
2.1.1. Линейно независимые векторы . . . . .	101
2.1.2. Двойственное пространство . . . . .	103
2.1.3. Двойственный оператор . . . . .	104
2.1.4. Отождествление $V$ и $V^*$ при наличии метрики . . . . .	104
2.1.5. Системы линейных уравнений . . . . .	105
2.1.6. Разрезание прямоугольника на квадраты . . . . .	106
2.1.7. Количество треугольников среди частей, на кото- рые прямые разрезают плоскость . . . . .	107
2.1.8. Аннулятор . . . . .	109
2.1.9. Скалярное произведение в пространстве матриц . . . . .	110
2.1.10. Аффинное пространство . . . . .	110
2.2. Ядро и образ оператора. Факторпространство . . . . .	114
2.2.1. Ядро и образ оператора . . . . .	114
2.2.2. Альтернатива Фредгольма . . . . .	114
2.2.3. Теорема Кронекера—Капелли . . . . .	116

2.2.4.	Факторпространство . . . . .	117
2.2.5.	Точные последовательности . . . . .	118
2.3.	Базисы. Линейная независимость . . . . .	119
2.3.1.	Характеристический многочлен оператора . . . . .	119
2.3.2.	Ориентация . . . . .	120
2.3.3.	Решётки . . . . .	121
2.3.4.	Свободные абелевы группы . . . . .	122
2.3.5.	Две теоремы об обмене векторов базисов . . . . .	123
2.3.6.	Разбиение $\mathbb{R}^n$ на многогранные углы . . . . .	125
2.3.7.	Линейная зависимость степеней оператора . . . . .	127
2.4.	Ранг матрицы . . . . .	129
2.4.1.	Неравенства для ранга матрицы . . . . .	129
2.4.2.	Другое определение ранга . . . . .	130
2.4.3.	Подпространства матриц ограниченного ранга . . . . .	130
2.5.	Подпространства. Ортогонализация . . . . .	134
2.5.1.	Размерность пересечения двух пространств . . . . .	134
2.5.2.	Ортогонализация Грама—Шмидта . . . . .	134
2.5.3.	Ортогональные проекции . . . . .	136
2.5.4.	Симметрия относительно подпространства . . . . .	137
2.5.5.	Угол между вектором и подпространством . . . . .	139
2.5.6.	Проекции ортонормированных базисов . . . . .	140
2.5.7.	Равные проекции ортогональных векторов . . . . .	142
2.5.8.	Определители, составленные из скалярных произведений . . . . .	143
2.5.9.	Системы подпространств . . . . .	145
2.5.10.	Теорема Шерка . . . . .	145
2.6.	Ортогональные многочлены . . . . .	149
2.6.1.	Общие свойства . . . . .	149
2.6.2.	Многочлены Лежандра . . . . .	150
2.6.3.	Многочлены Эрмита . . . . .	152
2.6.4.	Многочлены Чебышева . . . . .	153
2.7.	Комплексификация и о веществе. Эрмитовы пространства . . . . .	155
2.7.1.	Комплексификация . . . . .	155
2.7.2.	О веществе . . . . .	155
2.7.3.	Эрмитовы пространства . . . . .	156
2.7.4.	Унитарные, эрмитовы и косоэрмитовы операторы . . . . .	157
2.7.5.	Нормальные операторы . . . . .	159
2.7.6.	Комплексные структуры . . . . .	159
2.7.7.	Комплексно сопряжённое пространство . . . . .	161
2.7.8.	Метрика Фубини—Штуди . . . . .	162

2.7.9.	Кэлеров угол . . . . .	163
2.8.	Нормированные пространства . . . . .	166
2.8.1.	Тождество параллелограмма . . . . .	167
2.8.2.	Эквивалентные нормы . . . . .	168
2.8.3.	Метрическое определение середины отрезка . . . . .	170
2.8.4.	Изометрические отображения . . . . .	171
2.8.5.	Неравенства для норм . . . . .	172
2.8.6.	Двойственная норма . . . . .	174
Решения	. . . . .	176
<b>3.</b>	<b>Канонические формы матриц и линейных операторов</b>	<b>190</b>
3.1.	След и собственные значения оператора . . . . .	190
3.1.1.	След . . . . .	190
3.1.2.	Собственные векторы и собственные значения . . . . .	191
3.1.3.	Диагонализируемые операторы . . . . .	191
3.1.4.	Спектр матрицы Якоби . . . . .	193
3.1.5.	Собственные значения полинома от матрицы . . . . .	194
3.1.6.	Собственные значения матриц $AB$ и $BA$ . . . . .	196
3.1.7.	Собственные значения и суммы элементов столбцов . . . . .	196
3.1.8.	Матрица Каца . . . . .	198
3.1.9.	Теорема Годдарда—Шнайдера . . . . .	198
3.2.	Жорданова нормальная форма . . . . .	201
3.2.1.	Подобные матрицы . . . . .	201
3.2.2.	Существование и единственность жордановой формы . . . . .	202
3.2.3.	Возведение матрицы в степень . . . . .	204
3.2.4.	Жорданова форма над $\mathbb{R}$ . . . . .	204
3.2.5.	Аддитивная и мультипликативная запись жорданово- вой формы . . . . .	205
3.2.6.	Классификация Кронекера пар линейных операторов . . . . .	207
3.3.	Минимальный многочлен и характеристический многочлен	213
3.3.1.	Минимальный многочлен . . . . .	213
3.3.2.	Теорема Гамильтона—Кэли . . . . .	214
3.3.3.	Критерий совпадения характеристического много- члена и минимального многочлена . . . . .	215
3.3.4.	Обобщения теоремы Гамильтона—Кэли . . . . .	216
3.4.	Каноническая форма Фробениуса . . . . .	218
3.4.1.	Циклические клетки . . . . .	218
3.4.2.	Характеристический многочлен циклической клетки	219
3.5.	Приведение диагонали к удобному виду . . . . .	220

3.5.1.	Все диагональные элементы, кроме одного, равны нулю . . . . .	220
3.5.2.	Все диагональные элементы равны . . . . .	222
3.5.3.	Все диагональные элементы ненулевые . . . . .	223
3.6.	Полярное разложение . . . . .	224
3.6.1.	Существование полярного разложения . . . . .	225
3.6.2.	Разложение $UDW$ . . . . .	226
3.7.	Разложения матриц . . . . .	226
3.7.1.	Разложение Шура . . . . .	226
3.7.2.	Разложение Ланцоша . . . . .	227
3.7.3.	Произведение двух симметрических матриц . . . . .	228
3.8.	Нормальная форма Смита. Элементарные делители матриц . . . . .	229
3.8.1.	Нормальная форма Смита . . . . .	229
3.8.2.	Элементарные делители матриц . . . . .	231
3.8.3.	Целочисленные решения систем линейных уравнений . . . . .	232
3.8.4.	Классификация абелевых групп . . . . .	233
Решения	. . . . .	234
<b>4.</b>	<b>Матрицы специального вида</b>	<b>242</b>
4.1.	Симметрические и эрмитовы матрицы . . . . .	242
4.1.1.	Квадратичные и эрмитовы формы . . . . .	242
4.1.2.	Сигнатура квадратичной формы . . . . .	244
4.1.3.	Приведение к диагональному виду . . . . .	246
4.1.4.	Квадратный корень из неотрицательно определённой матрицы . . . . .	246
4.1.5.	Теорема Куранта—Фишера . . . . .	247
4.1.6.	Собственные значения произведения двух эрмитовых матриц . . . . .	248
4.2.	Одновременное приведение пары эрмитовых форм к диагональному виду . . . . .	250
4.2.1.	Случай положительно определённой матрицы . . . . .	250
4.2.2.	Случай знакоопределённой матрицы . . . . .	251
4.2.3.	Ещё один случай . . . . .	252
4.3.	Кососимметрические матрицы . . . . .	253
4.3.1.	Простейшие свойства . . . . .	254
4.3.2.	Канонический вид кососимметрической формы . . . . .	254
4.3.3.	Канонический вид кососимметрического оператора . . . . .	255
4.4.	Ортогональные матрицы и преобразование Кэли . . . . .	256
4.4.1.	Преобразование Кэли . . . . .	257
4.4.2.	Обобщённое преобразование Кэли . . . . .	257

4.5. Нормальные матрицы . . . . .	259
4.5.1. Ядро и образ нормального оператора . . . . .	260
4.5.2. Собственные векторы нормального оператора . . . . .	260
4.5.3. Полиномиальное выражение $A^*$ через $A$ . . . . .	261
4.6. Нильпотентные матрицы . . . . .	261
4.6.1. Характеристический и минимальный многочлены . . . . .	261
4.6.2. Критерии нильпотентности . . . . .	262
4.6.3. Диаграмма Юнга нильпотентной матрицы . . . . .	263
4.7. Проекторы . . . . .	264
4.7.1. Канонический вид матрицы проектора . . . . .	264
4.7.2. Эрмитовы проекторы . . . . .	265
4.7.3. Матрица проектора на подпространство . . . . .	266
4.7.4. Определитель суммы проекторов . . . . .	266
4.8. Инволюции . . . . .	269
4.8.1. Канонический вид матрицы инволюции . . . . .	269
4.8.2. Произведение двух инволюций . . . . .	269
Решения . . . . .	270
<b>5. Полилинейная алгебра . . . . .</b>	<b>279</b>
5.1. Полилинейные отображения. Тензорные произведения . . . . .	279
5.1.1. Определения . . . . .	279
5.1.2. Билинейные формы . . . . .	280
5.1.3. Изоморфизм $V^* \otimes W \rightarrow \text{Hom}(V, W)$ . . . . .	281
5.1.4. Валентность тензора . . . . .	282
5.1.5. Тензорное произведение отображений . . . . .	283
5.1.6. Матричные уравнения . . . . .	284
5.1.7. Подпространство, соответствующее полилинейной функции . . . . .	285
5.2. Симметрические и кососимметрические тензоры . . . . .	287
5.2.1. Симметризация и альтернирование . . . . .	287
5.2.2. Алгебра Грассмана . . . . .	288
5.2.3. Кососимметрические функции . . . . .	289
5.2.4. Свойства внешнего произведения . . . . .	291
5.2.5. Тензорная степень оператора . . . . .	292
5.2.6. Отображение Ходжа . . . . .	293
5.2.7. Векторное произведение . . . . .	295
5.3. Пфаффиан . . . . .	297
5.3.1. Определитель кососимметрической матрицы . . . . .	297
5.3.2. Пфаффиан матриц специального вида . . . . .	298
5.4. Разложимые тензоры . . . . .	300
5.4.1. Разложимый тензор определяет подпространство . . . . .	300



5.4.2.	Соотношения Плюккера . . . . .	301
5.4.3.	Альтернатива соотношениям Плюккера . . . . .	304
5.4.4.	Свойства отображения $d = i(v)$ . . . . .	305
5.5.	Тензорный ранг . . . . .	307
5.5.1.	Тензорный ранг в $V \otimes W$ . . . . .	307
5.5.2.	Алгоритм Штрассена . . . . .	308
5.5.3.	Тензорный ранг в $V_1 \otimes V_2 \otimes V_3$ . . . . .	309
5.6.	Линейные отображения пространств матриц . . . . .	310
5.6.1.	Отображения, сохраняющие ранг 1 . . . . .	310
5.6.2.	Отображения, сохраняющие определитель . . . . .	312
5.6.3.	Отображения, сохраняющие собственные значения . . . . .	312
5.6.4.	Отображения, сохраняющие невырожденность . . . . .	313
5.7.	Образ полилинейного отображения . . . . .	314
5.7.1.	Два примера . . . . .	315
5.7.2.	Образ отображения в 2-мерное подпространство . . . . .	315
	Решения . . . . .	316
<b>6.</b>	<b>Матричные неравенства</b>	<b>323</b>
6.1.	Неравенства для симметрических и эрмитовых матриц . . . . .	323
6.1.1.	Простейшие неравенства . . . . .	323
6.1.2.	Неравенство Адамара . . . . .	324
6.1.3.	Определитель линейной комбинации . . . . .	325
6.1.4.	Отношение миноров . . . . .	327
6.2.	Неравенства для собственных значений . . . . .	329
6.2.1.	Неравенство Шура . . . . .	329
6.2.2.	Сумма двух эрмитовых матриц . . . . .	330
6.2.3.	Произведение двух эрмитовых проекторов . . . . .	330
6.2.4.	Сингулярные значения . . . . .	331
6.2.5.	Круги Гершгорина . . . . .	332
6.3.	Неравенства для норм матриц . . . . .	337
6.3.1.	Операторная норма . . . . .	337
6.3.2.	Евклидова норма . . . . .	338
6.3.3.	Наилучшее приближение эрмитовой или унитарной матрицей . . . . .	338
6.3.4.	Наилучшее приближение невырожденной матрицы вырожденной матрицей . . . . .	339
6.4.	Дополнение по Шуру и произведение Адамара . . . . .	340
6.4.1.	Дополнение по Шуру положительно определённой матрицы . . . . .	340
6.4.2.	Произведение Адамара . . . . .	342
6.5.	Неотрицательные матрицы . . . . .	343

6.5.1. Неразложимые матрицы . . . . .	344
6.5.2. Экстремальные векторы . . . . .	344
6.5.3. Канонический вид неразложимой матрицы . . . . .	345
6.5.4. Примитивные матрицы . . . . .	347
6.5.5. Теорема Стильтеса—де Рама . . . . .	349
6.6. Дважды стохастические матрицы . . . . .	350
6.6.1. Простейшие свойства . . . . .	351
6.6.2. Теорема Биркгофа . . . . .	351
6.6.3. Неравенство Г. Вейля . . . . .	353
6.6.4. Теорема Гофмана—Виландта . . . . .	354
Решения . . . . .	356
<b>7. Коммутаторы</b>	<b>363</b>
7.1. Коммутирующие операторы . . . . .	363
7.1.1. Матрицы, перестановочные с данной . . . . .	363
7.1.2. Семейства коммутирующих операторов . . . . .	365
7.1.3. Перестановочность с $A$ влечёт перестановочность с $B$ . . . . .	367
7.1.4. Теорема Шура . . . . .	368
7.2. Свойства коммутаторов . . . . .	370
7.2.1. Определение и простейшие свойства . . . . .	370
7.2.2. Представление матриц в виде коммутаторов . . . . .	371
7.2.3. Равенство $\text{ad}_A^s X = 0$ влечёт равенство $\text{ad}_X^s B = 0$ . . . . .	372
7.2.4. Одновременная триангулируемость . . . . .	373
7.2.5. Триангулируемость операторов с рангом коммута- тора 1 . . . . .	375
7.3. Теория реплик . . . . .	376
7.3.1. Определение и основные свойства . . . . .	376
7.3.2. Полупростая и нильпотентная составляющая . . . . .	377
7.3.3. Комплексификация . . . . .	378
7.3.4. Критерий нильпотентности . . . . .	379
7.4. Элементы теории алгебр Ли . . . . .	380
7.4.1. Нильпотентные алгебры Ли . . . . .	382
7.4.2. Разрешимые алгебры Ли . . . . .	383
7.4.3. Представления алгебры Ли $\mathfrak{sl}(2, \mathbb{C})$ . . . . .	385
7.4.4. Критерий Картана . . . . .	386
Решения . . . . .	389
<b>8. Матрицы в алгебре и анализе</b>	<b>394</b>
8.1. Кватернионы и числа Кэли. Алгебры Клиффорда . . . . .	394
8.1.1. Удвоение алгебры . . . . .	394
8.1.2. Кватернионы . . . . .	395

8.1.3.	Кватернионы и простейшие алгебры Ли . . . . .	396
8.1.4.	Кватернионы и движения . . . . .	397
8.1.5.	Изоморфизм $\mathbb{H} \otimes \mathbb{H}$ и $M_4(\mathbb{R})$ . . . . .	399
8.1.6.	Числа Кэли . . . . .	400
8.1.7.	Векторное произведение чисел Кэли . . . . .	402
8.1.8.	Антикоммутирующие комплексные структуры . . .	402
8.1.9.	Доказательство теоремы о комплексных структурах	405
8.2.	Представления матричных алгебр . . . . .	408
8.2.1.	Представления матричных алгебр . . . . .	408
8.2.2.	Теорема Бернсайда . . . . .	409
8.3.	Конечные поля . . . . .	410
8.3.1.	Линейные пространства над конечными полями . .	410
8.3.2.	Комбинаторика подпространств . . . . .	411
8.3.3.	Минимальный многочлен вектора . . . . .	411
8.3.4.	Квадратичные формы над конечными полями . . .	412
8.3.5.	Квадратичные формы над полями характеристики 2	414
8.3.6.	Инвариант Арфа квадратичной формы . . . . .	417
8.4.	Результант . . . . .	419
8.4.1.	Матрица Сильвестра . . . . .	419
8.4.2.	Выражение результанта через корни . . . . .	419
8.4.3.	Матрица Безу . . . . .	420
8.4.4.	Понижение порядка матрицы для вычисления ре- зультанта . . . . .	422
8.4.5.	Дискриминант . . . . .	423
8.5.	Теорема Витта . . . . .	425
8.5.1.	Пространства Артина . . . . .	425
8.5.2.	Доказательство теоремы Витта . . . . .	426
8.6.	Обобщённая обратная матрица. Матричные уравнения . .	428
8.6.1.	Обобщённая обратная матрица . . . . .	428
8.6.2.	Решение несовместных систем уравнений . . . . .	429
8.6.3.	Матричные уравнения . . . . .	430
8.7.	Ганкелевы и тёплицевы матрицы . . . . .	433
8.7.1.	Корни многочленов . . . . .	433
8.7.2.	Ганкелевы матрицы и рациональные функции . . .	435
8.8.	Функции от матриц. Дифференцирование матриц . . . .	437
8.8.1.	Экспонента . . . . .	437
8.8.2.	Дифференцирование матриц . . . . .	438
8.8.3.	Системы дифференциальных уравнений . . . . .	439
8.8.4.	Вронскиан . . . . .	440
8.9.	Пары Лакса и интегрируемые системы . . . . .	443
8.9.1.	Уравнение Лакса . . . . .	443

8.9.2. Цепочка Тоды . . . . .	444
8.9.3. Уравнения движения твёрдого тела . . . . .	445
8.9.4. Уравнения Вольтерра . . . . .	446
8.10. Матрицы с предписанными собственными значениями . . . . .	447
8.10.1. Фиксированная диагональ . . . . .	447
8.10.2. Фиксированные внедиагональные элементы . . . . .	448
8.10.3. Предписанная диагональ . . . . .	450
8.11. Числовой образ оператора . . . . .	451
8.11.1. Теорема Тёплица–Хаусдорфа . . . . .	451
8.11.2. Числовой образ нормального оператора . . . . .	452
8.11.3. Числовой образ и нормальные расширения . . . . .	453
8.11.4. Приложение к коммутаторам . . . . .	454
8.11.5. Числовой образ вещественных квадратичных форм . . . . .	454
8.11.6. Отображение, заданное двумя эрмитовыми матри- цами . . . . .	456
Решения . . . . .	460
<b>9. Некоммутативная линейная алгебра . . . . .</b>	<b>472</b>
9.1. Матрицы с некоммутирующими элементами . . . . .	472
9.1.1. Векторные пространства над телами . . . . .	472
9.1.2. Элементарные преобразования матриц . . . . .	477
9.1.3. Нормальная форма Брюа . . . . .	479
9.1.4. Коммутант группы $GL(n, K)$ . . . . .	480
9.2. Определители и собственные значения . . . . .	482
9.2.1. Определитель Дьёдонне . . . . .	482
9.2.2. Тожество Капелли . . . . .	487
9.2.3. Собственные значения . . . . .	490
9.3. Кватернионные пространства . . . . .	492
9.3.1. Кватернионная структура . . . . .	493
9.3.2. Гиперэрмитовы матрицы . . . . .	494
9.3.3. Унитарные матрицы . . . . .	496
9.3.4. Правые собственные значения . . . . .	497
9.3.5. Разложение Шура . . . . .	498
9.3.6. Определитель Штуди . . . . .	500
9.3.7. Определитель Дьёдонне . . . . .	500
9.3.8. Определитель Мура . . . . .	502
9.3.9. Топология пространства $GL(n, \mathbb{H})$ . . . . .	503
9.3.10. Левые собственные значения . . . . .	505
Решения . . . . .	506

## Предисловие ко второму изданию

Это издание существенно переработано и расширено по сравнению с предыдущим, написанным более 15 лет назад. Добавлена даже целая новая глава, посвящённая некоммутативной линейной алгебре. Добавлены также параграфы, посвящённые ортогональным многочленам, нормированным пространствам, описанию образа полилинейного отображения, теории реплик и элементам теории алгебр Ли, ганкелевым и тёплицевым матрицам, числовому образу оператора. Гораздо более подробно, чем в первом издании, изложена линейная алгебра над конечными полями.

Ещё до выхода первого русского издания (Физматлит, 1996) появился перевод этой книги на английский язык (AMS, 1994). Недавно были изданы переводы на венгерский (Typotex, 2005) и французский (Cassini, 2007) языки.

## Предисловие к первому изданию

Книг по линейной алгебре написано очень много, и среди них есть поистине замечательные (см., например, список рекомендуемой литературы). Может сложиться мнение, будто других книг больше не нужно. Или несколько более осторожное мнение, будто в этих книгах сказано всё, что нужно, и сказано именно так, как нужно, а потому любая новая книга будет лишь повторением старых. Мнение очевидно неверное, но тем не менее почти общепринятое. Новые результаты в линейной алгебре появляются постоянно, и постоянно появляются новые, более простые и изящные доказательства известных теорем. Кроме того, есть не так уж мало интересных старых результатов, в учебную литературу не вошедших.

В этой книге я постарался собрать наиболее интересные задачи и теоремы линейной алгебры, доступные студентам младших курсов. Вычислительная линейная алгебра осталась при этом несколько в стороне. Значительную часть книги составляют результаты, известные лишь по журнальным публикациям. Мне кажется, что они будут интересны многим читателям.

Книга предполагает знакомство читателя с основными понятиями линейной алгебры: линейное пространство, базис, линейное отображение, определитель матрицы. Но все содержательные теоремы обычного курса линейной алгебры в книге приведены с полными доказательствами. При этом особое внимание обращено на нестандартные изящные

доказательства этих известных теорем.

Изложение ведется почти исключительно над полями действительных и комплексных чисел. Лишь изредка отмечены особенности случая полей конечной характеристики.

В линейной алгебре часто приходится осуществлять переходы от линейного оператора к матрице и обратно. Это делается без особых оговорок, но не должно приводить к недоразумениям.

Я благодарен рецензентам книги Д. В. Беклемишеву и А. И. Кострикину за ценные замечания и Д. Б. Фуксу за полезные обсуждения рукописи.

# Основные обозначения и соглашения

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} — матрица размера  $m \times n$ .$$

$a_{ij}$  — элемент  $i$ -й строки и  $j$ -го столбца матрицы  $A$ .

$\|a_{ij}\|$  — матрица  $A$ .

$\|a_{ij}\|_1^n$  — квадратная матрица  $\|a_{ij}\|$  порядка  $n$ ; для матрицы порядка  $n + 1$  бывает удобно также обозначение  $\|a_{ij}\|_0^n$ .

$\det(A)$ ,  $|A|$  — определитель матрицы  $A$ .

$|a_{ij}|_1^n$  — определитель матрицы  $\|a_{ij}\|_1^n$ .

$AB$  — произведение матрицы  $A$  размера  $p \times n$  на матрицу  $B$  размера  $n \times q$ ; является матрицей  $\|c_{ik}\|$  размера  $p \times q$ , где  $c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$  — произведение  $i$ -й строки первой матрицы на  $k$ -й столбец второй матрицы.

$\text{diag}(\lambda_1, \dots, \lambda_n)$  — диагональная матрица, т.е. матрица порядка  $n$  с элементами  $a_{ii} = \lambda_i$  и  $a_{ij} = 0$  при  $i \neq j$ .

$I = \text{diag}(1, \dots, 1)$  — единичная матрица; если необходимо указать порядок  $n$  единичной матрицы, то она обозначается  $I_n$ .

$A^T$  — транспонированная матрица;  $A^T = \|a'_{ij}\|$ , где  $a'_{ij} = a_{ji}$  для  $A = \|a_{ij}\|$ .

$\bar{A} = \|a'_{ij}\|$ , где  $a'_{ij} = \bar{a}_{ji}$  для  $A = \|a_{ij}\|$ .

$A^* = (A)^T$ .

$\sigma = \begin{pmatrix} 1 & \dots & n \\ k_1 & \dots & k_n \end{pmatrix}$  — перестановка;  $\sigma(i) = k_i$ .

$$(-1)^\sigma = \begin{cases} 1, & \text{если перестановка } \sigma \text{ чётная;} \\ -1, & \text{если перестановка } \sigma \text{ нечётная.} \end{cases}$$

Перестановка  $\sigma = \begin{pmatrix} 1 & \dots & n \\ k_1 & \dots & k_n \end{pmatrix}$  для краткости иногда обозначается

$(k_1, \dots, k_n)$ .

$\langle e_1, \dots, e_n \rangle$  — линейное пространство, порождённое векторами  $e_1, \dots, e_n$ .

Если в пространствах  $V^n$  и  $W^m$  заданы базисы  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_m$ , то матрице  $A$  соответствует линейное отображение  $A : V^n \rightarrow W^m$ ,

переводящее вектор  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  в вектор

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Так как  $y_i = \sum_{j=1}^n a_{ij}x_j$ , то  $A(\sum_{j=1}^n x_j e_j) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}x_j \varepsilon_i$ ; в частности,  $Ae_j = \sum_i a_{ij} \varepsilon_i$ .

$\text{rk } A$  — ранг матрицы  $A$ .

$\text{tr } A$  — след матрицы  $A$ .

$\delta_{ij} = 0$  при  $i \neq j$ ,  $\delta_{ij} = 1$  при  $i = j$ .

$\text{Hom}(V, W)$  — линейное пространство, состоящее из линейных отображений пространства  $V$  в пространство  $W$  (сложение двух линейных отображений и умножение линейного отображения на число определяются естественным образом).

$(x_1, \dots, \hat{x}_i, \dots, x_n) = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ .

$$\text{sign } x = \begin{cases} 1, & \text{если } x > 0; \\ 0, & \text{если } x = 0; \\ -1, & \text{если } x < 0. \end{cases}$$

$s_\lambda$  — функция Шура.

$(\alpha_1, \dots, \alpha_k | \beta_1, \dots, \beta_k)$  — обозначение Фробениуса для разбиений (см. с. 73).

$S^q B$  — симметрическая степень оператора  $B$  (см. с. 292).

$\Lambda^q B$  — внешняя степень оператора  $B$  (см. с. 292).

$A \twoheadrightarrow B$  — оператор  $B$  является репликой оператора  $A$  (см. с. 376).



# Глава 1.

## Определители

### Историческая справка

Бертран Рассел в книге «История западной философии» писал о Лейбнице: «Его лучшая мысль не доставила бы ему популярности, и он оставил свои рукописи, в которых излагалась эта мысль, неопубликованными. А то, что он опубликовал, имело целью заслужить одобрение государей и государынь. Следствием этого является то, что есть две системы философии, каждую из которых можно рассматривать как представляющую взгляды Лейбница: одна, которую он открыто провозглашал, была оптимистичной, ортодоксальной, фантастичной и мелкой; другая, которую извлекли из его рукописей относительно недавние издатели, была глубокой, ясной, во многом сходной с философией Спинозы и удивительно логичной».

Лейбниц оставил неопубликованными и свои работы, в которых он пришёл к понятию определителя. Впервые это понятие появилось у него в 1678 году; при этом он естественным образом пришёл к необходимости снабжать коэффициенты линейных уравнений двумя индексами. В 1693 году Лейбниц писал Лопиталю о пользе применения коэффициентов вида  $a_{10}$ ,  $a_{11}$  и т.д. (коэффициент  $a_{10}$  Лейбниц обозначал 10 и называл его числом, чтобы отличить от буквы, т.е. коэффициента без индекса): «... Раз Вы говорите, что Вам трудно поверить, что пользование числами носит столь же общий характер и столь же удобно, как и пользование буквами, значит, я нехорошо выразил свою мысль. Нельзя сомневаться в общности, если принять во внимание, что можно пользоваться 2, 3 и т.д. так же, как  $a$  или  $b$ , если только иметь в виду, что это не настоящие числа. Так,  $2 \cdot 3$  означает вовсе не 6, но то же, что  $ab$ . Что ка-

сается удобства, то оно очень велико, и поэтому я часто пользуюсь этим, особенно в длинных и трудных вычислениях, в которых легко ошибиться. Ибо, кроме удобства проверки с помощью чисел, а также отбрасывания девяток, я нахожу в этом очень важное преимущество даже для развития Анализа. Так как это открытие довольно своеобразно, я ещё не рассказывал о нём другим. Но вот в чём дело. Ведь верно, что когда приходится употреблять много букв, то эти буквы совсем не выражают отношений между обозначаемыми ими величинами; между тем, пользуясь числами, я могу это отношение выразить. Допустим, например, что предложены три простых уравнения с двумя неизвестными и требуется исключить эти два неизвестных, причём по общему правилу. Я полагаю  $10 + 11x + 12y = 0$  (1) и  $20 + 21x + 22y = 0$  (2) и  $30 + 31x + 32y = 0$  (3), где предлагаемые числа выражены каждое двумя знаками, из которых первый показывает мне, какому уравнению принадлежит число, а второй показывает, какой оно принадлежит букве. Вычисляя таким образом, везде открываешь гармонии, которые не только служат нам порукой, но и сразу показывают нам правила или теоремы. Например, исключая сперва из первого и второго уравнений  $y$ , мы получим

$$\begin{aligned} &+10 \cdot 22 + 11 \cdot 22x \\ &= 0, \\ &-12 \cdot 20 - 12 \cdot 21x \end{aligned} \quad (4)$$

а исключая его из первого и третьего, мы получим

$$\begin{aligned} &+10 \cdot 32 + 11 \cdot 32x \\ &= 0, \\ &-12 \cdot 30 - 12 \cdot 31x \end{aligned} \quad (5)$$

где легко заметить, что эти два уравнения отличаются лишь тем, что предыдущий знак 2 заменяется на предыдущий знак 3. Впрочем, в каждом члене каждого уравнения предыдущие знаки одинаковы, а последующие знаки образуют одну и ту же сумму. Теперь остаётся исключить из четвертого и пятого уравнений букву  $x$ , и тогда мы получим

$$\begin{array}{ccc} 10 & 21 & 32 \\ 11 & 22 & 30 \end{array} = \begin{array}{ccc} 10 & 22 & 31 \\ 11 & 20 & 32 \\ 12 & 20 & 31 \end{array},$$

что и представляет собой последнее уравнение, свободное от обоих неизвестных, которые желали исключить, и заключающее своё доказательство в самом себе, в силу заметных во всём гармоний, которые было бы весьма трудно открыть при употреблении букв  $a$ ,  $b$ ,  $c$ , особенно,

когда число букв и уравнений велико. Часть секрета Анализа состоит в характеристике, т.е. в искусстве хорошо употреблять применяемые знаки, и по этому малому образцу Вы видите, Сударь, что Виет и Декарт ещё не познали все его тайны. Незначительно продолжив это вычисление, можно прийти к общей теореме для любого произвольного числа букв и простых уравнений». Полученное Лейбницем условие совместности трёх линейных уравнений с двумя неизвестными можно записать в виде

$$\begin{vmatrix} a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \\ a_{30} & a_{31} & a_{32} \end{vmatrix} = 0.$$

Открытие Лейбница значительно опередило свое время. Первая публикация, посвященная определителям, появилась лишь через 72 года, в 1750 году; она принадлежит Г. Крамеру.

Почти одновременно с Лейбницем, в 1683 году, к понятию определителя пришел японский математик Секи Кова Шинсуке. Он усовершенствовал «фан-чэн» — метод решения системы  $n$  линейных уравнений с  $n$  неизвестными, содержащийся в древнекитайской «Математике в девяти книгах», датируемой между III в. до н. э. и I в. н. э.

## 1.1. Вычисление определителей

Методы вычисления определителей весьма разнообразны. В этом параграфе мы приведём некоторые наиболее важные из них на примере вычислений конкретных определителей. Начнём с того, что напомним, что такое определитель и каковы его основные свойства.

### 1.1.1. Определение

*Определитель*, или *детерминант*, — это функция, которая каждой квадратной матрице  $A$  сопоставляет число  $|A| = \det A$ . Эта функция характеризуется следующими тремя свойствами:

(a) Определитель линейно зависит от каждого столбца, т.е.

$$\begin{vmatrix} a_{11} & \dots & \lambda\alpha_{1i} + \mu\beta_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \lambda\alpha_{ni} + \mu\beta_{ni} & \dots & a_{nn} \end{vmatrix} = \\ = \lambda \begin{vmatrix} a_{11} & \dots & \alpha_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \alpha_{ni} & \dots & a_{nn} \end{vmatrix} + \mu \begin{vmatrix} a_{11} & \dots & \beta_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \beta_{ni} & \dots & a_{nn} \end{vmatrix}.$$

(b) Определитель матрицы с двумя одинаковыми столбцами равен 0.

(c) Определитель единичной матрицы  $I_n$  равен 1.

Нужно доказать, что такая функция  $\det$  существует и единственна. Для этого нам понадобится ещё одно свойство определителя:

(b') При перестановке двух столбцов определитель меняет знак.

Чтобы доказать это свойство, фиксируем в матрице все столбцы, кроме двух столбцов с номерами  $i$  и  $j$ , и будем рассматривать определитель матрицы как функцию  $f(u, v)$  этих двух столбцов. Тогда из свойств (a) и (b) следует, что  $0 = f(u + v, u + v) = f(u, u) + f(v, u) + f(u, v) + f(v, v) = f(v, u) + f(u, v)$ .

**Замечание.** Для матриц с элементами из поля характеристики 2 свойство (b) не следует из свойства (b').

Свойства (a), (b) и (c) позволяют вычислить определитель матрицы. Покажем это на примере матриц порядка 2. Запишем первый столбец  $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$  в виде  $\begin{pmatrix} a_{11} \cdot 1 + 0 \\ 0 + a_{21} \cdot 1 \end{pmatrix}$  и воспользуемся свойством (a):

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \begin{vmatrix} 1 & a_{12} \\ 0 & a_{22} \end{vmatrix} + a_{21} \begin{vmatrix} 0 & a_{12} \\ 1 & a_{22} \end{vmatrix}.$$

Затем в каждой из полученных матриц сделаем то же самое со вторым столбцом:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{12} \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} + a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{21}a_{12} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + a_{21}a_{22} \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix}.$$

Согласно свойству (b) определитель матрицы с одинаковыми столбцами равен 0, поэтому остается лишь вычислить определители  $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$  и  $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ .

Для этого воспользуемся свойствами (b') и (c). В результате получим

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Точно такие же рассуждения можно применить для вычисления определителя матрицы порядка  $n$ . В этом случае свойство (a) нужно применять для линейной комбинации не двух столбцов, а  $n$  столбцов. (Такое расширение свойства (a) доказывается очевидной индукцией по  $n$ .) В результате получаем

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma} (-1)^{\sigma} a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n},$$

где суммирование ведётся по всем подстановкам  $\sigma$ .

Мы доказали единственность функции  $\det$ . Докажем теперь, что функция  $\det$ , обладающая свойствами (a), (b) и (c), существует. Применим индукцию по порядку  $n$  матрицы  $A$ . При  $n = 1$  полагаем  $\det(a) = a$ . Если функция  $\det$  существует для матриц порядка  $n - 1$ , то для матриц порядка  $n$  мы определим её посредством формулы

$$\det A = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det A_{1j}, \quad (1.1)$$

где матрица  $A_{1j}$  получается из матрицы  $A$  вычёркиванием первой строки и  $j$ -го столбца. Формулу (1.1) называют *разложением определителя по первой строке*. Покажем, что функция  $\det$ , заданная формулой (1.1), обладает свойствами (a), (b) и (c).

Свойство (a) достаточно проверить для каждого члена  $a_{1j} \det A_{1j}$ . Если  $i \neq j$ , то  $a_{1j} = \text{const}$ , а  $\det A_{1j}$  линейно зависит от  $i$ -го столбца матрицы  $A$ . Если же  $i = j$ , то  $a_{1j}$  линейно зависит от  $i$ -го столбца, а  $\det A_{1j} = \text{const}$ .

Докажем теперь свойство (b). Предположим, что у матрицы  $A$  столбцы с номерами  $i$  и  $j$  одинаковые. Тогда

$$\det A = (-1)^{1+i} a_{1i} \det A_{1i} + (-1)^{1+j} a_{1j} \det A_{1j},$$

поскольку при  $k \neq i, j$  матрица  $A_{1k}$  имеет два одинаковых столбца. Пусть для определённости  $i < j$ . Тогда матрица  $A_{1i}$  получается из матрицы  $A_{1j}$  последовательной перестановкой столбцов  $i$  и  $i+1$ ,  $i+1$  и  $i+2$ , ...,  $j-2$  и  $j-1$ . Поэтому  $\det A_{1i} = (-1)^{j-i-1} \det A_{1j}$ , а значит,  $\det A = 0$ .

Свойство (c) очевидно:  $\det I_n = 1 \cdot \det I_{n-1} = 1$ .

**Замечание.** Определитель можно раскладывать не только по первой строке, но и по произвольной строке. Формула разложения определителя по  $i$ -й строке имеет вид

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij},$$

где матрица  $A_{ij}$  получается из матрицы  $A$  вычёркиванием  $i$ -й строки и  $j$ -го столбца.

### 1.1.2. Некоторые свойства определителя

**Теорема 1.1.2.1.**  $\det(A^T) = \det A$ .

**Доказательство.** Пусть  $A^T = B = (b_{ij})$ , т.е.  $b_{ij} = a_{ji}$ . Тогда

$$\begin{aligned} \det(A^T) &= \sum_{\sigma} (-1)^{\sigma} b_{\sigma(1)1} \dots b_{\sigma(n)n} = \\ &= \sum_{\sigma} (-1)^{\sigma} b_{1\sigma^{-1}(1)} \dots b_{n\sigma^{-1}(n)} = \\ &= \sum_{\sigma} (-1)^{\sigma} a_{\sigma^{-1}(1)1} \dots a_{\sigma^{-1}(n)n} = \det A, \end{aligned}$$

поскольку  $(-1)^{\sigma} = (-1)^{\sigma^{-1}}$ .  $\square$

**Замечание.** Теорема 1.1.2.1 показывает, что при вычислении определителя строки и столбцы матрицы играют одинаковые роли. Например, свойства (а) и (б) выполняются не только для столбцов, но и для строк.

**Теорема 1.1.2.2.** Пусть  $A$  и  $B$  — квадратные матрицы. Тогда определитель матрицы  $\begin{pmatrix} A & X \\ 0 & B \end{pmatrix}$  равен произведению определителей матриц  $A$  и  $B$ .

**Доказательство.** Фиксируем матрицы  $X$  и  $B$  и рассмотрим функцию  $f(A) = \det \begin{pmatrix} A & X \\ 0 & B \end{pmatrix}$ . Эта функция обладает свойствами (а) и (б), поэтому  $f(A) = (\det A)f(I_n)$ , где  $n$  — порядок матрицы  $A$ . Это доказывается точно так же, как мы доказывали единственность определителя.

Легко видеть, что  $\begin{vmatrix} I_n & X \\ 0 & B \end{vmatrix} = \begin{vmatrix} I_n & 0 \\ 0 & B \end{vmatrix}$ . Действительно, из свойств (а) и (б) следует, что определитель матрицы не изменяется, если к её столбцу добавить линейную комбинацию других столбцов. Добавляя линейные

комбинации первых  $n$  столбцов матрицы  $\begin{pmatrix} I_n & X \\ 0 & B \end{pmatrix}$  к остальным столбцам, можно получить матрицу  $\begin{pmatrix} I_n & 0 \\ 0 & B \end{pmatrix}$ .

Функция  $g(B) = \begin{vmatrix} I_n & 0 \\ 0 & B \end{vmatrix}$  обладает свойствами (а), (b) и (с), поэтому  $g(B) = \det B$ . В итоге получаем, что  $f(A) = (\det A)f(I_n) = (\det A)g(B) = (\det A)(\det B)$ .  $\square$

### 1.1.3. Определитель произведения матриц

Пусть  $A = \|a_{ij}\|_1^n$  и  $B = \|b_{ij}\|_1^n$ . Произведение  $AB$  матриц  $A$  и  $B$  — это матрица  $C = \|c_{ij}\|_1^n$ , где  $c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$  ( $i$ -я строка умножается на  $k$ -й столбец).

**Теорема 1.1.3.1.** Если  $A$  и  $B$  — матрицы порядка  $n$ , то  $\det(AB) = (\det A)(\det B)$ .

**Первое доказательство.** Фиксируем матрицу  $A$  и рассмотрим функцию  $f(B) = \det(AB)$ . Если в матрице  $B$  столбцы с номерами  $i$  и  $j$  одинаковые, то в матрице  $AB$  столбцы с номерами  $i$  и  $j$  тоже одинаковые. Поэтому функция  $f(B)$  обладает свойством (а). Если в матрице  $B$  столбец с номером  $i$  имеет вид  $\lambda a + \mu b$ , то в матрице  $AB$  столбец с номером  $i$  имеет вид  $\lambda Aa + \mu Ab$ . Поэтому функция  $f(B)$  обладает свойством (b). Следовательно,  $f(B) = (\det B)f(I_n)$ . Но  $f(I_n) = \det(AI_n) = \det A$ .  $\square$

**Второе доказательство.** Рассмотрим матрицу  $C = \begin{pmatrix} B & -I_n \\ 0 & A \end{pmatrix}$ . Согласно теореме 1.1.2.2 определитель матрицы  $C$  равен  $(\det B)(\det A)$ . Покажем теперь, что определитель матрицы  $C$  равен  $\det(AB)$ . Прежде всего заметим, что, прибавляя к последним  $n$  строкам матрицы  $C$  линейные комбинации её первых  $n$  строк, матрицу  $C$  можно заменить на матрицу  $\begin{pmatrix} B & -I_n \\ AB & 0 \end{pmatrix}$ ; при таких преобразованиях определитель не меняется. Перестановками столбцов полученную матрицу можно преобразовать в матрицу  $C' = \begin{pmatrix} -I_n & B \\ 0 & AB \end{pmatrix}$ ; при этом определитель умножится на  $(-1)^n$ . Снова применив теорему 1.1.2.2, получим, что  $\det C' = \det(-I_n) \det(AB) = (-1)^n \det(AB)$ .  $\square$

### 1.1.4. Правило Крамера

Прежде чем перейти непосредственно к примерам вычислений различных определителей, докажем *правило Крамера*, которое как раз и появилось в первой опубликованной работе об определителях.

Систему линейных уравнений

$$x_1 a_{i1} + \dots + x_n a_{in} = b_i, \quad i = 1, \dots, n,$$

можно записать в виде одного матричного уравнения  $x_1 A_1 + \dots + x_n A_n = B$ , где  $A_j$  —  $j$ -й столбец матрицы  $A = \|a_{ij}\|_1^n$ , а  $B$  — столбец свободных членов  $b_i$ .

**Теорема 1.1.4.1 (Крамер).** Для  $i = 1, \dots, n$  имеет место равенство

$$x_i \det(A_1, \dots, A_n) = \det(A_1, \dots, A_{i-1}, B, A_{i+1}, \dots, A_n).$$

**Доказательство.** Ясно, что

$$\begin{aligned} \det(A_1, \dots, B, \dots, A_n) &= \det(A_1, \dots, \sum x_j A_j, \dots, A_n) = \\ &= \sum x_j \det(A_1, \dots, A_j, \dots, A_n) = \\ &= x_i \det(A_1, \dots, A_i, \dots, A_n), \end{aligned}$$

поскольку при  $j \neq i$  получается определитель, содержащий два одинаковых столбца.  $\square$

Определитель  $\det(A_1, \dots, A_n)$  называют *определителем данной системы линейных уравнений*. В случае, когда этот определитель не равен 0, формула Крамера позволяет находить решение системы линейных уравнений.

### 1.1.5. Определитель Вандермонда

Один из наиболее часто встречающихся определителей — это *определитель Вандермонда*

$$V(x_1, \dots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}.$$

**Теорема 1.1.5.1.**  $V(x_1, \dots, x_n) = \prod_{i>j} (x_i - x_j)$ .



**Первое доказательство.** Применим индукцию по  $n$ . При  $n = 2$  равенство  $V(x_1, x_2) = x_2 - x_1$  очевидно. Вычтем в матрице Вандермонда из последнего столбца предпоследний столбец, домноженный на  $x_1$ , затем из предпоследнего столбца вычтем предшествующий ему столбец, домноженный на  $x_1$ , и т.д. В результате в первой строке все элементы, кроме первого, станут нулями. Таким образом, вычисление определителя Вандермонда порядка  $n$  сводится к вычислению некоторого определителя порядка  $n - 1$ . В этом новом определителе для каждой строки можно вынести множитель  $x_i - x_1$ , где  $i$  — номер строки. В результате получим

$$V(x_1, \dots, x_n) = \prod_{i>1} (x_i - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}.$$

Это доказывает шаг индукции.  $\square$

**Второе доказательство.** Определитель Вандермонда является многочленом от  $x_1, \dots, x_n$ , степень которого не превосходит  $1 + 2 + \dots + n - 1 = \frac{n(n-1)}{2}$ . Если  $x_i = x_j$  при  $i \neq j$ , то мы получаем определитель матрицы с двумя одинаковыми строками; он равен нулю. Следовательно, определитель Вандермонда делится на многочлен  $\prod_{i>j} (x_i - x_j)$ , степень которого равна  $\frac{n(n-1)}{2}$ . Следовательно, определитель Вандермонда равен  $C \prod_{i>j} (x_i - x_j)$ , где  $C$  — некоторая константа. Легко видеть, что коэффициенты при мономе  $x_2 x_3^2 \dots x_n^{n-1}$  в определителе Вандермонда и в многочлене  $\prod_{i>j} (x_i - x_j)$  равны 1, поэтому  $C = 1$ .  $\square$

Большинство применений определителя Вандермонда связано с тем, что  $V(x_1, \dots, x_n) = 0$  тогда и только тогда, когда среди чисел  $x_1, \dots, x_n$  есть два равных.

### 1.1.6. Интерполяционный многочлен Лагранжа

Определитель Вандермонда возникает, например, при решении следующей задачи. Пусть заданы два набора чисел  $x_1, \dots, x_{n+1}$  и  $a_1, \dots, a_{n+1}$ . Требуется найти многочлен  $P(x)$  степени не выше  $n$ , для которого  $P(x_i) = a_i$  при  $i = 1, \dots, n+1$ . Запишем искомый многочлен в виде  $P(x) = \xi_0 + \xi_1 x + \xi_2 x^2 + \dots + \xi_n x^n$ . Для коэффициентов  $\xi_0, \xi_1, \dots, \xi_n$  мы получаем  $n+1$  линейное уравнение  $\xi_0 + \xi_1 x_i + \xi_2 x_i^2 + \dots + \xi_n x_i^n = a_i$ ,  $i = 1, \dots, n+1$ . Определитель этой системы уравнений — определитель Вандермонда  $V(x_1, \dots, x_{n+1})$ . Поэтому если числа  $x_1, \dots, x_{n+1}$  попарно

различны, то эта система уравнений имеет единственное решение. Соответствующий многочлен  $P(x)$  называют *интерполяционным многочленом Лагранжа*.

Вычислим интерполяционный многочлен Лагранжа при  $n = 2$ . Сначала рассмотрим случай  $(a_1, a_2, a_3) = (1, 0, 0)$ . Решая в этом случае систему уравнений по правилу Крамера, легко находим многочлен

$$P_1(x) = \frac{x_2 x_3 - (x_2 + x_3)x + x^2}{(x_2 - x_1)(x_3 - x_1)} = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}.$$

Аналогично находятся интерполяционные многочлены  $P_2(x)$  и  $P_3(x)$  для наборов  $(0, 1, 0)$  и  $(0, 0, 1)$ . Ясно также, что интерполяционный многочлен для набора  $(a_1, a_2, a_3)$  равен  $a_1 P_1(x) + a_2 P_2(x) + a_3 P_3(x)$ . Таким образом, интерполяционный многочлен Лагранжа для набора  $(a_1, a_2, a_3)$  равен

$$\frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} a_1 + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} a_2 + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} a_3.$$

Теперь ясно, что для произвольного  $n$  интерполяционный многочлен Лагранжа имеет вид

$$\sum_{k=1}^{n+1} a_k \frac{(x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_{n+1})}{(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_{n+1})}.$$

**Замечание.** Мы решили систему линейных уравнений при  $n = 2$ , а потом угадали решение для произвольного  $n$ . Так мы обошлись без вычисления определителей, которые возникают при решении системы линейных уравнений по правилу Крамера. Непосредственные вычисления этих определителей приведены в решении задачи 1.17.

### 1.1.7. Определитель, сводящийся к определителю Вандермонда

К вычислению определителя Вандермонда сводится вычисление многих других определителей. Здесь мы приведём один такой пример.

**Теорема 1.1.7.1.** Пусть  $x$  и  $h$  — фиксированные числа,  $f(x) = a_0 t^n + a_1 t^{n-1} + \cdots + a_n$  и  $c_{ij} = f(t)$ , где  $t = x + h(i + j)$ . Тогда

$$|c_{ij}|_0^n = (-h^2)^{n(n+1)/2} (n! a_0)^{n+1}.$$

**Доказательство.** Для наглядности доказательство проведём при  $n = 2$  и укажем, какие изменения нужно сделать в общем случае. Разложение в ряд Тэйлора позволяет записать матрицу  $\|c_{ij}\|_0^n$  в виде произведения двух матриц

$$\begin{pmatrix} f''(x)/2 & f'(x) & f(x) \\ f''(x+h)/2 & f'(x+h) & f(x+h) \\ f''(x+2h)/2 & f'(x+2h) & f(x+2h) \end{pmatrix} \begin{pmatrix} 0 & h^2 & (2h)^2 \\ 0 & h & 2h \\ 1 & 1 & 1 \end{pmatrix}.$$

Первую из этих матриц тоже можно представить в виде произведения двух матриц

$$\begin{pmatrix} 1 & x & x^2 \\ 1 & x+h & (x+h)^2 \\ 1 & x+2h & (x+2h)^2 \end{pmatrix} \begin{pmatrix} a_0 & a_1 & a_2 \\ 0 & 2a_0 & a_1 \\ 0 & 0 & a_0 \end{pmatrix}.$$

В общем случае элементы второй матрицы имеют вид  $a_{ij} = a_{j-i} \binom{n+j-i}{i}$ . Для диагональных элементов (а нас интересуют только они) это следует из того, что  $\frac{1}{k!} (a_0 x^n)^{(k)} = \binom{n}{k} a_0 x^{n-k}$ .

Таким образом,

$$\begin{aligned} |c_{ij}|_0^n &= a_0^{n-1} \left( \binom{n}{0} \binom{n}{1} \cdots \binom{n}{n} \right) V(nh, (n-1)h, \dots, 0) \times \\ &\times V(x, x+h, \dots, x+nh) = \\ &= \frac{a_0^{n+1} (n!)^{n+1}}{(n!(n-1)! \cdots 1!)^2} \prod_{i>k} h^2 (k-i)(i-k). \end{aligned}$$

Остаётся заметить, что  $n!(n-1)! \cdots 1! = \prod_{0 \leq k < i \leq n} (i-k)$ .  $\square$

### 1.1.8. Определитель Коши

Определитель матрицы  $\|a_{ij}\|_1^n$ , где  $a_{ij} = \frac{1}{x_i + y_j}$ , называют *определителем Коши*.

**Теорема 1.1.8.1.** *Определитель Коши равен*

$$\prod_{i>j} (x_i - x_j)(y_i - y_j) / \prod_{i,j} (x_i + y_j).$$

**Первое доказательство.** Применим индукцию по  $n$ . При  $n = 1$  утверждение очевидно. Доказательство шага индукции проведём в два этапа.

Сначала вычтем последний столбец из всех остальных. В результате получим матрицу с элементами

$$a'_{ij} = \frac{1}{x_i + y_j} - \frac{1}{x_i + y_n} = \frac{y_n - y_j}{(x_i + y_j)(x_i + y_n)}, \quad j \neq n.$$

Из всех строк вынесем множители  $\frac{1}{x_i + y_n}$ , а из всех столбцов, кроме последнего, — множители  $y_n - y_j$ . В результате перейдём к определителю  $|b_{ij}|_1^n$ , где  $b_{ij} = a_{ij}$  при  $j \neq n$  и  $b_{in} = 1$ . Чтобы вычислить этот определитель, вычтем последнюю строку из всех остальных. Вынося из всех строк, кроме последней, множители  $x_n - x_i$ , а из всех столбцов, кроме последнего, — множители  $\frac{1}{x_n + y_j}$ , мы перейдём к определителю Коши  $|a_{ij}|_1^{n-1}$  меньшего размера.  $\square$

**Второе доказательство.** Определитель Коши представляет собой сумму  $n!$  слагаемых вида

$$P(x_1, \dots, x_n, y_1, \dots, y_n) / \prod_{i,j} (x_i + y_j),$$

где  $P$  — многочлен степени  $n(n-1)$ . Поэтому определитель Коши тоже имеет такой вид, причём  $P$  — многочлен степени не выше  $n(n-1)$ .

Если  $x_i = x_j$  или  $y_i = y_j$  при  $i \neq j$ , то мы получаем определитель с двумя равными строками или столбцами. Поэтому многочлен  $P$  делится на многочлен  $\prod_{i>j} (x_i - x_j)(y_i - y_j)$ , степень которого равна  $n(n-1)$ . Следовательно,  $P = \alpha_n \prod_{i>j} (x_i - x_j)(y_i - y_j)$ , где  $\alpha_n$  — константа.

Остаётся доказать, что  $\alpha_n = 1$  для всех  $n$ . При  $n = 1$  это очевидно. Поэтому достаточно проверить, что  $\alpha_n = \alpha_{n-1}$ . Фиксируем  $x_1, \dots, x_{n-1}, y_1, \dots, y_{n-1}$  и положим  $x_n = t$  и  $y_n = t^2$ . Умножим последний столбец матрицы  $A_n = \|a_{ij}\|_1^n$  на  $y_n$ . При  $t \rightarrow \infty$  полученная матрица стремится

к матрице  $\begin{pmatrix} & & 1 \\ & A_{n-1} & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}$ , поэтому её определитель стремится к

$|a_{ij}|_1^{n-1}$ . С другой стороны, отношение определителей этих матриц равно

$$\frac{\alpha_n t^2 \prod_{k=1}^{n-1} (t - x_k)(t^2 - y_k)}{\alpha_{n-1} (t^2 + t) \prod_{k=1}^{n-1} (t + x_k)(t^2 + y_k)} \rightarrow \frac{\alpha_n}{\alpha_{n-1}}.$$

$\square$

### 1.1.9. Матрица Фробениуса

Матрица  $A = \begin{pmatrix} 0 & & & \\ \vdots & & I_{n-1} & \\ 0 & & & \\ a_0 & a_1 & \dots & a_{n-1} \end{pmatrix}$  имеет много разных назва-

ний: *матрица Фробениуса, циклическая матрица, сопровождающая матрица* многочлена  $P(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - \dots - a_0$ . Применив индукцию по  $n$  и воспользовавшись разложением по первой строке, легко проверить, что  $\det(A - \lambda I) = (-1)^n(\lambda^n - a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - \dots - a_0) = (-1)^n P(\lambda)$ .

Если  $a_0 = 0$ , то  $\det A = 0$ . Если же  $a_0 \neq 0$ , то, как легко проверить, матрица

$$\begin{pmatrix} -\frac{a_1}{a_0} & -\frac{a_2}{a_0} & \dots & -\frac{a_{n-1}}{a_0} & \frac{1}{a_0} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

является обратной для матрицы  $A$ .

Легко проверить, что если  $p(\lambda_0) = 0$  и  $x$  — столбец  $(1, \lambda_0, \lambda_0^2, \dots, \lambda_0^{n-1})$ , то  $Ax = \lambda_0 x$ . Это означает, что  $x$  — собственный вектор матрицы  $A$ , соответствующий собственному значению  $\lambda_0$ .

### 1.1.10. Циркулянт

Матрицу  $\|a_{ij}\|_1^n$  называют *циркулянт*ом, если  $a_{ij} = b_{i-j}$  для некоторой последовательности чисел  $b_i$ , обладающих тем свойством, что  $b_k = b_i$  при  $k \equiv i \pmod{n}$ . При этом многочлен  $f(x) = b_0 + b_1x + b_2x^2 + \dots + b_{n-1}x^{n-1}$  называют *вспомогательным многочленом* для циркулянта.

**Теорема 1.1.10.1.** Пусть  $\varepsilon_1, \dots, \varepsilon_n$  — попарно различные корни степени  $n$  из единицы. Тогда определитель циркулянта, соответствующего последовательности  $b_0, b_1, \dots$ , равен  $f(\varepsilon_1)f(\varepsilon_2)\dots f(\varepsilon_n)$ , где  $f$  — вспомогательный многочлен.

**Доказательство.** Легко проверить, что при  $n = 3$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & \varepsilon_1 & \varepsilon_1^2 \\ 1 & \varepsilon_2 & \varepsilon_2^2 \end{pmatrix} \begin{pmatrix} b_0 & b_2 & b_1 \\ b_1 & b_0 & b_2 \\ b_2 & b_1 & b_0 \end{pmatrix} = \\ = \begin{pmatrix} f(1) & f(1) & f(1) \\ f(1) & f(\varepsilon_1) & f(\varepsilon_1^2) \\ f(1) & f(\varepsilon_2) & f(\varepsilon_2^2) \end{pmatrix} = f(1)f(\varepsilon_1)f(\varepsilon_1^2) \begin{pmatrix} 1 & 1 & 1 \\ 1 & \varepsilon_1 & \varepsilon_1^2 \\ 1 & \varepsilon_2 & \varepsilon_2^2 \end{pmatrix}.$$

Поэтому  $V(1, \varepsilon_1, \varepsilon_2)|a_{ij}|_1^3 = f(1)f(\varepsilon_1)f(\varepsilon_1^2)V(1, \varepsilon_1, \varepsilon_2)$ . Но определитель Вандермонда  $V(1, \varepsilon_1, \varepsilon_2)$  отличен от нуля, поэтому  $|a_{ij}|_1^3 = f(1)f(\varepsilon_1)f(\varepsilon_1^2)$ . В общем случае доказательство аналогично.  $\square$

**Замечание.** Справедливо даже более сильное утверждение: числа  $f(\varepsilon_1)$ ,  $\dots$ ,  $f(\varepsilon_n)$  являются собственными значениями циркулянта (теорема 3.1.5.2).

Рассмотрим циркулянт  $C_n$ , у которого  $c_{2,1} = c_{3,2} = \dots = c_{n,n-1} = c_{1,n} = 1$ , а все остальные элементы нулевые. Легко видеть, что матрица  $A$  порядка  $n$  является циркулянтом тогда и только тогда, когда  $A = b_0I + b_1C_n + b_2C_n^2 + \dots + b_{n-1}C_n^{n-1}$ .

**Теорема 1.1.10.2.** Матрица  $\|a_{ij}\|_1^n$  является циркулянтом тогда и только тогда, когда  $AC_n = C_nA$ .

**Доказательство.** Ясно, что если  $A = \sum b_k C_n^k$ , то  $AC_n = C_nA$ . Предположим теперь, что  $AC_n = C_nA$ . Пусть  $e_1$  — столбец  $(1, 0, \dots, 0)$ . Тогда

$$(A - a_{11}I - a_{21}C_n - \dots - a_{n1}C_n^{n-1})C_n^k e_1 = C_n^k (A - a_{11}I - a_{21}C_n - \dots - a_{n1}C_n^{n-1})e_1 = 0$$

для всех  $k$ . Из этого следует, что  $A = a_{11}I + a_{21}C_n + \dots + a_{n1}C_n^{n-1}$ , поскольку  $e_1, C_n e_1, \dots, C_n^{n-1} e_1$  — это столбцы  $(1, 0, \dots, 0)$ ,  $(0, 1, 0, \dots, 0)$ ,  $\dots$ ,  $(0, \dots, 0, 1)$ .  $\square$

**Теорема 1.1.10.3.** Вспомогательный многочлен произведения двух циркулянтов порядка  $n$  по модулю многочлена  $x^n - 1$  равен произведению их вспомогательных многочленов.

**Доказательство.** Циркулянтам  $\sum_{k=0}^{n-1} a_k C_n^k$  и  $\sum_{k=0}^{n-1} b_k C_n^k$  соответствуют вспомогательные многочлены  $\sum_{k=0}^{n-1} a_k x^k$  и  $\sum_{k=0}^{n-1} b_k x^k$ . При вычислении произведения циркулянтов нужно каждый раз заменять матрицу  $C_n^{n+k}$  на матрицу  $C_n^k$ . Для многочленов это соответствует приведению по модулю многочлена  $x^n - 1$ .  $\square$

Используя свойства циркулянта, можно получить обобщение так называемой теоремы Наполеона. Напомним, что *теоремой Наполеона* называют следующее утверждение: «Центры правильных треугольников, построенных внешним (соответственно, внутренним) образом на сторонах произвольного треугольника, образуют правильный треугольник.»

Рассмотрим на комплексной плоскости  $\mathbb{C}$  многоугольник  $A_1 \dots A_n$  (не обязательно выпуклый). Пусть  $\varepsilon_k = e^{2k\pi i/n}$  — корень степени  $n$  из единицы. Нас будут интересовать только корни  $\varepsilon_1, \dots, \varepsilon_{n-1}$ , отличные от 1. Сопоставим корню  $\varepsilon_k$  число  $c_k = \frac{1}{1-\varepsilon_k}$ . Фиксировав  $k$ , многоугольнику  $A_1 \dots A_n$  можно сопоставить многоугольник  $B_1 \dots B_n$ , где  $B_j = (1 - c_k)A_j + c_k A_{j+1}$ . Несложные тригонометрические преобразования показывают, что  $c_k = \frac{1}{2} \left(1 + i \operatorname{ctg} \frac{k\pi}{n}\right)$ , поэтому  $B_j$  — вершина равнобедренного треугольника с основанием  $A_j A_{j+1}$  и углом при вершине  $2k\pi/n$ .

Столбец чисел  $B_1, \dots, B_n$  получается из столбца чисел  $A_1, \dots, A_n$  умножением на циркулянт, которому соответствует вспомогательный многочлен  $1 - c_k + c_k x = \frac{x - \varepsilon_k}{1 - \varepsilon_k}$ . Поэтому если мы будем последовательно преобразовывать многоугольник  $A_1 \dots A_n$  в соответствии с числами  $c_1, \dots, c_{n-1}$  (в любом порядке), то в результате получим многоугольник, который получается при применении к столбцу чисел  $A_1, \dots, A_n$  циркулянта со вспомогательным многочленом

$$\frac{x - \varepsilon_1}{1 - \varepsilon_1} \cdot \frac{x - \varepsilon_2}{1 - \varepsilon_2} \dots \frac{x - \varepsilon_{n-1}}{1 - \varepsilon_{n-1}} = \frac{x^{n-1} + x^{n-2} + \dots + x + 1}{n}.$$

Каждый элемент такого циркулянта равен  $1/n$ , поэтому соответствующее преобразование многоугольников таково: многоугольнику  $A_1 \dots A_n$  сопоставляется его центр масс (с кратностью  $n$ ).

Мы выяснили, как устроена композиция всех  $n - 1$  преобразований, соответствующих корням степени  $n$  из единицы, отличных от 1. Ясно, что на предпоследнем шаге может получиться только многоугольник, который переходит в себя при повороте на угол  $2k\pi/n$ , где число  $k$  соответствует последнему шагу. Таким образом, если мы отбросим один из нетривиальных корней степени  $n$  из единицы и сделаем преобразования многоугольника, соответствующие всем остальным нетривиальным корням, то в результате получим правильный многоугольник или правильную звезду; число вершин полученного многоугольника — делитель числа  $n$ . Результат построений зависит только от отброшенного корня; от того порядка, в котором делаются преобразования, он не зависит (с геометрической точки зрения этот факт не очевиден).

Доказанное нами утверждение можно назвать *обобщённой теоремой Наполеона*, поскольку при  $n = 3$  мы получаем обычную теорему Наполеона.

**Теорема 1.1.10.4.** Пусть  $A = \|a_{ij}\|_1^n$  — циркулянт, для которого  $(b_1, b_2, \dots, b_n) = (1, \dots, 1, 0, \dots, 0)$  ( $k$  единиц, где  $k > 1$ ). В таком случае  $\det A = 0$  тогда и только тогда, когда числа  $k$  и  $n$  имеют общий делитель.

**Доказательство.** Пусть  $f(x) = 1 + x + x^2 + \dots + x^{k-1}$  и  $\varepsilon_k = \exp\left(\frac{2\pi pi}{n}\right)$ . Тогда  $\det A = f(1)f(\varepsilon_1) \dots f(\varepsilon_{n-1})$ . Ясно, что  $f(1) = k \neq 1$ . Поэтому  $\det A \neq 0$  тогда и только тогда, когда хотя бы одно из чисел  $f(\varepsilon_1), \dots, f(\varepsilon_{n-1})$  равно нулю.

Если  $x \neq 1$ , то  $f(x) = \frac{x^k - 1}{x - 1}$ ; в этом случае  $f(x) = 0$  тогда и только тогда, когда  $x^k = 1$ . Равенство  $\varepsilon_p^k = 1$  эквивалентно тому, что  $kp$  делится на  $n$  (здесь  $1 \leq p \leq n - 1$ ). Это возможно тогда и только тогда, когда  $k$  и  $n$  имеют общий делитель.  $\square$

### 1.1.11. Матрица Якоби

*Матрицей Якоби*, или *трёхдиагональной матрицей*, называют квадратную матрицу  $J = \|a_{ij}\|_1^n$  с элементами  $a_{ij}$ , равными нулю при  $|i - j| > 1$ . Матрица Якоби имеет вид

$$\begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ c_1 & a_2 & b_2 & \dots & 0 & 0 \\ 0 & c_2 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & c_{n-1} & a_n \end{pmatrix}.$$

Определитель матрицы Якоби можно вычислить с помощью следующего рекуррентного соотношения. Пусть  $\Delta_k = |a_{ij}|_1^k$  при  $k \geq 1$  и  $\Delta_0 = 1$ . Тогда

$$\Delta_k = a_k \Delta_{k-1} - b_{k-1} c_{k-1} \Delta_{k-2} \quad \text{при} \quad k \geq 2. \quad (1.1)$$

Это соотношение доказывается разложением определителя  $|a_{ij}|_1^k$  по последней строке.

Из рекуррентного соотношения (1.1) видно, в частности, что определитель матрицы Якоби зависит не от самих чисел  $b_i$  и  $c_i$ , а только от их произведений  $b_i c_i$ .



Величину

$$(a_1 \dots a_n) = \begin{vmatrix} a_1 & 1 & 0 & \dots & 0 & 0 \\ -1 & a_2 & 1 & \dots & 0 & 0 \\ 0 & -1 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & a_n \end{vmatrix}$$

называют *континуантой*. Это название связано с тем, что цепная дробь (по английски — continued fraction)

$$a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots + \frac{1}{a_n}}}$$

равна  $\frac{(a_1 a_2 \dots a_n)}{(a_2 a_3 \dots a_n)}$ . Докажем это утверждение по индукции. Ясно, что  $a_1 + \frac{1}{a_2} = \frac{(a_1 a_2)}{(a_2)}$ . Остаётся доказать, что

$$a_1 + \frac{1}{\frac{(a_2 a_3 \dots a_n)}{(a_3 a_4 \dots a_n)}} = \frac{(a_1 a_2 \dots a_n)}{(a_2 a_3 \dots a_n)},$$

т.е.  $a_1(a_2 \dots a_n) + (a_3 \dots a_n) = (a_1 \dots a_n)$ . Умножим справа и слева рассматриваемую матрицу на матрицу  $\|\delta_{i,n+1-j}\|_1^n$ . Требуемое соотношение следует теперь из соотношения (1.1).

### 1.1.12. Определитель Смита

*Определителем Смита* называют определитель матрицы  $\|a_{ij}\|_1^n$ , где  $a_{ij} = (i, j)$  — наибольший общий делитель чисел  $i$  и  $j$ .

**Теорема 1.1.12.1.** *Определитель Смита порядка  $n$  равен  $\varphi(1)\varphi(2)\dots\varphi(n)$ , где  $\varphi(m)$  — количество тех из чисел  $1, 2, \dots, m$ , которые взаимно просты с  $m$  (функция Эйлера).*

**Доказательство.** Для любого натурального  $n$  выполняется равенство  $n = \sum_{d|n} \varphi(d)$ , где запись  $d|n$  означает, что  $d$  делит  $n$ . Действительно, в группе  $\mathbb{Z}/n\mathbb{Z}$  есть единственная подгруппа порядка  $d$ ; в качестве образующей

этой группы можно выбрать один из  $\varphi(d)$  элементов. Кроме того, любой элемент группы  $\mathbb{Z}/n\mathbb{Z}$  порождает некоторую подгруппу.

Пусть

$$\Delta_{ab} = \begin{cases} 1, & \text{если } a \text{ делится на } b; \\ 0, & \text{если } a \text{ не делится на } b. \end{cases}$$

Тогда

$$(i, j) = \sum_{d \mid (i, j)} \varphi(d) = \sum_{d=1}^n \Delta_{id} \Delta_{jd} \varphi(d).$$

Это означает, что

$$\|(i, j)\|_1^n = \|\Delta_{id}\|_1^n (\|\Delta_{jd}\varphi(d)\|_1^n)^T.$$

Поэтому

$$|(i, j)|_1^n = (|\Delta_{id}|_1^n)^2 \varphi(1)\varphi(2) \cdots \varphi(n).$$

Остаётся заметить, что матрица  $\|\Delta_{id}\|_1^n$  треугольная с диагональю  $(1, 1, \dots, 1)$ .  $\square$

### 1.1.13. Определитель преобразованной матрицы

При умножении строки квадратной матрицы на число  $\lambda$  её определитель тоже умножается на  $\lambda$ , а при прибавлении к строке любой другой строки определитель не изменяется. Эти утверждения можно обобщить для преобразований сразу нескольких строк. Рассмотрим матрицу  $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ , где  $A_{11}$  и  $A_{21}$  — квадратные матрицы порядка  $m$  и  $n$  соответственно. Пусть  $D$  — квадратная матрица порядка  $m$ ;  $B$  — матрица размера  $m \times n$ .

**Теорема 1.1.13.1.**

$$\begin{vmatrix} DA_{11} & DA_{12} \\ A_{21} & A_{22} \end{vmatrix} = |D| \cdot |A| \text{ u } \begin{vmatrix} A_{11} & A_{12} \\ A_{21} + BA_{11} & A_{22} + BA_{12} \end{vmatrix} = |A|.$$

**Доказательство.**

$$\begin{pmatrix} DA_{11} & DA_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} D & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

и

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} + BA_{11} & A_{22} + BA_{12} \end{vmatrix} = \begin{vmatrix} I & 0 \\ B & I \end{vmatrix} \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}.$$

$\square$

### 1.1.14. Степени степенного ряда

Пусть  $f = 1 + a_1x + a_2x^2 + \dots$  — формальный степенной ряд. Произведение формальных рядов  $f$  и  $g = 1 + b_1x + b_2x^2 + \dots$  определяется следующим образом: коэффициент при  $x^k$  в нём равен  $a_k + a_{k-1}b_1 + \dots + b_k$ . Это определение позволяет определить по индукции  $i$ -ю степень  $f$ , которую мы обозначим  $f^i$ .

**Теорема 1.1.14.1.** Пусть  $c_{ij}$  — коэффициент при  $x^j$  формального ряда  $f^i$ . Тогда  $|c_{ij}|_0^n = a_1^{n(n+1)/2}$ .

**Доказательство.** [Wil] Рассмотрим вспомогательную нижнюю треугольную матрицу  $B = |b_{ij}|_0^n$ , где  $b_{ij} = (-1)^{i+j} \binom{i}{j}$ . Покажем, что  $BC$  — верхняя треугольная матрица  $D$  с элементами  $d_{ii} = a_1^i$  на диагонали. Действительно, пусть  $[x^j]f^i$  — коэффициент при  $x^j$  формального ряда  $f^i$ . Тогда

$$\begin{aligned} d_{ik} &= \sum_j b_{ij} c_{jk} = \sum_j (-1)^{i+j} \binom{i}{j} [x^k] f^j = \\ &= (-1)^i [x^k] \sum_j (-1)^j \binom{i}{j} [x^k] f^j = \\ &= (-1)^i [x^k] (1 - f)^i = [x^k] (a_1x + a_2x^2 + \dots)^i. \end{aligned}$$

Таким образом,  $d_{ii} = a_1^i$  и  $d_{ik} = 0$  при  $k < i$ .  $\square$

**Следствие 1.**  $\left| \binom{mi}{j} \right|_0^n = m^{n(n+1)/2}$ .

**Доказательство.** Пусть  $f = (1 + x)^m = 1 + mx + \dots$ . Тогда  $c_{ij}$  — коэффициент при  $x^j$  формального ряда  $f^i = (1 + x)^{mi}$ . Он равен  $\binom{mi}{j}$ . (Другое доказательство приведено в решении задачи 1.21.)  $\square$

**Следствие 2.** Пусть  $c_{ij}$  — количество представлений<sup>1</sup> целого числа  $j$  в виде суммы  $i$  квадратов целых неотрицательных чисел. Тогда  $|c_{ij}|_0^n = 1$ .

**Доказательство.** Пусть  $f = 1 + x + x^4 + x^9 + x^{16} + x^{25} + \dots$ . Тогда коэффициент при  $x^j$  формального ряда  $f^i$  равен  $c_{ij}$ .  $\square$

<sup>1</sup>Представления, отличающиеся порядком слагаемых, считаются различными.

### 1.1.15. Биномиальные определители

Пусть  $0 \leq a_1 < a_2 < \dots < a_k$  и  $0 \leq b_1 < b_2 < \dots < b_k$  — целые числа.

Биномиальным определителем называют определитель матрицы  $\left\| \binom{a_i}{b_j} \right\|_1^k$ .

Биномиальные определители имеют следующую комбинаторную интерпретацию. Рассмотрим на плоскости точки  $A_i = (0, a_i)$  и  $B_i = (b_i, b_i)$ ,  $i = 1, \dots, k$ . Будем считать, что на каждом шаге разрешено перемещаться на 1 вправо или вниз, т.е. из точки  $(x, y)$  — в точку  $(x + 1, y)$  или в точку  $(x, y - 1)$ . Пусть  $w_i$  — путь с таким свойством, идущий из  $A_i$  в  $B_i$ .

**Теорема 1.1.15.1 ([Ge]).** Биномиальный определитель  $\left| \binom{a_i}{b_j} \right|_1^k$  равен количеству наборов попарно непересекающихся путей  $(w_1, \dots, w_k)$ .

**Доказательство.** Рассмотрим множество наборов  $(\sigma, w_1, \dots, w_k)$ , где  $\sigma \in S_k$  — некоторая перестановка, а  $w_i$  — путь из точки  $A_i$  в точку  $B_{\sigma(i)}$ , обладающий указанным свойством. Ясно, что перестановка  $\sigma$  полностью определяется набором путей.

Количество путей из  $A_i$  в  $B_j$  равно  $\binom{a_i}{b_j}$ . Действительно, чтобы попасть из  $A_i$  в  $B_j$ , нужно сделать  $b_j$  шагов вправо и  $a_i - b_j$  шагов вниз. Поэтому из  $a_i$  шагов нужно выбрать те  $b_j$  шагов, которые делаются вниз. Таким образом,

$$\left| \binom{a_i}{b_j} \right|_1^k = \sum_{(\sigma, w_1, \dots, w_k)} (-1)^\sigma, \quad (1.1)$$

где суммирование ведётся по всем рассматриваемым наборам  $(\sigma, w_1, \dots, w_k)$ .

Покажем, что в сумме (1.1) взаимно сокращаются все слагаемые, кроме тех, которые соответствуют наборам путей из формулировки теоремы. Если перестановка  $\sigma$  не тождественна, то  $\sigma(i) > \sigma(j)$  для некоторых  $i, j$ , где  $1 \leq i < j \leq k$ . Легко видеть, что в таком случае пути  $w_i$  и  $w_j$  пересекаются. Поэтому нам нужно рассмотреть в точности те наборы путей, среди которых есть пересекающиеся пути. В таком наборе среди всех путей, пересекающих другие пути, можно выбрать путь  $w_{i_0}$  с наименьшим номером. Затем выберем первую точку  $C$  этого пути, через которую проходят какие-то другие пути, и среди них выберем путь  $w_{j_0}$  с наименьшим номером.

Преобразуем набор путей  $(w_1, \dots, w_k)$  следующим образом. Все пути, кроме  $w_{i_0}$  и  $w_{j_0}$ , оставим прежними. Путь  $w_{i_0}$  заменим на путь, который до точки  $C$  идёт по  $w_{i_0}$ , а после точки  $C$  — по  $w_{j_0}$ . А путь  $w_{j_0}$  заменим на путь, который до точки  $C$  идёт по  $w_{j_0}$ , а после точки  $C$  — по  $w_{i_0}$ . При

этом перестановка  $\sigma$  домножается на транспозицию  $(i, j)$ , поэтому знак перестановки изменяется. Легко видеть, что наше преобразование инволютивно, т.е., применив его дважды, мы получим исходный набор путей. Поэтому наборы  $(\sigma, w_1, \dots, w_k)$ , в которых встречаются пересекающиеся пути, разбиваются на пары так, что в каждой паре перестановки имеют разные знаки. В сумме (1.1) такие пары взаимно сокращаются.  $\square$

**Следствие 1.** Биномиальный определитель  $\left| \begin{pmatrix} a_i \\ b_j \end{pmatrix} \right|_1^k$  неотрицателен. Этот определитель положителен тогда и только тогда, когда  $b_i \leq a_i$  для  $i = 1, \dots, k$ .

**Следствие 2.** Если  $a_m = b_m$ , то

$$\left| \begin{pmatrix} a_i \\ b_j \end{pmatrix} \right|_1^k = \left| \begin{pmatrix} a_i \\ b_j \end{pmatrix} \right|_1^{m-1} \left| \begin{pmatrix} a_i \\ b_j \end{pmatrix} \right|_{m+1}^k.$$

## Задачи

**1.1.** Матрица  $A = \|a_{ij}\|_1^n$  кососимметрическая, т.е.  $a_{ij} = -a_{ji}$ . Докажите, что если  $n$  нечётно, то  $|A| = 0$ .

**1.2.** Докажите, что определитель кососимметрической матрицы чётного порядка не изменится, если ко всем её элементам прибавить одно и то же число  $x$ .

**1.3.** Вычислите определитель кососимметрической матрицы  $A_n = \|a_{ij}\|_1^{2n}$  с элементами  $a_{ij} = 1$  при  $i < j$ .

**1.4.** Докажите, что все члены разложения определителя порядка  $n$  при  $n \geq 3$  не могут быть одновременно положительными.

**1.5.** Вычислите определитель  $|a_{ij}|_1^n$ , где  $a_{ij} = a^{|i-j|}$ .

**1.6.** Пусть

$$\Delta_3 = \begin{vmatrix} 1 & -1 & 0 & 0 \\ x & h & -1 & 0 \\ x^2 & hx & h & -1 \\ x^3 & hx^2 & hx & h \end{vmatrix};$$

$\Delta_n$  определяется аналогично. Докажите, что  $\Delta_n = (x + h)^n$ .

**1.7.** Вычислите определитель  $|c_{ij}|_1^n$ , где  $c_{ij} = a_i b_j$  при  $i \neq j$  и  $c_{ii} = x_i$ .

**1.8.** Пусть  $a_{12} = c_1, a_{23} = c_2, \dots, a_{n-1,n} = c_{n-1}, a_{n1} = c_n$ , а все остальные элементы матрицы  $A = \|a_{ij}\|_1^n$  нулевые. Докажите, что определитель матрицы  $I + A + A^2 + \dots + A^{n-1}$  равен  $(1 - c)^{n-1}$ , где  $c = c_1 \cdots c_n$ .

**1.9.** Вычислите определитель  $|a_{ij}|_1^n$ , где  $a_{ij} = \frac{1}{1 - x_i y_j}$ .

**1.10.** Пусть  $a_{ij} = \binom{n+i}{j}$ . Докажите, что  $|a_{ij}|_0^m = 1$ .

**1.11.** Докажите, что для любых действительных чисел  $a, b, c, d, e$  и  $f$

$$\begin{vmatrix} (a+b)de - (d+e)ab & ab - de & a + b - d - e \\ (b+c)ef - (e+f)bc & bc - ef & b + c - e - f \\ (c+d)fa - (f+a)cd & cd - fa & c + d - f - a \end{vmatrix} = 0.$$

**1.12.** ([Мо3]) Пусть  $a_{ij}$ ,  $1 \leq i, j \leq 3$ , — комплексные числа, модули которых равны 1. Докажите, что если определитель матрицы  $\|a_{ij}\|_1^3$  равен 0, то две строки или два столбца этой матрицы пропорциональны.

**1.13.** Пусть  $a_1, \dots, a_n$  — целые числа с наибольшим общим делителем  $d$ . Докажите, что существует матрица с целочисленными элементами, первая строка которой равна  $(a_1, \dots, a_n)$ , а определитель равен  $d$ .

### Определитель Вандермонда

**1.14.** Вычислите определитель

$$\begin{vmatrix} 1 & x_1 & \dots & x_1^{n-2} & (x_2 + x_3 + \dots + x_n)^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-2} & (x_1 + x_2 + \dots + x_{n-1})^{n-1} \end{vmatrix}.$$

**1.15.** Вычислите определитель

$$\begin{vmatrix} 1 & x_1 & \dots & x_1^{n-2} & x_2 x_3 \dots x_n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-2} & x_1 x_2 \dots x_{n-1} \end{vmatrix}.$$

**1.16.** Вычислите определитель  $|a_{ik}|_0^n$ , где  $a_{ik} = \lambda_i^{n-k}(1 + \lambda_i^2)^k$ .

**1.17.** Рассмотрим матрицу Вандермонда  $V = \|a_{ij}\|$ , где  $a_{ij} = x_i^{j-1}$ . Пусть  $V_k$  — матрица, полученная из  $V$  вычёркиванием  $(k+1)$ -го столбца (состоящего из  $k$ -х степеней) и добавлением  $n$ -го столбца, состоящего из  $n$ -х степеней. Докажите, что  $\det V_k = \sigma_{n-k}(x_1, \dots, x_n) \det V$ , где  $\sigma_{n-k}$  — элементарная симметрическая функция.

**1.18.** Пусть  $P_k(x)$  — многочлен степени  $k-1$  со старшим коэффициентом  $a_k$ , т.е.  $P_k(x) = a_k x^{k-1} + \dots$ . Вычислите определитель  $|a_{ij}|_1^n$ , где  $a_{ij} = P_j(x_i)$ .

**1.19.** [Gr4] Пусть  $f(x) = (x - x_1) \cdots (x - x_n)$ , причём числа  $x_1, \dots, x_n$  попарно различны. Докажите, что:

- а)  $f'(x_i) = (-1)^{n-i} \frac{V(x_1, \dots, x_n)}{V(x_1, \dots, \hat{x}_i, \dots, x_n)}$ ;  
 б)  $\sum_{i=1}^n \frac{x_i^k}{f'(x_i)} = \frac{1}{V(x_1, \dots, x_n)} \begin{vmatrix} 1 & x_1 & \dots & x_1^{n-2} & x_1^k \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-2} & x_n^k \end{vmatrix}$ ;  
 в)  $\sum_{i=1}^n \frac{x_i^k}{f'(x_i)} = \begin{cases} 0 & \text{при } 0 \leq k \leq n-2, \quad k \in \mathbb{Z}; \\ 1 & \text{при } k = n-1. \end{cases}$

**1.20.** Пусть  $\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}$ . Докажите, что

$$\begin{vmatrix} 1 & \binom{x_1}{1} & \dots & \binom{x_1}{n-1} \\ 1 & \binom{x_2}{1} & \dots & \binom{x_2}{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & \binom{x_n}{1} & \dots & \binom{x_n}{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} \frac{x_i - x_j}{i - j}.$$

**1.21.** Пусть  $a_{ij} = \binom{in}{j}$ . Докажите, что  $|a_{ij}|_1^n = n^{r(r+1)/2}$  при  $r \leq n$ .

**1.22.** Пусть  $x_1, \dots, x_n$  — произвольные целые числа. Докажите, что число  $\prod_{1 \leq i < j \leq n} \frac{x_i - x_j}{i - j}$  целое.

**1.23.** Даны целые числа  $k_1, \dots, k_n$ . Вычислите определитель  $|a_{ij}|_1^n$ , где  $a_{ij} = 1/(k_i + j - i)!$  при  $k_i + j - i \geq 0$  и  $a_{ij} = 0$  при  $k_i + j - i < 0$ .

**1.24.** Пусть  $s_k = p_1 x_1^k + \dots + p_n x_n^k$ ,  $a_{ij} = s_{i+j}$ . Докажите, что  $|a_{ij}|_1^n = p_1 \cdots p_n \prod_{i > j} (x_i - x_j)^2$ .

**1.25.** Пусть  $s_k = x_1^k + \dots + x_n^k$ . Вычислите определитель

$$\begin{vmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} & 1 \\ s_1 & s_2 & s_3 & \dots & s_n & y \\ s_2 & s_3 & s_4 & \dots & s_{n+1} & y^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ s_n & s_{n+1} & s_{n+2} & \dots & s_{2n-1} & y^n \end{vmatrix}.$$

**1.26.** Пусть  $a_{ij} = (x_i + y_j)^n$ . Докажите, что

$$|a_{ij}|_0^n = \binom{n}{1} \cdots \binom{n}{n} \prod_{i > k} (x_i - x_k)(y_k - y_i).$$

**1.27.** Найдите над полем  $\mathbb{C}$  все решения системы уравнений  $\lambda_1^k + \dots + \lambda_n^k = 0$  ( $k = 1, \dots, n$ ).

**1.28.** Пусть  $\sigma_k(x_0, \dots, x_n)$  —  $k$ -я элементарная симметрическая функция,  $\sigma_0 = 1$ ;  $\sigma_k(\hat{x}_i) = \sigma_k(x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ . Докажите, что если  $a_{ij} = \sigma_i(\hat{x}_j)$ , то  $|a_{ij}|_0^n = \prod_{i < j} (x_i - x_j)$ .

### Соотношения между определителями

**1.29.** а) Пусть  $b_{ij} = (-1)^{i+j} a_{ij}$ . Докажите, что  $|a_{ij}|_1^n = |b_{ij}|_1^n$ .

б) Пусть  $b_{ij} = (-1)^{m_i+n_j} a_{ij}$ . Докажите, что  $|a_{ij}|_1^n = (-1)^{M+N} |b_{ij}|_1^n$ , где  $M = \sum m_i$  и  $N = \sum n_i$ .

**1.30.** Пусть  $b_{ij} = a_{n+1-i, n+1-j}$ . Докажите, что  $\|b_{ij}\|_1^n = \|a_{ij}\|_1^n$ .

**1.31.** Докажите, что

$$\begin{vmatrix} a_1 c_1 & a_2 d_1 & a_1 c_2 & a_2 d_2 \\ a_3 c_1 & a_4 d_1 & a_3 c_2 & a_4 d_2 \\ b_1 c_3 & b_2 d_3 & b_1 c_4 & b_2 d_4 \\ b_3 c_3 & b_4 d_3 & b_3 c_4 & b_4 d_4 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ a_3 & a_4 \end{vmatrix} \cdot \begin{vmatrix} b_1 & b_2 \\ b_3 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & c_2 \\ c_3 & c_4 \end{vmatrix} \cdot \begin{vmatrix} d_1 & d_2 \\ d_3 & d_4 \end{vmatrix}.$$

**1.32.** Докажите, что

$$\begin{vmatrix} a_1 & 0 & 0 & b_1 & 0 & 0 \\ 0 & a_2 & 0 & 0 & b_2 & 0 \\ 0 & 0 & a_3 & 0 & 0 & b_3 \\ b_{11} & b_{12} & b_{13} & a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} & a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} & a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_1 a_{11} - b_1 b_{11} & a_2 a_{12} - b_2 b_{12} & a_3 a_{13} - b_3 b_{13} \\ a_1 a_{21} - b_1 b_{21} & a_2 a_{22} - b_2 b_{22} & a_3 a_{23} - b_3 b_{23} \\ a_1 a_{31} - b_1 b_{31} & a_2 a_{32} - b_2 b_{32} & a_3 a_{33} - b_3 b_{33} \end{vmatrix}.$$

**1.33.** Пусть  $s_k = \sum_{i=1}^n a_{ki}$ . Докажите, что

$$\begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} \\ \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} \end{vmatrix} = (-1)^{n-1} (n-1) \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

**1.34.** Докажите, что

$$\begin{vmatrix} \binom{n}{m_1} & \binom{n}{m_1-1} & \dots & \binom{n}{m_1-k} \\ \dots & \dots & \dots & \dots \\ \binom{n}{m_k} & \binom{n}{m_k-1} & \dots & \binom{n}{m_k-k} \end{vmatrix} = \begin{vmatrix} \binom{n}{m_1} & \binom{n+1}{m_1} & \dots & \binom{n+k}{m_1} \\ \dots & \dots & \dots & \dots \\ \binom{n}{m_k} & \binom{n+1}{m_k} & \dots & \binom{n+k}{m_k} \end{vmatrix}.$$



**1.35.** Пусть  $\Delta_n(k) = |a_{ij}|_0^n$ , где  $a_{ij} = \binom{k+i}{2j}$ . Докажите, что

$$\Delta_n(k) = \frac{k(k+1) \dots (k+n-1)}{1 \cdot 3 \dots (2n-1)} \Delta_{n-1}(k-1).$$

**1.36.** Пусть  $D_n = |a_{ij}|_0^n$ , где  $a_{ij} = \binom{n+1}{2j-1}$ . Докажите, что  $D_n = 2^{n(n+1)/2}$ .

**1.37.** Даны числа  $a_0, a_1, \dots, a_n$ . Пусть  $b_k = \sum_{i=0}^k (-1)^i \binom{k}{i} a_i$  ( $k = 0, \dots, 2n$ ),  $a_{ij} = a_{i+j}$ ,  $b_{ij} = b_{i+j}$ . Докажите, что  $|a_{ij}|_0^n = |b_{ij}|_0^n$ .

**1.38.** [Fil] Пусть  $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  и  $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ , где  $A_{11}$  и  $B_{11}$ , а также  $A_{22}$  и  $B_{22}$  — квадратные матрицы одного порядка соответственно, причем  $\text{rk } A_{11} = \text{rk } A$  и  $\text{rk } B_{22} = \text{rk } B$ . Докажите, что

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \cdot \begin{vmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{vmatrix} = |A+B| \cdot |A_{11}| \cdot |B_{22}|.$$

**1.39.** Пусть  $A$  и  $B$  — квадратные матрицы порядка  $n$ . Докажите, что  $|| \cdot || = \sum_{k=1}^n |A_k| \cdot |B_k|$ , где матрицы  $A_k$  и  $B_k$  получены из  $A$  и  $B$  обменом столбцов с номерами 1 и  $k$  соответственно (первый столбец матрицы  $A$  и  $k$ -й столбец матрицы  $B$  меняются местами).

**1.40.** Даны столбцы  $A_1, \dots, A_p, B_1, \dots, B_{n-p}, C_1, \dots, C_p$  длины  $n$ . Пусть  $a_{ij} = |\hat{A}_i B C_j|$  — определитель матрицы  $(A_1 \dots \hat{A}_i \dots A_p B_1 \dots B_{n-p} C_j)$ . Докажите, что  $|a_{ij}|_1^p = |AB|^{p-1} |BC|$  (тождество Базена).

### Числа Фибоначчи

Числами Фибоначчи называют последовательность чисел  $F_1 = 1, F_2 = 1, F_n = F_{n-1} + F_{n-2}$  при  $n \geq 3$ . Начало этой последовательности имеет вид 1, 1, 2, 3, 5, 8, 13, 21, ...

**1.41.** Докажите, что при  $n \geq 2$

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}^n = (-1)^n \begin{pmatrix} F_{n+1} & -F_n \\ -F_n & F_{n-1} \end{pmatrix}.$$

**1.42.** Докажите, что определитель

$$\begin{vmatrix} x & a & 0 & \dots & 0 & 0 \\ -a & x & a & \dots & 0 & 0 \\ 0 & -a & x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x & a \\ 0 & 0 & 0 & \dots & -a & x \end{vmatrix}$$

порядка  $n$  равен

$$x^n + \binom{n-1}{1} a^2 x^{n-2} + \binom{n-2}{2} a^4 x^{n-4} + \binom{n-3}{3} a^6 x^{n-6} + \dots$$

**1.43.** Докажите, что

$$F_{n+1} = 1 + \binom{n-1}{1} + \binom{n-2}{2} + \binom{n-3}{3} + \dots$$

## 1.2. Миноры и алгебраические дополнения

### 1.2.1. Определение минора

Во многих случаях для матрицы  $A$  бывает полезно рассмотреть определитель матрицы, элементы которой стоят на пересечениях некоторых  $p$  строк и  $p$  столбцов матрицы  $A$ . Такой определитель называют *минором*  $p$ -го порядка матрицы  $A$ . Для удобства введем следующее обозначение:

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{vmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \dots & a_{i_1 k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_p k_1} & a_{i_p k_2} & \dots & a_{i_p k_p} \end{vmatrix}.$$

Если  $i_1 = k_1, \dots, i_p = k_p$ , то минор называют *главным*, а если  $i_1 = k_1 = 1, \dots, i_p = k_p = p$ , то — *угловым*.

### 1.2.2. Ранг матрицы

Ненулевой минор максимального порядка называют *базисным*, а его порядок называют *рангом* матрицы.

**Теорема 1.2.2.1.** Если  $A \begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix}$  — базисный минор матрицы  $A$ , то все её строки являются линейными комбинациями строк с номерами  $i_1, \dots, i_p$ , причём сами эти строки линейно независимы.

**Доказательство.** Линейная независимость строк с номерами  $i_1, \dots, i_p$  очевидна, так как определитель матрицы с линейно зависимыми строками равен нулю. Очевиден также случай, когда размер матрицы  $A$  равен  $m \times p$  или  $p \times m$ .

Доказательство достаточно провести для базисного минора  $A \begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix}$ . Определитель  $\begin{vmatrix} a_{11} & \dots & a_{1p} & a_{1j} \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & a_{pj} \\ a_{i1} & \dots & a_{ip} & a_{ij} \end{vmatrix}$  равен нулю как при  $j \leq p$ , так и при  $j > p$ . Разложив его по последнему столбцу, получим соотношение вида

$$a_{1j}c_1 + a_{2j}c_2 + \dots + a_{pj}c_p + a_{ij}c = 0,$$

где числа  $c_1, \dots, c_p, c$  не зависят от  $j$  (но зависят от  $i$ ), причём  $c = A \begin{pmatrix} 1 & \dots & p \\ 1 & \dots & p \end{pmatrix} \neq 0$ . Значит,  $i$ -я строка равна линейной комбинации первых  $p$  строк с коэффициентами  $-c_1/c, \dots, -c_p/c$ .  $\square$

**Следствие 1.** Ранг матрицы  $A$  равен наибольшему числу её линейно независимых строк.

**Доказательство.** Если  $A \begin{pmatrix} i_1 & \dots & i_p \\ k_1 & \dots & k_p \end{pmatrix}$  — базисный минор, то все строки матрицы  $A$  лежат в линейном подпространстве, порожденном строками с номерами  $i_1, \dots, i_p$ .  $\square$

**Следствие 2.** Ранг матрицы равен наибольшему числу её линейно независимых столбцов.

**Доказательство.** Аналогичные рассуждения можно провести и для столбцов.  $\square$

### 1.2.3. Формула Бине—Коши

Определитель произведения двух прямоугольных матриц можно выразить через миноры сомножителей.

**Теорема 1.2.3.1 (формула Бине–Коши).** Пусть  $A$  и  $B$  — матрицы размера  $n \times m$  и  $m \times n$  соответственно, причём  $n \leq m$ . Тогда

$$\det AB = \sum_{1 \leq k_1 < \dots < k_n \leq m} A_{k_1 \dots k_n} B^{k_1 \dots k_n},$$

где  $A_{k_1 \dots k_n}$  — минор, полученный из столбцов матрицы  $A$  с номерами  $k_1, \dots, k_n$ , а  $B^{k_1 \dots k_n}$  — минор, полученный из строк матрицы  $B$  с номерами  $k_1, \dots, k_n$ .

**Доказательство.** Пусть  $C = AB$ ,  $c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}$ . Тогда

$$\begin{aligned} \det C &= \sum_{\sigma} (-1)^{\sigma} \sum_{k_1} a_{1k_1} b_{k_1 \sigma(1)} \dots \sum_{k_n} a_{nk_n} b_{k_n \sigma(n)} = \\ &= \sum_{k_1, \dots, k_n=1}^m a_{1k_1} \dots a_{nk_n} \sum_{\sigma} (-1)^{\sigma} b_{k_1 \sigma(1)} \dots b_{k_n \sigma(n)} = \\ &= \sum_{k_1, \dots, k_n=1}^m a_{1k_1} \dots a_{nk_n} B^{k_1 \dots k_n}. \end{aligned}$$

Минор  $B^{k_1 \dots k_n}$  отличен от нуля, только если числа  $k_1, \dots, k_n$  попарно различны, поэтому суммирование можно вести по попарно различным  $k_1, \dots, k_n$ . А так как  $B^{\tau(k_1) \dots \tau(k_n)} = (-1)^{\tau} B^{k_1 \dots k_n}$  для любой перестановки  $\tau$  чисел  $k_1, \dots, k_n$ , то

$$\begin{aligned} \sum_{k_1, \dots, k_n=1}^m a_{1k_1} \dots a_{nk_n} B^{k_1 \dots k_n} &= \sum_{k_1 < k_2 < \dots < k_n} (-1)^{\tau} a_{1\tau(1)} \dots a_{n\tau(n)} B^{k_1 \dots k_n} = \\ &= \sum_{1 \leq k_1 < k_2 < \dots < k_n \leq m} A_{k_1 \dots k_n} B^{k_1 \dots k_n}. \end{aligned}$$

□

**Замечание.** Другое доказательство формулы Бине–Коши см. в решении задачи 2.8.

#### 1.2.4. Алгебраическое дополнение минора

Напомним формулу разложения определителя по  $i$ -й строке:

$$|a_{ij}|_1^n = \sum_{j=1}^n (-1)^{i+j} a_{ij} M_{ij}, \quad (1.1)$$

где  $M_{ij}$  — определитель матрицы, полученной из матрицы  $A = |a_{ij}|_1^n$  вычёркиванием  $i$ -й строки и  $j$ -го столбца. Число  $A_{ij} = (-1)^{i+j} M_{ij}$  называют *алгебраическим дополнением* элемента в матрице  $A$ ; ясно, что  $A_{ij}$  —

сумма всех сомножителей при  $a_{ij}$  в разложении определителя матрицы  $A$ .

Определитель можно раскладывать не только по одной строке, но и по нескольким. Фиксируем строки с номерами  $i_1, \dots, i_p$ , где  $i_1 < i_2 < \dots < i_p$ . В разложение определителя матрицы  $A$  входят произведения членов разложения минора  $A \begin{pmatrix} i_1 & \dots & i_p \\ j_1 & \dots & j_p \end{pmatrix}$ , где  $j_1 < \dots < j_p$ , на члены разложения минора  $A \begin{pmatrix} i_{p+1} & \dots & i_n \\ j_{p+1} & \dots & j_n \end{pmatrix}$ , где  $i_{p+1} < \dots < i_n$ , причём никаких других членов в разложении определителя матрицы  $A$  нет. Чтобы вычислить знаки при этих произведениях, переставим строки и столбцы так, чтобы минор  $A \begin{pmatrix} i_1 & \dots & i_p \\ j_1 & \dots & j_p \end{pmatrix}$  стал угловым. Для этого нужно совершить  $(i_1 - 1) + \dots + (i_p - p) + (j_1 - 1) + \dots + (j_p - p) \equiv i + j \pmod{2}$  перестановок, где  $i = i_1 + \dots + i_p$ ,  $j = j_1 + \dots + j_p$ . Число  $(-1)^{i+j} A \begin{pmatrix} i_{p+1} & \dots & i_n \\ j_{p+1} & \dots & j_n \end{pmatrix}$  называют *алгебраическим дополнением* минора  $A \begin{pmatrix} i_1 & \dots & i_p \\ j_1 & \dots & j_p \end{pmatrix}$ . Мы доказали следующее утверждение.

**Теорема 1.2.4.1 (Лаплас).** *Фиксируем  $p$  строк матрицы  $A$ . Тогда сумма произведений миноров порядка  $p$ , принадлежащих этим строкам, на их алгебраические дополнения равна определителю матрицы  $A$ .*

### 1.2.5. Присоединённая матрица

Матрицу  $\text{adj } A = \|A_{ij}\|^T$  называют *присоединённой* к  $A$ . Докажем, что  $A(\text{adj } A) = |A| \cdot I$ . Для этого нужно проверить, что  $\sum_{j=1}^n a_{ij} A_{kj} = \delta_{ki} |A|$ . При  $k = i$  эта формула совпадает с (1.1). Если же  $k \neq i$ , то заменим  $k$ -ю строку матрицы  $A$  на  $i$ -ю. Определитель полученной матрицы равен нулю; разлагая его по  $k$ -й строке, приходим к требуемому равенству:  $0 = \sum_{j=1}^n a'_{kj} A_{kj} = \sum_{j=1}^n a_{ij} A_{kj}$ .

Аналогично можно доказать, что  $(\text{adj } A)A = |A| \cdot I$ , т.е.  $\sum_{j=1}^n A_{ij} a_{kj} = \delta_{ki} |A|$ . При  $k = i$  равенство то же самое, а если  $k \neq i$ , то заменим  $i$ -ю строку матрицы  $A$  на  $k$ -ю. Разлагая определитель полученной матрицы  $i$ -й строке, приходим к требуемому равенству:  $0 = \sum_{j=1}^n a'_{ij} A_{kj} = \sum_{j=1}^n A_{ij} a_{kj}$ .

Пусть  $A$  — квадратная матрица. Матрицу  $B$  называют *обратной* к  $A$ , если  $AB = I$ . (Точнее говоря, такую матрицу называют *правой* обратной, но вскоре мы убедимся, что левая обратная совпадает с правой обратной.)

Легко видеть, что матрица  $B$  единственна. Действительно, пусть  $B_1 \neq B_2$  и  $AB_1 = AB_2 = I$ . Тогда  $A(B_1 - B_2) = 0$ , т.е.  $AX = 0$  для некоторой ненулевой матрицы  $X$ . Пусть  $x$  — ненулевой столбец матрицы  $X$ . Равенство  $Ax = 0$  означает, что столбцы матрицы  $A$  линейно зависимы. Тогда строки матрицы  $A$  тоже линейно зависимы, поэтому строки матрицы  $I = AB$  тоже линейно зависимы. Приходим к противоречию.

Если у матрицы  $A$  есть обратная матрица  $B$ , то  $\det A \cdot \det B = \det I = 1$ , поэтому  $\det A \neq 0$ . Наоборот, если матрица  $A$  невырожденная, то обратная матрица  $B$  задаётся формулой  $B = (\text{adj } A)/|A|$ . В таком случае  $BA = I$ , поскольку  $\sum_{j=1}^n A_{ij}a_{kj} = \sum_{j=1}^n a_{kj}A_{ij} = \delta_{ik}|A|$ .

Обратную матрицу обозначают  $A^{-1}$ .

**Теорема 1.2.5.1.** *Операция  $\text{adj}$  обладает следующими свойствами:*

- а)  $\text{adj}(AB) = \text{adj } B \cdot \text{adj } A$ ;
- б)  $\text{adj}(XAX^{-1}) = X(\text{adj } A)X^{-1}$ ;
- в) если  $A$  — матрица порядка  $n$ , а  $\lambda$  — число, то  $\text{adj}(\lambda A) = \lambda^{n-1} \text{adj } A$ ;
- г) если  $AB = BA$ , то  $(\text{adj } A)B = B(\text{adj } A)$ .

**Доказательство.** а) Так как для обратимых матриц выполнено равенство  $(AB)^{-1} = B^{-1}A^{-1}$ , то свойство а) выполнено для невырожденных матриц. Обе части равенства а) непрерывно зависят от элементов матриц  $A$  и  $B$ , поэтому оно выполняется для всех матриц.

Свойство б) очевидно следует из а).

в) Если матрица  $A$  невырожденная, то  $\text{adj}(\lambda A) = (\lambda A)^{-1}|\lambda A| = \lambda^{-1}A^{-1}\lambda^n|A| = \lambda^{n-1}\text{adj } A$ . Обе части равенства  $\text{adj}(\lambda A) = \lambda^{n-1}\text{adj } A$  непрерывно зависят от элементов матрицы  $A$ , поэтому оно выполняется для всех матриц.

г) Если  $AB = BA$  и матрица  $A$  обратима, то  $A^{-1}(AB)A^{-1} = A^{-1}(BA)A^{-1}$ , т.е.  $BA^{-1} = A^{-1}B$ . Поэтому свойство в) выполняется для обратимых матриц  $A$ . Для вырожденной матрицы  $A$  рассмотрим семейства матриц  $X(\lambda) = [\text{adj}(A + \lambda I)]B$  и  $Y(\lambda) = B[\text{adj}(A + \lambda I)]$ . Для почти всех  $\lambda$  матрица  $A + \lambda I$  невырождена. Действительно,  $|A + \lambda I| = \lambda^n + \dots$  — многочлен от  $\lambda$  со старшим коэффициентом 1; этот многочлен имеет конечное число корней. Если матрица  $A + \lambda I$  невырождена, то выполняется равенство  $X(\lambda) = Y(\lambda)$ . Элементы матриц  $X(\lambda)$  и  $Y(\lambda)$  являются многочленами от  $\lambda$ , поэтому  $X(\lambda) = Y(\lambda)$  для всех  $\lambda$ . В частности,  $X(0) = Y(0)$ , т.е.  $(\text{adj } A)B = B(\text{adj } A)$ .  $\square$

### 1.2.6. Миноры присоединённой матрицы. Тождество Якоби

Миноры матрицы  $A$  и дополнительные к ним миноры матрицы  $(\text{adj } A)^T$  связаны весьма простыми соотношениями. Мы сначала рас-

смотрим угловые миноры, а уже затем произвольные.

**Теорема 1.2.6.1.** Пусть  $A = \|a_{ij}\|_1^n$ ,  $(\text{adj } A)^T = \|A_{ij}\|_1^n$  и  $p$  — натуральное число, причём  $1 \leq p < n$ . Тогда

$$\begin{pmatrix} A_{11} & \dots & A_{1p} \\ \dots & \dots & \dots \\ A_{p1} & \dots & A_{pp} \end{pmatrix} = |A|^{p-1} \begin{pmatrix} a_{p+1,p+1} & \dots & a_{p+1,n} \\ \dots & \dots & \dots \\ a_{n,p+1} & \dots & a_{nn} \end{pmatrix}.$$

**Доказательство.** При  $p = 1$  утверждение совпадает с определением алгебраического дополнения  $A_{11}$ . Пусть теперь  $p > 1$ . Легко проверить, что

$$\begin{pmatrix} A_{11} & \dots & A_{1p} & A_{1,p+1} & \dots & A_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ A_{p1} & \dots & A_{pp} & A_{p1} & \dots & A_{pn} \\ \dots & 0 & \dots & I & \dots & \dots \end{pmatrix} A = \begin{pmatrix} |A| & \dots & 0 & \dots & 0 \\ 0 & \dots & |A| & \dots & 0 \\ a_{1,p+1} & \dots & \dots & \dots & a_{n,p+1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{1n} & \dots & \dots & \dots & a_{nn} \end{pmatrix}.$$

Следовательно,

$$\begin{pmatrix} A_{11} & \dots & A_{1p} \\ \dots & \dots & \dots \\ A_{p1} & \dots & A_{pp} \end{pmatrix} \cdot |A| = |A|^p \begin{pmatrix} a_{p+1,p+1} & \dots & a_{p+1,n} \\ \dots & \dots & \dots \\ a_{n,p+1} & \dots & a_{nn} \end{pmatrix}.$$

Если  $|A| \neq 0$ , то мы можем сократить на  $|A|$ . В результате получим требуемое. В случае, когда  $|A| = 0$ , можно воспользоваться тем, что обе части требуемого равенства непрерывно зависят от  $a_{ij}$ , а любая вырожденная матрица является пределом невырожденных матриц.  $\square$

Выясним теперь, как преобразуется присоединённая матрица при перестановке двух строк (столбцов).

**Теорема 1.2.6.2.** При перестановке двух строк (столбцов) матрицы  $A$  в присоединённой матрице происходит точно такая же перестановка столбцов (строк) и, кроме того, все её элементы меняют знак.

**Доказательство.** Рассмотрим сначала случай, когда матрица  $A$  невырожденная. Ясно, что при перестановке двух строк (столбцов) матрицы  $A$  в матрице  $A^{-1}$  происходит точно такая же перестановка столбцов (строк). Кроме того, при этой перестановке определитель  $|A|$  меняет знак. Остаётся заметить, что  $\text{adj } A = |A|A^{-1}$ . Для вырожденной матрицы можно воспользоваться теми же рассуждениями, что и при доказательстве теоремы 1.2.6.1  $\square$

**Следствие.** Для вырожденной матрицы  $A$  ранг матрицы  $\text{adj } A$  не превосходит 1.

**Доказательство.** Применив теорему 1.2.6.1 при  $p = 2$ , получим

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = |A| \cdot \begin{pmatrix} a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots \\ a_{n3} & \dots & a_{nn} \end{pmatrix} = 0.$$

Затем, воспользовавшись теоремой 1.2.6.2, аналогично доказываем, что все остальные миноры порядка 2 матрицы  $\text{adj } A$  тоже равны 0.  $\square$

**Теорема 1.2.6.3 (тождество Якоби).** Пусть  $A = \|a_{ij}\|_1^n$ ,  $(\text{adj } A)^T = \|A_{ij}\|_1^n$ ,  $p$  — натуральное число, причём  $1 \leq p < n$ ,  $\sigma = \begin{pmatrix} i_1 & \dots & i_n \\ j_1 & \dots & j_n \end{pmatrix}$  — произвольная перестановка. Тогда

$$\begin{pmatrix} A_{i_1 j_1} & \dots & A_{i_1 j_p} \\ \dots & \dots & \dots \\ A_{i_p j_1} & \dots & A_{i_p j_p} \end{pmatrix} = (-1)^\sigma |A|^{p-1} \begin{pmatrix} a_{i_{p+1}, j_{p+1}} & \dots & a_{i_{p+1}, j_n} \\ \dots & \dots & \dots \\ a_{i_n, j_{p+1}} & \dots & a_{i_n, j_n} \end{pmatrix}.$$

**Доказательство.** Применим теорему 1.2.6.1 к матрице  $B = \|b_{kl}\|_1^n$ , где  $b_{kl} = a_{i_k j_l}$ . В результате получим

$$\begin{pmatrix} B_{11} & \dots & B_{1p} \\ \dots & \dots & \dots \\ B_{p1} & \dots & B_{pp} \end{pmatrix} = |B|^{p-1} \begin{pmatrix} b_{p+1, p+1} & \dots & b_{p+1, n} \\ \dots & \dots & \dots \\ b_{n, p+1} & \dots & b_{nn} \end{pmatrix}. \quad (1)$$

Согласно теореме 1.2.6.2  $B_{kl} = (-1)^\sigma A_{i_k j_l}$ . Ясно также, что  $|B| = (-1)^\sigma |A|$ . Поэтому, сокращая обе части равенства (1) на  $((-1)^\sigma)^p$ , получаем требуемое.  $\square$

### 1.2.7. Тождество Льюиса Кэрролла

Интересным частным случаем тождества Якоби является *тождество Льюиса Кэрролла*. Оно соответствует случаю  $p = 2$ ,  $\sigma = \begin{pmatrix} 1 & n & 2 & \dots & n-1 \\ 1 & n & 2 & \dots & n-1 \end{pmatrix}$ . Это тождество записывается следующим образом. Пусть  $A_{i_1 \dots i_k}^{j_1 \dots j_k}$  — минор, полученный из матрицы  $A$  вычёркиванием строк с номерами  $i_1, \dots, i_k$  и столбцов с номерами  $j_1, \dots, j_k$ . Тогда

$$|A| A_{1n}^{1n} = A_1^1 A_n^n - A_n^1 A_1^n.$$



Тождество Льюиса Кэрролла появилось в статье Чарльза Доджсона [Do], более известного под псевдонимом Льюис Кэрролл. Простое комбинаторное доказательство этого тождества приведено в [Ze2].

Тождество Льюиса Кэрролла является весьма эффективным средством для вычисления определителей. В качестве примера вычислим один определитель, впервые вычисленный Мак-Магоном.

**Теорема 1.2.7.1.** *Определитель матрицы  $\|a_{ij}\|_1^n$ , где  $a_{ij} = \binom{a+i}{b+j}$ , равен*

$$\frac{(a+n)!!(n-1)!!(a-b-1)!!(b)!!}{(a)!!(a-b+n-1)!!(b+n)!!}. \quad (1)$$

**Доказательство.** [Ze1] Пусть  $X_n(a, b) = \left| \binom{a+i}{b+j} \right|_1^n$ , а  $Y_n(a, b)$  — выражение (1). В данном случае тождество Льюиса Кэрролла записывается так:

$$X_n(a, b)X_{n-2}(a+1, b+1) = X_{n-1}(a, b)X_{n-1}(a+1, b+1) - X_{n-1}(a+1, b)X_{n-1}(a, b+1). \quad (2)$$

Проверим, что  $Y_n(a, b)$  тоже удовлетворяют такому рекуррентному соотношению. Сократив на общее выражение

$$\frac{(a+n)!!(a+n-1)!!(b)!!(b+1)!!}{(a)!!(a+1)!!(b+n)!!(b+n-1)!!},$$

а затем поделив обе части на  $(n-2)!!(n-3)!!(a-b-1)!!(a-b-2)!!$  и умножив на  $(a-b+n-3)!!(a-b+n-2)!!$ , получим соотношение

$$\frac{(n-1)!(a-b-1)!}{(a-b+n-1)!} = \frac{(n-2)!(a-b-1)!}{(a-b+n-2)!} - \frac{(n-2)!(a-b)!}{(a-b+n-1)!}$$

(мы воспользовались тем, что  $\frac{m!!}{(m-1)!!} = m!$ ). Поделим обе части этого соотношения на  $(n-2)!(a-b-1)!$  и умножим на  $(a-b+n-2)!$ . В результате получим соотношение

$$\frac{n-1}{a-b+n-1} = 1 - \frac{a-b}{a-b+n-1},$$

которое легко проверяется.

Соотношение (2) позволяет выразить  $X_n$  через  $X_{n-1}$  и  $X_{n-2}$ . Поэтому остаётся доказать, что  $X_1(a, b) = Y_1(a, b)$  и  $X_2(a, b) = Y_2(a, b)$ . Ясно, что

$$X_1(a, b) = \frac{(a+1)!}{(a-b)!(b+1)!} = Y_1(a, b),$$

$$\begin{aligned} X_2(a, b) &= \frac{(a+1)!(a+2)!}{(b+1)!(b+2)!(a-b)!(a-b-1)!} \left( \frac{1}{a-b} - \frac{1}{a-b+1} \right) = \\ &= \frac{(a+1)!(a+2)!}{(b+1)!(b+2)!(a-b)!(a-b+1)!} = Y_2(a, b). \end{aligned}$$

□

### 1.2.8. Ассоциированная матрица

Кроме присоединённой матрицы иногда рассматривают *ассоциированную* матрицу  $\|M_{ij}\|_1^n$ , составленную из миноров порядка  $n - 1$  матрицы  $A$ . Определители присоединённой и ассоциированной матриц равны (см., например, задачу 1.29).

Можно также рассмотреть для матрицы  $A$  размера  $m \times n$  матрицу, составленную из миноров порядка  $r$ , где  $r \leq \min(m, n)$ ; миноры при этом упорядочиваются лексикографически. Полученную матрицу  $C_r(A)$  называют  *$r$ -й ассоциированной* для матрицы  $A$ .

Используя формулу Бине–Коши, можно проверить, что  $C_r(AB) = C_r(A)C_r(B)$ .

Для квадратной матрицы  $A$  порядка  $n$  определители матриц  $A$  и  $C_r(A)$  связаны соотношением

$$\det C_r(A) = (\det A)^p, \text{ где } p = \binom{n-1}{r-1}.$$

Наиболее простое доказательство этого утверждения использует понятие внешней степени (теорема 5.2.5.3).

### 1.2.9. Обобщённое тождество Сильвестра

Пусть  $1 \leq m \leq r < n$ ,  $A = \|a_{ij}\|_1^n$ ,  $A_n = |A| = |a_{ij}|_1^n$  и  $A_m = |a_{ij}|_1^m$ . Рассмотрим матрицу  $S_{m,n}^r$ , составленную из миноров порядка  $r$  матрицы  $A$ , содержащих угловой минор  $A_m$ . Определитель этой матрицы является минором порядка  $\binom{n-m}{r-m}$  матрицы  $C_r(A)$ . Определитель матрицы  $S_{m,n}^r$  выражается через  $A_n$  и  $A_m$  следующим образом.

**Теорема 1.2.9.1 (обобщённое тождество Сильвестра).**

$$|S_{m,n}^r| = A_m^p A_n^q, \text{ где } p = \binom{n-m-1}{r-m}, \quad q = \binom{n-m-1}{r-m-1}. \quad (1.1)$$

**Доказательство.** [Мо1] Докажем равенство (1.1) индукцией по  $n$ . При  $n = 2$  оно очевидно. Матрица  $S_{0,n}^r$  совпадает с  $C_r(A)$ . Согласно теореме 5.2.5.3  $|C_r(A)| = A_n^q$ , где  $q = \binom{n-1}{r-1}$ . Поэтому равенство (1.1) справедливо при  $m = 0$  (мы считаем, что  $A_0 = 1$ ). Обе части равенства (1.1) являются непрерывными функциями от переменных  $a_{ij}$ , поэтому доказательство шага индукции достаточно провести в случае, когда  $a_{11} \neq 0$ .

Все рассматриваемые миноры содержат первую строку, поэтому из строк с номерами  $2, \dots, n$  можно вычесть первую строку, домноженную на произвольный коэффициент, и определитель матрицы  $S_{m,n}^r$  при этом не изменится. С помощью такой операции все элементы первого столбца матрицы  $A$ , кроме элемента  $a_{11}$ , можно сделать нулевыми. Пусть  $\bar{A}$  — матрица, полученная из новой матрицы вычёркиванием первой строки и первого столбца;  $\bar{S}_{m-1,n-1}^{r-1}$  — матрица, составленная из миноров порядка  $r-1$  матрицы  $\bar{A}$ , содержащих её угловой минор порядка  $m-1$ . Очевидно, что  $S_{m,n}^r = a_{11} \bar{S}_{m-1,n-1}^{r-1}$ , причём к матрице  $\bar{S}_{m-1,n-1}^{r-1}$  можно применить предположение индукции (случай  $m-1=0$  мы уже рассмотрели отдельно). Кроме того, если  $\bar{A}_{m-1}$  и  $\bar{A}_{n-1}$  — угловые миноры порядка  $m-1$  и  $n-1$  матрицы  $\bar{A}$ , то  $A_m = a_{11} \bar{A}_{m-1}$  и  $A_n = a_{11} \bar{A}_{n-1}$ . Поэтому

$$|S_{m,n}^r| = a_{11}^t \bar{A}_{m-1}^{p_1} \bar{A}_{n-1}^{q_1} = a_{11}^{t-p_1-q_1} \bar{A}_m^{p_1} \bar{A}_n^{q_1},$$

где  $t = \binom{n-m}{r-m}$ ,  $p_1 = \binom{n-m-1}{r-m} = p$  и  $q_1 = \binom{n-m-1}{r-m-1} = q$ . Учитывая, что  $t = p + q$ , получаем требуемое.  $\square$

Классическое *тождество Сильвестра* получается при  $r = m + 1$ . Оно имеет вид  $|S_{m,n}^{m+1}| = A_m A^{n-m-1} A_n$ . Отметим, что при  $n = m + 2$  тождество Сильвестра превращается в тождество Льюиса Кэрролла. Нужно только вместо миноров, окаймляющих угловой минор, взять миноры, окаймляющие центральный минор (т.е. минор, полученный при выбрасывании первых и последних строк и столбцов).

### 1.2.10. Миноры матрицы Якоби

Любой минор матрицы Якоби  $J$  равен некоторому произведению её главных миноров и элементов  $b_i, c_i$ . А именно, справедливо следующее утверждение.

**Теорема 1.2.10.1.** Пусть  $1 \leq i_1 < i_2 < \dots < i_m \leq n$ ,  $1 \leq j_1 < j_2 < \dots < j_m \leq n$  и  $i_1 = j_1, \dots, i_{p_1} = j_{p_1}, i_{p_1+1} \neq j_{p_1+1}, i_{p_2} \neq j_{p_2}, i_{p_2+1} = j_{p_2+1}, \dots, i_{p_3} = j_{p_3}, \dots$  Тогда

$$J \begin{pmatrix} i_1 \dots i_m \\ j_1 \dots j_m \end{pmatrix} = J \begin{pmatrix} i_1 \dots i_{p_1} \\ j_1 \dots j_{p_1} \end{pmatrix} J \begin{pmatrix} i_{p_1+1} \\ j_{p_1+1} \end{pmatrix} \dots J \begin{pmatrix} i_{p_2} \\ j_{p_2} \end{pmatrix} J \begin{pmatrix} i_{p_2+1} \dots i_{p_3} \\ j_{p_2+1} \dots j_{p_3} \end{pmatrix} \dots$$

**Доказательство.** Достаточно проверить, что если  $i_p \neq j_p$ , то

$$J \begin{pmatrix} i_1 \dots i_m \\ j_1 \dots j_m \end{pmatrix} = J \begin{pmatrix} i_1 \dots i_{p-1} \\ j_1 \dots j_{p-1} \end{pmatrix} J \begin{pmatrix} i_p \\ j_p \end{pmatrix} J \begin{pmatrix} i_{p+1} \dots i_m \\ j_{p+1} \dots j_m \end{pmatrix}.$$

Это равенство следует из того, что матрица, соответствующая рассматриваемому минору, имеет блочно треугольный вид.  $\square$

### 1.2.11. Миноры матрицы Грина

Пусть  $a_1, \dots, a_n$  и  $b_1, \dots, b_n$  — некоторые ненулевые числа. Матрицей Грина называют матрицу  $G = \|g_{ij}\|_1^n$ , где

$$g_{ij} = \begin{cases} a_i b_j & \text{при } i \leq j; \\ a_j b_i & \text{при } i \geq j. \end{cases}$$

**Теорема 1.2.11.1.** Пусть  $1 \leq i_1 < i_2 < \dots < i_m \leq n$  и  $1 \leq j_1 < j_2 < \dots < j_m \leq n$ . Тогда если  $i_1, j_1 < i_2, j_2 < \dots < i_m, j_m$ , то

$$G \begin{pmatrix} i_1 \dots i_m \\ j_1 \dots j_m \end{pmatrix} = a_{k_1} \begin{vmatrix} a_{k_2} & a_{l_1} \\ b_{k_2} & b_{l_1} \end{vmatrix} \dots \begin{vmatrix} a_{k_m} & a_{l_{m-1}} \\ b_{k_m} & b_{l_{m-1}} \end{vmatrix} b_{l_m},$$

где  $k_p = \min(i_p, j_p)$  и  $l_p = \max(i_p, j_p)$ . Если же условие  $i_1, j_1 < i_2, j_2 < \dots < i_m, j_m$  не выполняется, то указанный минор равен 0.

**Доказательство.** Если  $i_1 < i_2 \leq j_1$ , то рассматриваемый минор является определителем матрицы, у которой первые две строки пропорциональны. Если же  $j_1 < j_2 \leq i_1$ , то пропорциональны первые два столбца. Поэтому можно считать, что  $\max(i_1, j_1) < \min(i_2, j_2)$ . Пусть для определённости  $i_2 \leq j_2$  (если  $i_2 \geq j_2$ , то вместо операций с первыми двумя строками нужно проделать аналогичные операции с первыми двумя столбцами.) Вычтем из первой строки вторую строку, умноженную на  $-\frac{a_{i_1}}{a_{i_2}}$ . В

результате все элементы первой строки, кроме элемента первого столбца, станут нулевыми. А первый элемент будет равен  $a_{k_1} b_{l_1} - \frac{a_{i_1}}{a_{i_2}} b_{i_2} a_{j_1} =$

$a_{k_1} \begin{vmatrix} a_{k_2} & a_{l_1} \\ b_{k_2} & b_{l_1} \end{vmatrix} \frac{1}{a_{k_2}}$ . Таким образом,

$$G \begin{pmatrix} i_1 \dots i_m \\ j_1 \dots j_m \end{pmatrix} = a_{k_1} \begin{vmatrix} a_{k_2} & a_{l_1} \\ b_{k_2} & b_{l_1} \end{vmatrix} \frac{1}{a_{k_2}} G \begin{pmatrix} i_2 \dots i_m \\ j_2 \dots j_m \end{pmatrix}.$$

Требуемое утверждение теперь доказывается очевидной индукцией.  $\square$

### 1.2.12. Вполне положительные матрицы

Матрицу называют *вполне положительной* (вполне неотрицательной), если все<sup>2</sup> её миноры положительны (неотрицательны). Это определение имеет смысл и для прямоугольных матриц, но обычно рассматривают только квадратные матрицы.

<sup>2</sup>Имеются в виду миноры порядка от 1 до наибольшего возможного размера.

Далеко не очевидно, что существуют вполне положительные матрицы сколь угодно большого порядка. Поэтому приведём некоторые примеры вполне положительных матриц.

**Теорема 1.2.12.1.** Матрица Коши  $\left\| \frac{1}{x_i + y_j} \right\|_1^n$  вполне положительна, если  $0 < x_1 < \dots < x_n$  и  $0 < y_1 < \dots < y_n$ .

**Доказательство.** Напомним, что определитель матрицы Коши равен

$$\frac{\prod_{i>j}(x_i - x_j)(y_i - y_j)}{\prod_{i,j}(x_i + y_j)}.$$

При указанных условиях он, очевидно, положителен. Ясно также, что если мы выберем из данной матрицы Коши  $k$  строк и  $k$  столбцов, то снова получим матрицу Коши, которая обладает теми же самыми свойствами.  $\square$

Пусть  $x_1, \dots, x_n$  и  $\alpha_1, \dots, \alpha_n$  — вещественные числа. Назовём матрицу  $V = \|a_{ij}\|_1^n$ , где  $a_{ij} = x_i^{\alpha_j}$ , обобщённой матрицей Вандермонда.

**Теорема 1.2.12.2.** а) Если  $0 < x_1 < \dots < x_n$  и  $\alpha_1 < \dots < \alpha_n$ , то определитель обобщённой матрицы Вандермонда не равен нулю.

б) Если помимо этого  $\alpha_1 > 0$ , то обобщённая матрица Вандермонда является вполне положительной.

**Доказательство.** а) Предположим, что определитель матрицы  $V$  равен 0. Тогда можно подобрать константы  $c_1, \dots, c_n$  так, что не все из них равны 0 и функция  $f(x) = \sum_{j=1}^n c_j x^{\alpha_j}$  имеет по крайней мере  $n$  положительных нулей  $x = x_1, \dots, x_n$ . Докажем индукцией по  $n$ , что если  $0 < x_1 < \dots < x_n$  и  $\alpha_1 < \dots < \alpha_n$ , то такого быть не может. При  $n = 1$  получаем  $f(x) = c_1 x^{\alpha_1}$ , где  $c_1 \neq 0$ . Ясно, что эта функция не обращается в нуль при  $x > 0$ . Шаг индукции доказывается следующим образом. Если функция  $f(x) = \sum_{j=1}^n c_j x^{\alpha_j}$  имеет  $n$  положительных нулей, то по теореме Ролля функция  $f_1(x) = (x^{-\alpha_1} f(x))'$  имеет по крайней мере  $n - 1$  положительный нуль. Но эта функция имеет вид  $f_1(x) = \sum_{j=2}^n d_j x^{\alpha_j}$ , где не все константы  $d_2, \dots, d_n$  равны 0.

б) Достаточно доказать, что если  $0 < x_1 < \dots < x_n$  и  $0 < \alpha_1 < \dots < \alpha_n$ , то определитель обобщённой матрицы Вандермонда положителен. Прежде всего заметим, что если  $0 < x_1 < \dots < x_n$ , то определитель обычной матрицы Вандермонда положителен, поскольку он равен  $\prod_{i>j}(x_i - x_j)$ . Рассмотрим семейство обобщённых матриц Вандермонда

$V(t) = \|x_i^{\alpha_j(t)}\|_1^n$ , где  $\alpha_j(t) = (i-1)t + (1-t)\alpha_j$ . Ясно, что  $\alpha_j(0) = \alpha_j$  и  $\alpha_j(1) = j-1$ , т.е.  $V(0)$  — рассматриваемая обобщённая матрица Вандермонда, а  $V(1)$  — обычная матрица Вандермонда. Ясно также, что если  $0 < t \leq 1$ , то  $0 < \alpha_1(t) < \dots < \alpha_n(t)$ , поэтому  $|V(t)| \neq 0$ . А так как  $|V(1)| > 0$ , то и  $|V(0)| > 0$ .  $\square$

### 1.2.13. Теорема Митчелла

Для обобщённой матрицы Вандермонда  $\|x_i^{a_j}\|_1^n$ , где числа  $a_1, \dots, a_n$  целые неотрицательные, можно доказать более точное утверждение, чем теорема 1.2.12.2. Чтобы сформулировать это утверждение, введём следующие обозначения. Пусть  $a = (a_1, \dots, a_n)$ ,  $V_a(x_1, \dots, x_n) = |x_i^{a_j}|_1^n$  и  $V(x_1, \dots, x_n) = \prod_{i < j} (x_j - x_i)$  — определитель Вандермонда. Если числа  $a_1, \dots, a_n$  целые неотрицательные, то  $V_a(x_1, \dots, x_n)$  — многочлен от переменных  $x_1, \dots, x_n$ , обращающийся в нуль при  $x_i = x_j$ . Поэтому  $P_a = V_a/V$  — многочлен; ясно, что этот многочлен однородный с целыми коэффициентами.

**Теорема 1.2.13.1 (Митчелл [Mi5]).** *Предположим, что  $0 \leq a_1 < a_2 < \dots < a_n$  и числа  $a_1, \dots, a_n$  целые. Тогда коэффициенты многочлена  $P_a$  неотрицательны и их сумма равна  $\prod_{i < j} \frac{a_j - a_i}{j - i}$ .*

**Доказательство.** [Ev1] Сначала докажем следующее вспомогательное утверждение.

**Лемма.** *Если  $a_1 = 0$ , то*

$$V_a(1, x_2, \dots, x_n) = (x_2 - 1) \cdots (x_n - 1) \sum_b V_b(x_2, \dots, x_n),$$

где суммирование ведётся по всем наборам целых чисел  $b = (b_1, \dots, b_{n-1})$ , для которых  $a_i \leq b_i < a_{i+1}$ .

**Доказательство.** Многочлен  $V_a(1, x_2, \dots, x_n)$  является определителем матрицы, у которой каждый элемент первой строки и первого столбца равен 1. Вычтем первую строку из всех остальных. В результате перейдём к определителю

$$\begin{aligned} |x_i^{a_j} - 1|_2^n &= (x_2 - 1) \cdots (x_n - 1) |x_i^{a_j-1} + x_i^{a_j-2} + \dots + x_i + 1|_2^n = \\ &= (x_2 - 1) \cdots (x_n - 1) |x_i^{a_j-1} + x_i^{a_j-2} + \dots + x_i^{a_j-1}|_2^n; \end{aligned}$$

последнее равенство мы получаем, вычитая из каждого столбца предыдущий столбец. Теперь требуемое равенство следует из того, что определитель линейно зависит от каждого столбца.  $\square$

Из доказанной леммы немедленно следует, что

$$P_a(1, x_2, \dots, x_n) = \sum_b P_b(x_2, \dots, x_n), \quad (1.1)$$

если  $a = (0, a_2, \dots, a_n)$ .

Доказательство неотрицательности коэффициентов многочлена  $P_a$  достаточно провести в случае, когда  $a_1 = 0$ . Действительно, поделив  $P_a$  на  $(x_1 \cdots x_n)^{a_1}$ , мы приходим к многочлену  $P_{a'}$ , где  $a'_i = a_i - a_1$ . Применим индукцию по  $n$ . При  $n = 1$  неотрицательность коэффициентов очевидна. Шаг индукции доказывается следующим образом. Формула (1.1) показывает, что коэффициенты многочлена  $P_a(1, x_2, \dots, x_n)$ , который мы рассматриваем как многочлен от переменных  $x_2, \dots, x_n$ , неотрицательны. Но многочлен  $P_a(x_1, \dots, x_n)$  однородный, поэтому он полностью определяется своей степенью и многочленом  $P_a(1, x_2, \dots, x_n)$ .

Перейдём теперь к вычислению суммы коэффициентов многочлена  $P_a$ . Ясно, что

$$V_a(1, x, x^2, \dots, x^{n-1}) = |x^{(i-1)a_j}|_1^n = |x^{(j-1)a_i}|_1^n = V(x^{a_1}, \dots, x^{a_n}).$$

Поэтому  $P_a(1, x, x^2, \dots, x^{n-1})V(1, x, x^2, \dots, x^{n-1}) = V(x^{a_1}, \dots, x^{a_n})$ , т.е.

$$P_a(1, x, x^2, \dots, x^{n-1}) = \frac{V(x^{a_1}, \dots, x^{a_n})}{V(1, x, x^2, \dots, x^{n-1})} = \prod_{i < j} \frac{x^{a_j} - x^{a_i}}{x^j - x^i}.$$

Ясно, что

$$\lim_{x \rightarrow 1} \frac{x^m - x^n}{x - 1} = \lim_{x \rightarrow 1} (x^{m-1} + x^{m-2} + \dots + x + 1) = m - n.$$

Поэтому

$$P_a(1, 1, \dots, 1) = \prod_{i < j} \lim_{x \rightarrow 1} \frac{x^{a_i} - x^{a_j}}{x - 1} \cdot \frac{x - 1}{x^j - x^i} = \prod_{i < j} \frac{a_j - a_i}{j - i}.$$

Остаётся заметить, что  $P_a(1, 1, \dots, 1)$  — это и есть сумма коэффициентов многочлена  $P_a$ .  $\square$

### 1.2.14. Теорема Чеботарёва

С помощью теоремы Митчелла легко получить доказательство следующего утверждения.

**Теорема 1.2.14.1 (Чеботарёв [Ч]).** Пусть  $p$  — простое число и  $\varepsilon = e^{2\pi i/p}$ . Тогда все миноры матрицы Вандермонда  $\|a_{ij}\|_0^{p-1}$ , где  $a_{ij} = \varepsilon^{ij}$ , отличны от нуля.

**Доказательство.** Любой минор рассматриваемой матрицы можно записать в виде  $|\varepsilon^{a_i b_j}|_1^n$ , где  $0 \leq a_1 < a_2 < \dots < a_n \leq p-1$  и  $0 \leq b_1 < b_2 < \dots < b_n \leq p-1$ . Таким образом, в обозначениях теоремы Митчелла нужно доказать, что  $V_a(\varepsilon^{b_1}, \dots, \varepsilon^{b_n}) \neq 0$ . Числа  $\varepsilon^{b_1}, \dots, \varepsilon^{b_n}$  попарно различны, поэтому  $V(\varepsilon^{b_1}, \dots, \varepsilon^{b_n}) \neq 0$ . Следовательно, достаточно доказать, что  $P_a(\varepsilon^{b_1}, \dots, \varepsilon^{b_n}) \neq 0$ .

Предположим, что  $P_a(\varepsilon^{b_1}, \dots, \varepsilon^{b_n}) = 0$ . Сопоставим многочлену  $P_a(x_1, \dots, x_n)$  многочлен  $f(x) = P_a(x^{b_1}, \dots, x^{b_n})$ . Тогда  $f(\varepsilon) = P_a(\varepsilon^{b_1}, \dots, \varepsilon^{b_n}) = 0$  и  $f(1) = P_a(1, \dots, 1)$ . Многочлен  $f$ , как и многочлен  $P_a$ , имеет целые коэффициенты. Многочлен  $x^{p-1} + x^{p-2} + \dots + x + 1$ , корнем которого является  $\varepsilon$ , неприводим (см. [П2], пример 6.2 на с. 62). Поэтому если  $f(\varepsilon) = 0$ , то  $f(x)$  делится на  $x^{p-1} + x^{p-2} + \dots + x + 1$ . В частности,  $f(1)$  делится на  $p$ . Следовательно,  $P_a(1, \dots, 1) = \prod_{i < j} \frac{a_j - a_i}{j - i}$  делится на  $p$ , но при этом ни одно из чисел  $a_j - a_i$  не делится на  $p$ . Получено противоречие.  $\square$

## Задачи

**2.1.** Пусть  $A$  — матрица порядка  $n$ . Докажите, что  $|A + \lambda I| = \lambda^n + \sum_{k=1}^n J_k \lambda^{n-k}$ , где  $J_k$  — сумма всех  $\binom{n}{k}$  главных миноров  $k$ -го порядка матрицы  $A$ .

**2.2.** Докажите, что

$$\begin{vmatrix} a_{11} & \dots & a_{1n} & x_1 \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} & x_n \\ y_1 & \dots & y_n & 0 \end{vmatrix} = - \sum_{i,j} x_i y_j A_{ij},$$

где  $A_{ij}$  — алгебраическое дополнение элемента  $a_{ij}$  в матрице  $\|a_{ij}\|_1^n$

**2.3.** Пусть  $\|a_{ij}\|_1^n$  — симметрическая матрица,

$$A_i(x) = \sum_{s=1}^n a_{is}, \quad A(x, x) = \sum_{s,t=1}^n a_{st} x_s x_t.$$



Докажите, что если  $1 \leq p < n$ , то

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & A_1(x) \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & A_p(x) \\ A_1(x) & \dots & A_n(x) & 0 \end{vmatrix} = -A(x, x)\Delta,$$

где  $\Delta = \begin{vmatrix} a_{11} & \dots & a_{1p} \\ \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} \end{vmatrix}.$

**2.4.** Докажите, что сумма главных  $k$ -миноров матрицы  $A^T A$  равна сумме квадратов всех  $k$ -миноров матрицы  $A$ .

**2.5.** Докажите, что

$$\begin{vmatrix} u_1 a_{11} & \dots & u_n a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + \dots + \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ u_1 a_{n1} & \dots & u_n a_{nn} \end{vmatrix} = (u_1 + \dots + u_n)|A|.$$

**2.6.** На плоскости даны точки  $A, B, C, D, E$  и  $F$ . Докажите, что для ориентированных площадей треугольников выполняется соотношение

$$S_{AEF}S_{DBC} + S_{BEF}S_{DCA} + S_{CEF}S_{DAB} = S_{DEF}S_{ABC}.$$

### Обратные и присоединённые матрицы

**2.7.** Пусть  $A$  и  $B$  — квадратные матрицы порядка  $n$ . Вычислите матрицу  $\begin{pmatrix} I & A & 0 \\ 0 & I & B \\ 0 & 0 & I \end{pmatrix}^{-1}$ .

**2.8.** Докажите, что матрица, обратная к невырожденной верхней треугольной матрице, тоже будет верхней треугольной.

**2.9.** Пусть  $A$  и  $B$  — матрицы размером  $m \times n$  и  $n \times m$ . Докажите, что если матрица  $I_m - AB$  обратима, то матрица  $I_n - BA$  тоже обратима.

**2.10.** Приведите пример матрицы порядка  $n$ , присоединённая к которой имеет лишь один ненулевой элемент, причём этот элемент стоит в  $i$ -й строке и  $j$ -м столбце, где  $i$  и  $j$  заданы.

**2.11.** Матрица  $A$  кососимметрична, причем её порядок равен  $n$ . Докажите, что матрица  $\text{adj } A$  симметрична при нечётном  $n$  и кососимметрична при чётном  $n$ .

**2.12.** Пусть  $A_n$  — кососимметрическая матрица порядка  $n$  с элементами  $-1$  над главной диагональю. Вычислите  $\text{adj } A_n$ .

**2.13.** Матрицу  $\text{adj}(A - \lambda I)$  можно записать в виде суммы  $\sum_{k=0}^{n-1} \lambda^k A_k$ , где  $n$  — порядок матрицы  $A$ . Докажите, что: а) для любого  $k = 1, 2, \dots, n-1$  матрица  $A_k A - A_{k-1}$  диагональна; б) матрица  $A_{n-s}$  выражается как многочлен степени  $s-1$  от  $A$ .

**2.14.** Найдите все матрицы  $A$ , для которых все элементы матриц  $A$  и  $A^{-1}$  неотрицательны.

**2.15.** Докажите, что площадь треугольника, образованного прямыми  $a_i x + b_i y = c_i$ ,  $i = 1, 2, 3$ , равна

$$\pm \frac{1}{2} \frac{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}^2}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} a_3 & b_3 \\ a_1 & b_1 \end{vmatrix}}.$$

**2.16.** Пусть  $\varepsilon$  — первообразный корень  $n$ -й степени из 1;  $A = \|a_{ij}\|_1^n$ , где  $a_{ij} = \varepsilon^{ij}$ . Вычислите матрицу  $A^{-1}$ .

**2.17.** Вычислите матрицу, обратную к матрице Вандермонда

**2.18.** Вычислите матрицу, обратную матрице Коши  $\|a_{ij}\|_1^n$ , где  $a_{ij} = (x_i + y_j)^{-1}$ .

**2.19.** Вычислите матрицу, присоединённую к матрице  $A = \begin{pmatrix} I_n & -[2007/09/17v0.1\text{FontdefinitionsforT1/Newton}]\mathbf{1} \\ 0 & 0 \end{pmatrix}$ , где  $-\mathbf{1}$  — столбец чисел  $-1$ .

**2.20.** Сумма элементов каждой строки матрицы  $A$  равна 0. Докажите, что каждый столбец матрицы  $\text{adj } A$  состоит из равных чисел.

## 1.3. Дополнение по Шуру

### 1.3.1. Определитель блочной матрицы

Квадратную матрицу  $X$  можно записать в блочном виде:  $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ , где  $A$  и  $D$  — квадратные матрицы, сумма порядков которых

равна порядку матрицы  $X$ .

**Теорема 1.3.1.1.** Если  $|A| \neq 0$ , то

$$|X| = |A| \cdot |D - CA^{-1}B|.$$

**Первое доказательство.** Достаточно заметить, что

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & D - CA^{-1}B \end{pmatrix}$$

□

**Второе доказательство.** Матрицу  $X$  можно записать в виде

$$X = \begin{pmatrix} I & 0 \\ U & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} I & V \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & AV \\ UA & UAV + W \end{pmatrix}.$$

Для этого нужно положить  $U = CA^{-1}$ ,  $V = A^{-1}B$  и  $W = D - CA^{-1}B$ . □

Аналогично доказывается следующее утверждение.

**Теорема 1.3.1.2.** Если  $|D| \neq 0$ , то

$$|X| = |D| \cdot |A - BD^{-1}C|.$$

Матрицы  $(X|A) = D - CA^{-1}B$  и  $(X|D) = A - BD^{-1}C$  называют *дополнениями по Шуру* матриц  $A$  и  $D$  в матрице  $X$ .

**Теорема 1.3.1.3.** Если  $A$  и  $D$  — матрицы одного порядка и  $AC = CA$ , то

$$|X| = |AD - CB|.$$

**Доказательство.** Рассмотрим сначала случай, когда  $|A| \neq 0$ . Согласно теореме 1.3.1.1

$$|X| = |AD - ACA^{-1}B| = |AD - CB|.$$

Теперь рассмотрим случай, когда  $|A| = 0$ . Равенство  $|X| = |AD - CB|$  представляет собой полиномиальное тождество для элементов матрицы  $X$ . Поэтому если существуют невырожденные матрицы  $A_\varepsilon$ , для которых  $A_\varepsilon C = CA_\varepsilon$  и  $\lim_{\varepsilon \rightarrow 0} A_\varepsilon = A$ , то требуемое равенство верно и для матрицы  $A$ . Положим  $A_\varepsilon = A + \varepsilon I$ . Ясно, что  $|A + \varepsilon I_n| = \varepsilon^n + \dots$  — многочлен степени  $n$  от  $\varepsilon$ . Этот многочлен имеет конечное число корней, поэтому для почти всех  $\varepsilon$  матрица  $A + \varepsilon I_n$  невырожденная. В частности, эта матрица невырожденная для достаточно малых ненулевых  $\varepsilon$ . □

**Замечание.** По поводу обобщения теоремы 1.3.1.3 см. задачу 1.11.

**Теорема 1.3.1.4.** Если  $a$  — число,  $u$  — строка,  $v$  — столбец, то

$$\begin{vmatrix} A & v \\ u & a \end{vmatrix} = a|A| - u(\text{adj } A)v.$$

**Доказательство.** Предположим сначала, что  $|A| \neq 0$ . Тогда согласно теореме 1.3.1.1

$$\begin{vmatrix} A & v \\ u & a \end{vmatrix} = |A|(a - uA^{-1}v) = a|A| - u(\text{adj } A)v.$$

Обе части полученного равенства непрерывно зависят от элементов матрицы  $A$ , поэтому оно остаётся справедливым и для вырожденных матриц. Действительно, любая вырожденная матрица является пределом невырожденных матриц.  $\square$

Следующее утверждение показывает, что, делая предельные переходы (как в теоремах 1.3.1.3 и 1.3.1.4), нужно соблюдать осторожность.

**Теорема 1.3.1.5.** а) Если  $A$  и  $D$  — матрицы одного порядка,  $|A| \neq 0$  и  $AB^T = -BA^T$ , то

$$|X| = |DA^T + CB^T|.$$

б) Без условия  $|A| \neq 0$  утверждение (а) неверно.

**Доказательство.** а) Если  $|A| \neq 0$ , то

$$\begin{aligned} |X| &= |D - CA^{-1}B| \cdot |A^T| = |DA^T - CA^{-1}BA^T| = \\ &= |DA^T + CA^{-1}AB^T| = |DA^T + CB^T|. \end{aligned}$$

б) Матрицы  $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ ,  $D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  удовлетворяют соотношению  $AB^T = -BA^T$ , но равенство  $|X| = |DA^T + CB^T|$  для них не выполняется.

Это явление связано с тем, что при нашем выборе матрицы  $B$  равенство  $ZB^T = -BZ^T$  выполняется только для матриц вида  $Z = \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}$ .

Это лишает нас возможности приблизить матрицу  $A$  невырожденными матрицами, удовлетворяющими такому равенству.  $\square$

### 1.3.2. Обращение блочной матрицы

Дополнение по Шуру естественным образом возникает также при блочной записи обратной матрицы.

**Теорема 1.3.2.1.** Пусть  $A$  и  $D$  — невырожденные матрицы. Тогда

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix},$$

где  $A_1^{-1} = A - BD^{-1}C$  и  $D_1^{-1} = D - CA^{-1}B$ .

**Доказательство.**

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} = \begin{pmatrix} AA_1 + BC_1 & AB_1 + BD_1 \\ CA_1 + DC_1 & CB_1 + DD_1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

поэтому  $AB_1 + BD_1 = 0$  и  $CA_1 + DC_1 = 0$ , а значит,  $B_1 = -A^{-1}BD_1$  и  $C_1 = -D^{-1}CA_1$ . Далее,  $AA_1 + BC_1 = I$  и  $CB_1 + DD_1 = I$ . Следовательно,  $AA_1 - BD^{-1}CA_1 = I$  и  $-CA^{-1}BD_1 + DD_1 = I$ , т.е.  $A_1 = (A - BD^{-1}C)^{-1}$  и  $D_1 = (D - CA^{-1}B)^{-1}$ .  $\square$

**Замечание.** Если  $A$  и  $D$  — невырожденные матрицы, то аналогично получаем  $D_1 = -B^{-1}AB_1$  и  $A_1 = -C^{-1}DC_1$ . Далее,  $-AC^{-1}DC_1 + BC_1 = I$  и  $CB_1 - DB^{-1}AB_1 = I$ , т.е.  $C_1 = (B - AC^{-1}D)^{-1}$  и  $B_1 = (C - DB^{-1}A)^{-1}$ .

### 1.3.3. Теорема Хэйнсворт

Пусть  $A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$ ,  $B = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  и  $C = A_{11}$  — квадратные матрицы, причём матрицы  $B$  и  $C$  невырождены. Матрицу  $(B|C) = A_{22} - A_{21}A_{11}^{-1}A_{12}$  можно рассматривать как подматрицу матрицы  $(A|C) = \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} - \begin{pmatrix} A_{21} \\ A_{31} \end{pmatrix} A_{11}^{-1} (A_{12} \ A_{13})$ .

**Теорема 1.3.3.1 (Хэйнсворт).**  $(A|B) = ((A|C)|(B|C))$ .

**Доказательство.** [Os] Для матрицы  $A$  можно записать следующие разложения:

$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{pmatrix} \begin{pmatrix} I & * & * \\ 0 & (A|C) & \\ 0 & & \end{pmatrix}, \quad (1.1)$$

$$A = \begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & I \end{pmatrix} \begin{pmatrix} I & 0 & * \\ 0 & I & * \\ 0 & 0 & (A|B) \end{pmatrix}. \quad (1.2)$$

Рассматривая дополнение по Шуру матрицы  $A_{11}$  в левом сомножителе разложения (1.2), можно записать

$$\begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & I \end{pmatrix} = \begin{pmatrix} A_{11} & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{pmatrix} \begin{pmatrix} I & X_1 & X_2 \\ 0 & X_3 & X_4 \\ 0 & X_5 & X_6 \end{pmatrix} = \begin{pmatrix} * & * & A_{11}X_2 \\ * & * & A_{21}X_2 + X_4 \\ * & * & A_{31}X_2 + X_6 \end{pmatrix}. \quad (1.3)$$

Следовательно,  $A_{11}X_2 = 0$ , а так как матрица  $A_{11}$  невырождена, то  $X_2 = 0$ . Поэтому  $0 = A_{21}X_2 + X_4 = X_4$  и  $I = A_{31}X_2 + X_6 = X_6$ . Ясно также, что  $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ A_{21} & I \end{pmatrix} \begin{pmatrix} I & X_1 \\ 0 & X_3 \end{pmatrix}$ , т.е.  $X_3 = (B|C)$ . Подставим теперь разложение (1.3) в (1.2) и сравним полученное выражение с (1.1):

$$\begin{pmatrix} A_{11} & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{pmatrix} \begin{pmatrix} I & * & * \\ 0 & (A|C) & \end{pmatrix} = \begin{pmatrix} A_{11} & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{pmatrix} \begin{pmatrix} I & * & 0 \\ 0 & (B|C) & 0 \\ 0 & * & I \end{pmatrix} \begin{pmatrix} I & 0 & * \\ 0 & I & * \\ 0 & 0 & (A|B) \end{pmatrix}.$$

Сокращая обе части этого равенства слева на невырожденную матрицу, получаем, в частности,

$$(A|C) = \begin{pmatrix} (B|C) & 0 \\ * & I \end{pmatrix} \begin{pmatrix} I & * \\ 0 & (A|B) \end{pmatrix}.$$

Разложение такого вида единственно.  $\square$

### 1.3.4. Пфаффиан

Используя дополнение по Шуру, можно получить простое доказательство того, что определитель кососимметрической матрицы  $A$  чётного порядка  $n$  является квадратом некоторого многочлена  $\text{Pf}(A)$  (называемого *пфаффианом*) от переменных  $a_{ij}$ ,  $1 \leq i < j \leq n$ . Такое доказательство приведено в [Le].

Доказательство проведём индукцией по  $n$ . При  $n = 2$  утверждение очевидно. Запишем матрицу  $A$  в виде  $A = \begin{pmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{pmatrix}$ , где  $A_{11} = \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix}$ . Мы рассматриваем  $a_{ij}$  как независимые переменные, поэтому можно считать, что  $a_{12} \neq 0$ . В таком случае  $|A| = |A_{11}| \cdot |A_{22} - A_{21}A_{11}^{-1}A_{12}|$  (теорема 1.3.1.1). Легко проверить, что матрица  $B = A_{22} -$

$A_{21}A_{11}^{-1}A_{12}$  кососимметрична. Поэтому  $|A| = (a_{12} \operatorname{Pf}(B))^2$ , где  $\operatorname{Pf}(B)$  по предположению индукции — многочлен от переменных  $a_{12}^{-1}$  и  $a_{ij}$ . Но мы знаем, что  $|A|$  — многочлен от переменных  $a_{ij}$ . Поэтому  $a_{12} \operatorname{Pf}(B)$  — многочлен от переменных  $a_{ij}$ .

### Задачи

**3.1.** Пусть  $A$  — квадратная матрица. Докажите, что

$$\begin{vmatrix} I & A \\ A^T & I \end{vmatrix} = 1 - \sum M_1^2 + \sum M_2^2 - \sum M_3^2 + \dots,$$

где  $\sum M_k^2$  — сумма квадратов всех миноров порядка  $k$  матрицы  $A$ .

**3.2.** Пусть  $A$  — матрица размером  $n \times m$ ,  $B$  — матрица размером  $m \times n$ . Докажите, что

$$|I_n - AB| = |I_m - BA|.$$

**3.3.** Пусть  $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  и  $A^{-1} = \begin{pmatrix} A'_{11} & A'_{12} \\ A'_{21} & A'_{22} \end{pmatrix}$ , где  $A_{11}$  и  $A'_{11}$  — матрицы порядка  $m$ ,  $A_{22}$  и  $A'_{22}$  — матрицы порядка  $n$ . Докажите, что  $\det A \det A'_{22} = \det A_{11}$ .

**3.4.** Пусть

$$A_4 = \begin{pmatrix} 1 & 2 & 4 & 7 \\ 3 & 5 & 8 & 11 \\ 6 & 9 & 12 & 14 \\ 10 & 13 & 15 & 16 \end{pmatrix};$$

матрица  $A_n$  для произвольного  $n$  определяется аналогично (на диагоналях, перпендикулярных главной, стоят последовательные натуральные числа). Докажите, что  $|A_{2k}| = (-1)^k k(k+1)$  и  $|A_{2k+1}| = (-1)^k (2k^2 + 2k + 1)$ .

## 1.4. Симметрические функции. Суммы степеней. Числа Бернулли

Этот параграф посвящён различным детерминантным выражениям для симметрических функций, сумм степеней и чисел Бернулли. Большинство из этих выражений возникает из соотношений, являющихся системами линейных уравнений, при решении этих систем по правилу

Крамера. В большинстве из рассматриваемых случаев однородной части системы линейных уравнений соответствует треугольная матрица, поэтому в выражении остаётся только один определитель.

Для получения требуемых соотношений используются некоторые известные факты о симметрических функциях и числах Бернулли. Все эти факты приведены с полными доказательствами.

### 1.4.1. Примеры симметрических многочленов

Многочлен  $f(x_1, \dots, x_n)$  называют *симметрическим*, если для любой подстановки  $\sigma \in S_n$  выполняется равенство

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n).$$

Основным примером симметрических многочленов служат *элементарные* симметрические многочлены

$$\sigma_k(x_1, \dots, x_n) = \sum_{i_1 < \dots < i_k} x_{i_1} \cdots x_{i_k},$$

где  $1 \leq k \leq n$ ; удобно считать, что  $\sigma_0 = 1$  и  $\sigma_k(x_1, \dots, x_n) = 0$  при  $k > n$ .

Элементарные симметрические многочлены можно задавать с помощью *производящей функции*

$$\sigma(t) = \sum_{k=0}^{\infty} \sigma_k t^k = \prod_{i=1}^n (1 + tx_i).$$

Если  $x_1, \dots, x_n$  — корни многочлена  $x^n + a_1 x^{n-1} + \dots + a_n$ , то  $\sigma_k(x_1, \dots, x_n) = (-1)^k a_k$ .

Другим примером симметрических многочленов служат *полные однородные* симметрические многочлены

$$p_k(x_1, \dots, x_n) = \sum_{i_1 + \dots + i_n = k} x_1^{i_1} \cdots x_n^{i_n}.$$

Им соответствует производящая функция

$$p(t) = \sum_{k=0}^{\infty} p_k t^k = \prod_{i=1}^n (1 - tx_i)^{-1}.$$

Важным примером симметрических многочленов служат также *степенные суммы*

$$s_k(x_1, \dots, x_n) = x_1^k + \dots + x_n^k.$$



Им соответствует производящая функция

$$s(t) = \sum_{k=0}^{\infty} s_k t^{k-1} = \sum_{i=1}^n x_i (1 - tx_i)^{-1}.$$

Иногда используются *мономиальные* симметрические многочлены

$$m_{i_1 \dots i_n}(x_1, \dots, x_n) = \sum_{\sigma \in S_n} x_{\sigma(1)}^{i_1} \cdots x_{\sigma(n)}^{i_n}.$$

Производящие функции  $\sigma(t)$  и  $p(t)$  связаны соотношением  $\sigma(t)p(-t) = 1$ . Приравнявая коэффициенты при  $t^n$ ,  $n \geq 1$ , в левой и правой части, получаем

$$\sum_{r=0}^n (-1)^r \sigma_r p_{n-r} = 0. \quad (1)$$

Производящая функция  $s(t)$  выражается через  $p(t)$  и  $\sigma(t)$  следующим образом:

$$s(t) = \frac{d}{dt} \ln p(t) = \frac{p'(t)}{p(t)}, \quad \text{т. е.} \quad s(t)p(t) = p'(t);$$

$$s(-t) = -\frac{d}{dt} \ln \sigma(t) = -\frac{\sigma'(t)}{\sigma(t)}, \quad \text{т. е.} \quad s(-t)\sigma(t) = -\sigma'(t).$$

Приравнявая коэффициенты при  $t^{n+1}$ , получаем

$$np_n = \sum_{r=1}^n s_r p_{n-r}, \quad (2)$$

$$n\sigma_n = \sum_{r=1}^n (-1)^{r-1} s_r \sigma_{n-r}. \quad (3)$$

Соотношения (3) называют *формулами Ньютона*.

Запишем соотношения (1) для  $n = 1, \dots, k$ . При фиксированных  $\sigma_1, \dots, \sigma_k$  эти соотношения можно рассматривать как систему линейных уравнений для  $p_1, \dots, p_k$ , а при фиксированных  $p_1, \dots, p_k$  — как систему уравнений для  $\sigma_1, \dots, \sigma_k$ . Решая эти системы, находим

$$\sigma_k = \begin{vmatrix} p_1 & 1 & 0 & \dots & 0 \\ p_2 & p_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ p_{k-1} & p_{k-2} & \dots & \dots & 1 \\ p_k & p_{k-1} & \dots & \dots & p_1 \end{vmatrix}, \quad p_k = \begin{vmatrix} \sigma_1 & 1 & 0 & \dots & 0 \\ \sigma_2 & \sigma_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_{k-1} & \sigma_{k-2} & \dots & \dots & 1 \\ \sigma_k & \sigma_{k-1} & \dots & \dots & \sigma_1 \end{vmatrix}.$$

Аналогично с помощью соотношений (2) получаем

$$s_k = (-1)^{k-1} \begin{vmatrix} p_1 & 1 & 0 & \dots & 0 \\ 2p_2 & p_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ kp_k & p_{k-1} & p_{k-2} & \dots & p_1 \end{vmatrix},$$

$$p_k = \frac{1}{k!} \begin{vmatrix} s_1 & -1 & 0 & \dots & 0 \\ s_2 & s_1 & -2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ s_{k-1} & s_{k-2} & \dots & \dots & -k+1 \\ s_k & s_{k-1} & \dots & \dots & s_1 \end{vmatrix}.$$

С помощью соотношений (3) получаем

$$s_k = \begin{vmatrix} \sigma_1 & 1 & 0 & \dots & 0 \\ 2\sigma_2 & \sigma_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ k\sigma_k & \sigma_{k-1} & \sigma_{k-2} & \dots & \sigma_1 \end{vmatrix}, \quad \sigma_k = \frac{1}{k!} \begin{vmatrix} s_1 & 1 & 0 & \dots & 0 \\ s_2 & s_1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ s_{k-1} & s_{k-2} & \dots & \dots & k-1 \\ s_k & s_{k-1} & \dots & \dots & s_1 \end{vmatrix}.$$

### 1.4.2. Основная теорема о симметрических многочленах

**Теорема 1.4.2.1.** Пусть  $f(x_1, \dots, x_n)$  — симметрический многочлен. Тогда существует такой многочлен  $g(y_1, \dots, y_n)$ , что  $f(x_1, \dots, x_n) = g(\sigma_1, \dots, \sigma_n)$ . При этом многочлен  $g$  единствен.

**Доказательство.** Достаточно рассмотреть случай, когда  $f$  — однородный многочлен. Будем говорить, что моном  $x_1^{\lambda_1} \dots x_n^{\lambda_n}$  имеет более высокий порядок, чем моном  $x_1^{\mu_1} \dots x_n^{\mu_n}$ , если  $\lambda_1 = \mu_1, \dots, \lambda_k = \mu_k$  и  $\lambda_{k+1} > \mu_{k+1}$  (возможно  $k = 0$ ). Пусть  $ax_1^{\lambda_1} \dots x_n^{\lambda_n}$  — старший моном многочлена  $f$ . Тогда  $\lambda_1 \geq \dots \geq \lambda_n$ . Рассмотрим симметрический многочлен

$$f_1 = f - a\sigma_1^{\lambda_1 - \lambda_2} \sigma_2^{\lambda_2 - \lambda_3} \dots \sigma_n^{\lambda_n}. \quad (1)$$

Старший член монома  $\sigma_1^{\lambda_1 - \lambda_2} \dots \sigma_n^{\lambda_n}$  равен

$$x_1^{\lambda_1 - \lambda_2} (x_1 x_2)^{\lambda_2 - \lambda_3} \dots (x_1 \dots x_n)^{\lambda_n} = x_1^{\lambda_1} x_2^{\lambda_2} \dots x_n^{\lambda_n},$$

поэтому порядок старшего монома многочлена  $f_1$  строго ниже порядка старшего монома многочлена  $f$ . Применим к многочлену  $f_1$  снова операцию (1) и т.д. Ясно, что после конечного числа таких операций придём к нулевому многочлену.

Докажем теперь единственность представления  $f(x_1, \dots, x_n) = g(\sigma_1, \dots, \sigma_n)$ . Достаточно проверить, что если  $g(y_1, \dots, y_n) = \sum a_{i_1 \dots i_n} y_1^{i_1} \dots y_n^{i_n}$  — ненулевой многочлен, то после подстановки  $y_1 = \sigma_1 = x_1 + \dots + x_n, \dots, y_n = \sigma_n = x_1 \dots x_n$  этот многочлен останется ненулевым. Ограничимся рассмотрением старших мономов

$$a_{i_1 \dots i_n} x_1^{i_1 + \dots + i_n} x_2^{i_2 + \dots + i_n} \dots x_n^{i_n},$$

получающихся в результате подстановки. Ясно, что самый старший среди этих мономов ни с чем сократиться не может.  $\square$

Из доказательства теоремы 1.4.2.1 видно, что если  $f(x_1, \dots, x_n)$  — симметрический многочлен с целыми коэффициентами, то  $f(x_1, \dots, x_n) = g(\sigma_1, \dots, \sigma_n)$ , где  $g$  — тоже многочлен с целыми коэффициентами. Детерминантное выражение  $\sigma_k$  через  $p_1, \dots, p_k$  показывает, что для полных однородных многочленов справедливо аналогичное утверждение. Что же касается степенных сумм, то для них выражение вида  $f(x_1, \dots, x_n) = g(s_1, \dots, s_n)$  тоже существует, но при этом коэффициенты многочлена  $g$  не обязательно целые. Например,

$$x_1 x_2 = \frac{(x_1 + x_2)^2 - (x_1^2 + x_2^2)}{2} = \frac{s_1^2 - s_2}{2}.$$

Из основной теоремы о симметрических многочленах следует, что если  $x_1, \dots, x_n$  — корни многочлена  $x^n + a_1 x^{n-1} + \dots + a_n$ , то величина

$$D = \prod_{i < j} (x_i - x_j)^2,$$

представляющая собой симметрический многочлен от  $x_1, \dots, x_n$ , полиномиально выражается через  $a_1, \dots, a_n$ . Эту величину называют *дискриминантом* многочлена.

Назовём многочлен  $f(x_1, \dots, x_n)$  *кососимметрическим*, если

$$f(\dots x_i \dots x_j \dots) = f(\dots x_j \dots x_i \dots),$$

т. е. при транспозиции любых двух переменных  $x_i$  и  $x_j$  многочлен меняет знак. Примером кососимметрического многочлена служит  $\Delta = \prod_{i < j} (x_i - x_j)$ . Ясно, что  $\Delta^2 = D$ .

**Теорема 1.4.2.2.** *Любой кососимметрический многочлен  $f(x_1, \dots, x_n)$  можно представить в виде*

$$\Delta(x_1, \dots, x_n) g(x_1, \dots, x_n),$$

где  $g$  — симметрический многочлен.

**Доказательство.** Достаточно проверить, что  $f$  делится на  $\Delta$ . В самом деле, если  $f/\Delta$  — многочлен, то этот многочлен по очевидным причинам симметрический. Покажем, например, что  $f$  делится на  $x_1 - x_2$ . Сделаем замену  $x_1 = u + v$ ,  $x_2 = u - v$ . В результате получим

$$f(x_1, x_2, x_3, \dots, x_n) = f_1(u, v, x_3, \dots, x_n).$$

Если  $x_1 = x_2$ , то  $u = 0$ . Поэтому  $f_1(0, v, x_3, \dots, x_n) = 0$ . Это означает, что многочлен  $f_1$  делится на  $u$ , т. е. многочлен  $f$  делится на  $x_1 - x_2$ . Аналогично доказывается, что  $f$  делится на  $x_i - x_j$  при всех  $i < j$ .  $\square$

Равенство  $\Delta^2 = D$  показывает, что представление  $f = \Delta g$  не единственно.

### 1.4.3. Сумма степеней

Перейдём к рассмотрению свойств суммы степеней первых членов натурального ряда. Пусть  $S_n(k) = 1^n + 2^n + \dots + (k-1)^n$ . Докажем, что

$$S_{n-1}(k) = \frac{1}{n!} \begin{vmatrix} k^n & \binom{n}{n-2} & \binom{n}{n-3} & \dots & \binom{n}{1} & 1 \\ k^{n-1} & \binom{n-1}{n-2} & \binom{n-1}{n-3} & \dots & \binom{n-1}{1} & 1 \\ k^{n-2} & 0 & \binom{n-2}{n-3} & \dots & \binom{n-2}{1} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ k & 0 & 0 & \dots & 0 & 1 \end{vmatrix}.$$

Заметим сначала, что этот определитель не изменится, если из первого столбца вычесть последний.

Сложим равенства  $(x+1)^m - x^m = \sum_{i=0}^{m-1} \binom{m}{i} x^i$  для  $x = 1, 2, \dots, k-1$ . В итоге получим  $k^m - 1 = \sum_{i=0}^{m-1} \binom{m}{i} S_i(k)$ . Набор таких равенств для  $m = 1, 2, \dots, n$  можно рассмотреть как систему линейных уравнений относительно  $S_i(k)$ . Из этой системы получаем требуемое выражение для  $S_{n-1}(k)$ .

Из полученного выражения для  $S_{n-1}(k)$  видно, что  $S_{n-1}(k)$  — многочлен степени  $n$  от  $k$ .

### 1.4.4. Теорема Фаульгабера—Якоби

Фаульгабер в своей книге [Fa1] отметил, что все суммы  $S_n(k)$  полиномиально выражаются через первые две суммы  $S_1(k)$  и  $S_2(k)$ . Двести лет спустя, в 1834 г., Якоби переоткрыл теорему Фаульгабера.

**Теорема 1.4.4.1 (Фаульгабер–Якоби).** Пусть  $U = S_1(x)$  и  $V = S_2(x)$ . Тогда при  $k \geq 1$  существуют такие многочлены  $P_k$  и  $Q_k$  с рациональными коэффициентами, что  $S_{2k+1}(x) = U^2 P_k(U)$  и  $S_{2k}(x) = V Q_k(U)$ .

**Доказательство.** Чтобы получить выражение для  $S_{2k+1}$ , воспользуемся равенством

$$\begin{aligned} [n(n-1)]^r &= \sum_{x=1}^{n-1} (x^r(x+1)^r - x^r(x-1)^r) = \\ &= 2 \left[ \binom{r}{1} \sum x^{2r-1} + \binom{r}{3} \sum x^{2r-3} + \binom{r}{5} \sum x^{2r-5} + \dots \right], \end{aligned} \quad (1)$$

т. е.

$$[n(n-1)]^{i+1} = \sum \binom{i+1}{2(i-j)+1} S_{2j+1}(n).$$

Эти равенства для можно записать в матричном виде:

$$\begin{pmatrix} n^2(n-1)^2 \\ n^3(n-1)^3 \\ n^4(n-1)^4 \\ \vdots \end{pmatrix} = 2 \begin{pmatrix} 2 & 0 & 0 & \dots \\ 1 & 3 & 0 & \dots \\ 0 & 4 & 4 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} S_3(n) \\ S_5(n) \\ S_7(n) \\ \vdots \end{pmatrix}.$$

В полученной бесконечной матрице все главные миноры конечного порядка невырожденные, поэтому

$$\begin{pmatrix} S_3(n) \\ S_5(n) \\ S_7(n) \\ \vdots \end{pmatrix} = \frac{1}{2} \|a_{ij}\|^{-1} \begin{pmatrix} n^2(n-1)^2 \\ n^3(n-1)^3 \\ n^4(n-1)^4 \\ \vdots \end{pmatrix}, \quad \text{где } a_{ij} = \binom{i+1}{2(i-j)+1}.$$

Эта формула показывает, что  $S_{2k+1}(n)$  выражается через  $n(n-1) = 2U(n)$  и делится на  $[n(n-1)]^2$ .

Чтобы получить выражение для  $S_{2k}$ , воспользуемся равенством

$$\begin{aligned}
 n^{r+1}(n-1)^r &= \sum_{x=1}^{n-1} (x^r(x+1)^{r+1} - (x-1)^r x^{r+1}) = \\
 &= \sum x^{2r} \left[ \binom{r+1}{1} + \binom{r}{1} \right] + \sum x^{2r-1} \left[ \binom{r+1}{2} - \binom{r}{2} \right] + \\
 &+ \sum x^{2r-2} \left[ \binom{r+1}{3} + \binom{r}{3} \right] + \sum x^{2r-3} \left[ \binom{r+1}{4} - \binom{r}{4} \right] + \\
 + \dots &= \left[ \binom{r+1}{1} + \binom{r}{1} \right] \sum x^{2r} + \left[ \binom{r+1}{3} + \binom{r}{3} \right] \sum x^{2r-2} + \\
 + \dots &+ \binom{r}{1} \sum x^{2r-1} + \binom{r}{3} \sum x^{2r-3} + \dots
 \end{aligned}$$

Суммы нечётных степеней можно уничтожить с помощью (1). В результате получим

$$\begin{aligned}
 n^{r+1}(n-1)^r &= \frac{n^r(n-1)^r}{2} + \left[ \binom{r+1}{1} + \binom{r}{1} \right] \sum x^{2r} + \\
 &+ \left[ \binom{r+1}{3} + \binom{r}{3} \right] \sum x^{2r-2} + \dots,
 \end{aligned}$$

т. е.

$$n^i(n-1)^i \left( \frac{2n-1}{2} \right) = \sum \left[ \binom{i+1}{2(i-j)+1} + \binom{i}{2(i-j)+1} \right] S_{2j}(n).$$

Теперь аналогично предыдущему случаю получаем

$$\begin{pmatrix} S_2(n) \\ S_4(n) \\ S_6(n) \\ \vdots \end{pmatrix} = \frac{2n-1}{2} \|b_{ij}\|^{-1} \begin{pmatrix} n(n-1) \\ n^2(n-1)^2 \\ n^3(n-1)^3 \\ \vdots \end{pmatrix},$$

$$\text{где } b_{ij} = \binom{i+1}{2(i-j)+1} + \binom{i}{2(i-j)+1}.$$

Простые вычисления показывают, что  $S_2(n) = \frac{2n-1}{2} \cdot \frac{n(n-1)}{3}$ . Поэтому многочлены  $S_4(n)$ ,  $S_6(n)$ , ... делятся на  $S_2(n)$ , причем  $S_{2k}(n)/S_2(n)$  является многочленом от  $n(n-1) = 2U(n)$ .  $\square$

### 1.4.5. Числа Бернулли

Во многих теоремах анализа и теории чисел встречаются *числа Бернулли*  $B_k$ , возникающие при разложении в ряд функции

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!} \quad (\text{при } |x| < 2\pi).$$

Легко проверить, что  $B_0 = 1$  и  $B_1 = -1/2$ . С помощью чисел Бернулли можно представить  $S_k(n) = 1^k + 2^k + \dots + (n-1)^k$  в виде многочлена от  $n$ .

**Теорема 1.4.5.1.**  $(m+1)S_m(n) = \sum_{k=0}^m \binom{m+1}{k} B_k n^{m+1-k}$ .

**Доказательство.** Представим произведение рядов  $\sum_{k=0}^{\infty} \frac{B_k x^k}{k!} \sum_{s=1}^{\infty} \frac{(nx)^s}{s!}$  в виде ряда двумя способами. С одной стороны, это произведение равно

$$\begin{aligned} \frac{x}{e^x - 1} (e^{nx} - 1) &= x \sum_{r=1}^{n-1} e^{rx} = nx + \sum_{m=1}^{\infty} \left( \sum_{r=1}^{n-1} r^m \right) \frac{x^{m+1}}{m!} = \\ &= nx + \sum_{m=1}^{\infty} \frac{(m+1)S_m(n)}{(m+1)!} x^{m+1}. \end{aligned}$$

С другой стороны, это произведение равно

$$\sum_{k=0, s=1}^{\infty} \frac{B_k n^s x^{s+k}}{k! s!} = B_0 nx + \sum_{m=1}^{\infty} \sum_{k=0}^m \frac{\binom{m+1}{k} B_k n^{m+1-k}}{(m+1)!} x^{m+1}.$$

□

Приведём некоторые детерминантные выражения для  $B_k$ . Пусть  $b_k = B_k/k!$ . Тогда согласно определению

$$\begin{aligned} x &= (e^x - 1) \left( \sum_{k=0}^{\infty} b_k x^k \right) = \\ &= \left( x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \right) (1 + b_1 x + b_2 x^2 + b_3 x^3 + \dots), \end{aligned}$$

т.е.

$$\begin{aligned} b_1 &= -\frac{1}{2!}, \\ \frac{b_1}{2!} + b_2 &= -\frac{1}{3!}, \\ \frac{b_1}{3!} + \frac{b_2}{2!} + b_3 &= -\frac{1}{4!}, \\ \dots &\dots \end{aligned}$$

Решая эту систему линейных уравнений по правилу Крамера, получаем

$$B_k = k!b_k = (-1)^k k! \begin{vmatrix} 1/2! & 1 & 0 & \dots & 0 \\ 1/3! & 1/2! & 1 & \dots & 0 \\ 1/4! & 1/3! & 1/2! & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1/(k+1)! & 1/k! & 1/(k-1)! & \dots & 1/2! \end{vmatrix}.$$

Докажем теперь, что  $B_{2k+1} = 0$  при  $k \geq 1$ . Пусть  $\frac{x}{e^x-1} = -\frac{x}{2} + f(x)$ . Тогда  $f(x) - f(-x) = \frac{x}{e^x-1} + \frac{x}{e^{-x}-1} + x = 0$ , т.е.  $f$  — чётная функция. Пусть  $c_k = b_{2k}/(2k)!$ . Тогда

$$x = \left( x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \right) \left( 1 - \frac{x}{2} + c_1 x^2 + c_2 x^4 + c_3 x^6 + \dots \right).$$

Приравнивая коэффициенты при  $x^3, x^5, x^7, \dots$  и учитывая, что  $\frac{1}{2(2n)!} - \frac{1}{(2n+1)!} = \frac{2n-1}{2(2n+1)!}$ , получаем

$$\begin{aligned} c_1 &= -\frac{1}{2 \cdot 3!}, \\ \frac{c_1}{3!} + c_2 &= -\frac{3}{2 \cdot 5!}, \\ \frac{c_1}{5!} + \frac{c_2}{3!} + c_3 &= -\frac{5}{2 \cdot 7!}, \\ \dots &\dots \end{aligned}$$

Следовательно,

$$\begin{aligned} B_{2k} &= (2k)!c_k = \\ &= \frac{(-1)^{k+1}(2k)!}{2} \begin{vmatrix} 1/3! & 1 & 0 & 0 & \dots & 0 \\ 3/5! & 1/3! & 1 & 0 & \dots & 0 \\ 5/7! & 1/5! & 1/3! & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ (2k-1)/(2k+1)! & 1/(2k-1)! & \dots & \dots & \dots & 1/3! \end{vmatrix} \end{aligned}$$

#### 1.4.6. Функции Шура

Упорядоченный набор  $\lambda$  целых неотрицательных чисел  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$  называют *разбиением* числа  $|\lambda| = \lambda_1 + \dots + \lambda_n$ . Число  $n$  называют при этом *длиной разбиения*.

Разбиению  $\lambda$  можно сопоставить *диаграмму Юнга*, состоящую из  $|\lambda|$  клеток, расположенных таким образом, что в  $i$ -й строке стоит  $\lambda_i$  клеток, причём первые клетки всех строк находятся в первом столбце. Например, разбиению  $(4, 2, 1)$  соответствует диаграмма Юнга, изображённая



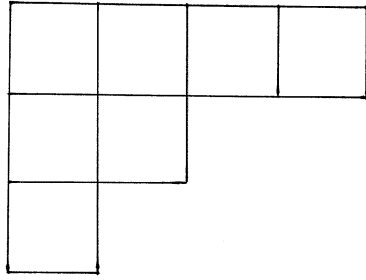


Рис. 1.1.

на рис. 1.1. Диаграмму Юнга, соответствующую разбиению  $\lambda$ , мы будем тоже обозначать  $\lambda$ .

Разбиение  $\lambda'$  называют *сопряжённым* к  $\lambda$ , если его диаграмма Юнга получается из диаграммы Юнга разбиения  $\lambda$  транспонированием относительно главной диагонали.

Для диаграмм Юнга, у которых количество клеток на главной диагонали мало по сравнению с количеством всех клеток, удобно использовать следующее обозначение Фробениуса. Диаграмму Юнга, у которой на главной диагонали находится  $k$  клеток, обозначим  $(\alpha_1, \dots, \alpha_k | \beta_1, \dots, \beta_k)$ , где  $\alpha_i$  — количество клеток в  $i$ -м столбце, стоящих ниже клетки главной диагонали, а  $\beta_i$  — количество клеток в  $i$ -й строке, стоящих правее клетки главной диагонали. Например, в этих обозначениях разбиение  $(4, 2, 1)$  записывается так:  $(2, 0 | 3, 0)$ .

Пусть  $a_{ij} = p_{j-i}$ , где  $p_i = p_i(x_1, \dots, x_n)$  — полный однородный многочлен степени  $n$ . Рассмотрим бесконечную матрицу

$$P = \|a_{ij}\|_0^\infty = \begin{vmatrix} p_0 & p_1 & p_2 & \dots \\ 0 & p_0 & p_1 & \dots \\ 0 & 0 & p_0 & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}.$$

*Функцией Шура*, или *S-функцией*, соответствующей разбиению  $\lambda$ , называют минор матрицы  $P$ , образованный строками  $0, 1, \dots, n-1$  и столбцами  $\lambda_n, \lambda_{n-1} + 1, \dots, \lambda_1 + n - 1$ . Эту симметрическую функцию от переменных  $x_1, \dots, x_n$  обозначают  $s_\lambda$ .

Например,  $s_{1,1} = p_1^2 - p_2$ ,  $s_2 = p_2$ . Легко видеть, что при добавлении к разбиению  $\lambda$  нескольких нулей выражение  $s_\lambda$  через  $p_i$  не меняется; меняется только число переменных  $x_1, \dots, x_n$ .

В виде определителя функция  $s_\lambda$  записывается следующим образом:

$$s_\lambda = |p_{\lambda_{n+1-j}+j-i}|_1^n = |p_{\lambda_j+i-j}|_1^n;$$

при записи последнего равенства мы воспользовались задачей 1.30. Переходя к транспонированной матрице, можно также написать  $s_\lambda = |p_{\lambda_i+j-i}|_1^n$ .

*Косой функцией Шура*  $s_{\lambda,\mu}$ , соответствующей паре разбиений  $\lambda$  и  $\mu$ , называют минор матрицы  $P$ , образованный строками  $\mu_n, \mu_{n-1} + 1, \dots, \mu_1 + n - 1$  и столбцами  $\lambda_n, \lambda_{n-1} + 1, \dots, \lambda_1 + n - 1$ . Ясно, что при этом  $s_\lambda = s_{\lambda,0}$ . В виде определителя косая функция Шура записывается следующим образом:

$$s_{\lambda,\mu} = |p_{\lambda_i-\mu_j+j-i}|_1^n.$$

Разбиение  $\mu$  называют *подразбиением* разбиения  $\lambda$ , если  $\lambda_i \geq \mu_i$  при  $i = 1, \dots, n$ . Если  $\mu$  не является подразбиением  $\lambda$ , то  $s_{\lambda,\mu} = 0$ . Действительно, если  $\lambda_i < \mu_i$ , то матрица, образованная указанными строками и столбцами, имеет вид  $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ , где  $A$  и  $C$  — квадратные матрицы, причём первый столбец матрицы  $C$  (он соответствует столбцу  $\lambda_i + n - i$ ) нулевой.

Первоначально функции Шура (еще до Шура) были введены Якоби как отношения кососимметрических функций определенного вида. Пусть  $\alpha = (\alpha_1, \dots, \alpha_n)$  — некоторое разбиение,  $a_\alpha$  — антисимметризация одночлена  $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ , т. е.

$$a_\alpha = \sum_{\omega \in S_n} (-1)^\omega \omega(x^\alpha),$$

где  $(-1)^\omega$  — знак подстановки  $\omega$  и  $\omega(x^\alpha) = x_{\omega(1)}^{\alpha_1} \cdots x_{\omega(n)}^{\alpha_n}$ . Ясно, что многочлен  $a_\alpha(x_1, \dots, x_n)$  равен определителю  $|x_i^{\alpha_j}|_1^n$ ; в частности, этот многочлен кососимметрический. Поэтому если  $\alpha_i = \alpha_{i+1}$  для некоторого  $i$ , то  $a_\alpha = 0$ . Таким образом, можно считать, что  $\alpha = \lambda + \delta$ , где  $\delta = (n-1, n-2, \dots, 1, 0)$ .

**Теорема 1.4.6.1 (тождество Якоби–Труди).** Пусть  $\delta = (n-1, n-2, \dots, 1, 0)$ . Тогда

$$s_\lambda = a_{\lambda+\delta}/a_\delta, \quad \text{т. е.} \quad |p_{\lambda_i+j-i}|_1^n = |x_i^{\lambda_j+n-j}|_1^n / |x_i^{n-j}|_1^n.$$

**Доказательство.** Пусть  $\alpha = (\alpha_1, \dots, \alpha_n)$  — некоторое разбиение. Рассмотрим матрицы  $A_\alpha = \|x_j^{\alpha_i}\|_1^n$  и  $H_\alpha = \|p_{\alpha_i-n+j}\|_1^n$ . Рассмотрим также

матрицу  $M = \|(-1)^{n-i} \sigma_{n-i}(\hat{x}_j)\|_1^n$ , где  $\hat{x}_j = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ . Покажем, что эти три матрицы связаны соотношением

$$H_\alpha M = A_\alpha. \quad (1)$$

Пусть  $\sigma^{(j)}(t) = \sum_{k=0}^{n-1} \sigma_k(\hat{x}_j) t^k = \prod_{l \neq j} (1 + x_l t)$  и  $p(t) = \sum_{k=0}^{\infty} p_k t^k = \prod_{l=1}^n (1 - x_l t)^{-1}$ . Тогда

$$p(t) \sigma^{(j)}(-t) = (1 - x_j t)^{-1}.$$

Сравнивая коэффициенты при  $t^{\alpha_i}$  в обеих частях этого равенства, получаем

$$\sum_{l=1}^n p_{\alpha_i - n + l} (-1)^{n-l} \sigma_{n-l}(\hat{x}_j) = x_j^{\alpha_i}.$$

Это и есть требуемое соотношение (1).

Из (1), в частности, следует, что

$$\det H_\alpha \det M = \det A_\alpha. \quad (2)$$

Чтобы вычислить  $\det M$ , положим  $\alpha = \delta = (n-1, n-2, \dots, 1, 0)$ . В таком случае матрица  $H_\alpha$  имеет вид  $\|p_{j-i}\|_i^n$ . Эта матрица треугольная с элементами  $p_0 = 1$  на диагонали. Поэтому  $\det H_\delta = 1$ , а значит,  $\det M = \det A_\delta = a_\delta$ . А так как  $\det H_\alpha = s_{\alpha-\delta}$ , то при  $\alpha = \lambda + \delta$  равенство (2) принимает вид  $s_\lambda a_\delta = a_{\lambda+\delta}$ , т. е.  $s_\lambda = a_{\lambda+\delta}/a_\delta$ .  $\square$

**Теорема 1.4.6.2.** *Любой симметрический многочлен с вещественными коэффициентами однозначно представляется в виде  $\sum t_\lambda s_\lambda$ ,  $t_\lambda \in \mathbb{R}$ . При этом если коэффициенты многочлена целые, то  $t_\lambda \in \mathbb{Z}$ .*

**Доказательство.** Согласно теореме 1.4.2.2 отображение  $f \mapsto a_\delta f$  устанавливает взаимно однозначное соответствие между симметрическими и кососимметрическими многочленами. Поэтому согласно тождеству Якоби–Труди достаточно проверить, что любой кососимметрический многочлен однозначно представляется в виде  $\sum t_\lambda a_{\lambda+\delta}$ , но это очевидно.  $\square$

### 1.4.7. Формула Джамбелли

Функцию Шура  $s_\lambda$  можно выразить не только как минор матрицы  $\|p_{j-i}\|_0^\infty$ , построенный по разбиению  $\lambda$ , но и как минор матрицы  $\|\sigma_{j-i}\|_0^\infty$ , построенный по сопряжённому разбиению  $\lambda'$ .

**Теорема 1.4.7.1 (формула Джамбелли).** а) Пусть  $\lambda'$  — разбиение, сопряжённое с  $\lambda$ ;  $m$  — его длина. Тогда  $s_\lambda = |\sigma_{\lambda'_i + j - i}|_1^m$ .

б) Для косой функции Шура имеет место равенство  $s_{\lambda, \mu} = |\sigma_{\lambda'_i - \mu'_j + j - i}|_1^m$ , где  $m$  — наибольшая из длин разбиений  $\lambda'$  и  $\mu'$ .

**Доказательство.** Мы сразу будем рассматривать случай косых функций Шура. Пусть  $N + 1 = m + n$ ,  $P_N = \|p_{j-i}\|_0^N$  и  $\Sigma_N = \|(-1)^{j-i} \sigma_{j-i}\|_0^N$ . Матрицы  $P_N$  и  $\Sigma_N$  — верхние треугольные с единицами на диагонали. Легко проверить, что эти матрицы взаимно обратные. Действительно, из соотношения  $p(t)\sigma(-t) = 1$  следует, что  $\sum_r (-1)^r \sigma_r p_{n-r} = 0$  при  $n \geq 1$ , а это как раз и означает, что все элементы матрицы  $P_N \Sigma_N$ , лежащие вне диагонали, равны 0.

Рассмотрим минор  $|\sigma_{\lambda_j - \mu_i + i - j}|_1^n$  матрицы  $P_N$ , образованный строками  $\mu_i + n - i$  и столбцами  $\lambda_i + n - i$ ,  $i = 1, \dots, n$ . Прежде всего докажем, что алгебраическое дополнение к этому минору образовано строками  $n - 1 + j - \mu'_j$  и столбцами  $n - 1 + j - \lambda'_j$ ,  $j = 1, \dots, m$ , а затем вычислим соответствующий знак.

**Лемма 1.** Объединение наборов чисел  $\lambda_i + n - i$  и  $n - 1 + j - \lambda'_j$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ , составляет в точности набор  $\{0, 1, \dots, n + m - 1\}$ .

**Доказательство.** Рассмотрим квадрат со стороной  $n + m$  и поместим в его верхнем левом углу диаграмму  $\lambda$  (рис. 1.2). Ломаная, отделяющая  $\lambda$  от остальной части квадрата, состоит из  $m$  горизонтальных отрезков длины 1 и  $n$  вертикальных. Занумеруем эти отрезки числами  $0, \dots, n + m - 1$ , начиная с самого нижнего и самого левого отрезка. Вертикальные отрезки занумерованы числами  $\lambda_i + n - i$ ,  $i = 1, \dots, n$ . Действительно, вертикальному отрезку, пересекающему  $i$ -ю строку, предшествуют  $\lambda_i$  горизонтальных отрезков и  $n - i$  вертикальных. Аналогично горизонтальные отрезки занумерованы числами  $(n + m - 1) - (\lambda'_j + m - j) = n - 1 + j - \lambda'_j$ ,  $j = 1, \dots, m$ .  $\square$

**Лемма 2.** Знак перестановки  $(\lambda_i + n - i, n - 1 + j - \lambda'_j)$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ , равен  $(-1)^{|\lambda|}$ .

**Доказательство.** Снова обратимся к рис. 1.2. Возьмём произвольную клетку диаграммы Юнга, у которой две стороны принадлежат рассматриваемой ломаной. Уничтожение этой клетки диаграммы Юнга соответствует транспозиции, меняющей номера вертикального и горизонтального отрезка. После того как мы уничтожим все клетки, мы получим тождественную перестановку, поэтому знак исходной перестановки равен  $(-1)^{|\lambda|}$ .  $\square$



поэтому достаточно доказать, что

$$\left| \frac{1}{1 - x_i y_j} \right|_1^n = \sum_{\lambda} a_{\lambda+\delta}(x) a_{\lambda+\delta}(y).$$

Равенство

$$\frac{1}{1 - x_i y_j} = 1 + x_i y_j + x_i^2 y_j^2 + x_i^3 y_j^3 + \dots$$

показывает, что в выражении для определителя  $\left| \frac{1}{1 - x_i y_j} \right|_1^n$  коэффициент при  $y_1^{\lambda_1} y_2^{\lambda_2} \dots y_n^{\lambda_n}$  равен  $|x_i^{\lambda_j}|_1^n$ . Ясно также, что этот определитель симметричен по  $x$  и  $y$ , поэтому

$$\left| \frac{1}{1 - x_i y_j} \right|_1^n = |x_i^{\lambda_j}|_1^n |y_i^{\lambda_j}|_1^n = \sum_{\lambda} a_{\lambda}(x) a_{\lambda}(y).$$

Если  $\lambda$  нельзя представить в виде  $\bar{\lambda} + \delta$ , то  $a_{\lambda} = 0$ . Поэтому  $\sum_{\lambda} a_{\lambda}(x) a_{\lambda}(y) = \sum_{\lambda} a_{\lambda+\delta}(x) a_{\lambda+\delta}(y)$ .  $\square$

Каждому разбиению  $\lambda = (\lambda_1, \dots, \lambda_n)$  можно сопоставить симметрические многочлены  $p_{\lambda} = p_{\lambda_1} \dots p_{\lambda_n}$  и  $m_{\lambda} = m_{\lambda_1 \dots \lambda_n}$ . Если  $\lambda$  пробегает все разбиения длины  $n$  числа  $d = |\lambda|$ , то как многочлены  $p_{\lambda}$ , так и многочлены  $m_{\lambda}$ , образуют базис в пространстве  $V_d$  симметрических многочленов степени  $d$  от  $n$  переменных. Поэтому формула  $B(p_{\lambda}, m_{\mu}) = \delta_{\lambda, \mu} = \delta_{\lambda_1, \mu_1} \dots \delta_{\lambda_n, \mu_n}$  определяет билинейную форму  $B$  в пространстве  $V_d$ .

**Теорема 1.4.8.2.** *Функции Шура образуют ортонормированный базис относительно формы  $B$ , т.е.  $B(s_{\lambda}, s_{\mu}) = \delta_{\lambda, \mu}$ .*

**Доказательство.** Прежде всего заметим, что

$$\begin{aligned} \prod_{i,j=1}^n \frac{1}{1 - x_i y_j} &= \prod_{j=1}^n \prod_{i=1}^n (1 + x_i y_j + x_i^2 y_j^2 + \dots) = \\ &= \prod_{j=1}^n \sum_{k=0}^{\infty} p_k(x) y_j^k = \sum_{\lambda} p_{\lambda}(x) m_{\lambda}(y), \end{aligned}$$

где суммирование ведётся по всем разбиениям  $\lambda$ . Поэтому согласно теореме 1.4.8.1  $\sum_{\lambda} s_{\lambda}(x) s_{\lambda}(y) = \sum_{\lambda} p_{\lambda}(x) m_{\lambda}(y)$ ; такое же равенство верно и в том случае, когда  $\lambda$  пробегает все разбиения фиксированного числа  $d$ .

Разложим  $s_{\lambda}$  по базису  $\{p_{\mu}\}$  и по базису  $\{m_{\nu}\}$ :  $s_{\lambda} = \sum_{\mu} a_{\lambda\mu} p_{\mu}$  и  $s_{\lambda} = \sum_{\mu} b_{\nu\lambda} m_{\nu}$ , где  $a_{\lambda\mu}$  и  $b_{\nu\lambda}$  — некоторые числа.

$$B(s_{\lambda}, s_{\mu}) = \sum_{\rho, \tau} a_{\lambda\rho} b_{\tau\mu} \delta_{\rho\tau} = \sum_{\rho} a_{\lambda\rho} b_{\rho\mu}.$$

Тождество  $\sum_{\lambda} s_{\lambda}(x)s_{\lambda}(y) = \sum_{\lambda} p_{\lambda}(x)m_{\lambda}(y)$  позволяет получить соотношение между числами  $a_{\lambda\mu}$  и  $b_{\nu\lambda}$ . Действительно,

$$\sum_{\lambda} s_{\lambda}(x)s_{\lambda}(y) = \sum_{\lambda, \mu, \nu} a_{\lambda\mu} b_{\nu\lambda} p_{\mu}(x)m_{\nu}(y),$$

поэтому  $\sum_{\lambda} a_{\lambda\mu} b_{\nu\lambda} = \delta_{\mu\nu}$ . Таким образом, если  $B = \|b_{ij}\|$  и  $A = \|a_{ij}\|$ , то  $BA = I$ . Но тогда  $AB = I$ , а значит,  $B(s_{\lambda}, s_{\mu}) = \delta_{\lambda, \mu}$ , что и требовалось.  $\square$

**Следствие.** Форма  $B$  симметрическая положительно определённая, поэтому она задаёт скалярное произведение в пространстве  $V_d$ .

#### 1.4.9. Комбинаторное определение функций Шура

Функции Шура играют важную роль в комбинаторике. Это связано с тем, что функция Шура  $s_{\lambda}$  допускает следующее комбинаторное определение. Назовём *таблицей формы  $\lambda$*  расположение чисел  $1, 2, \dots, n$  в клетках диаграммы Юнга  $\lambda$  (возможно с повторениями), у которого в каждой строке числа идут (слева направо) в неубывающем порядке, а в каждом столбце (сверху вниз) — в строго возрастающем. Таблице  $T$  можно сопоставить моном  $x^{w(T)} = x_1^{T_1} \dots x_n^{T_n}$ , где  $T_k$  — количество чисел  $k$  в таблице  $T$ .

**Теорема 1.4.9.1.**  $s_{\lambda}(x_1, \dots, x_n) = \sum_T x^{w(T)}$ , где суммирование ведётся по всем таблицам  $T$  формы  $\lambda$ .

**Доказательство.** [Pr] Сначала докажем следующее рекуррентное выражение для функций Шура.

**Лемма.**  $s_{\lambda}(x_1, \dots, x_n) = \sum_{\mu} s_{\mu}(x_1, \dots, x_{n-1})x_n^{|\lambda|-|\mu|}$ , где суммирование ведётся по всем разбиениям  $\mu$  длины  $n-1$ , для которых  $\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \dots \geq \mu_{n-1} \geq \lambda_n$ .

**Доказательство.** Рассмотрим сначала один конкретный случай:  $n = 3$ ,  $\lambda_1 = 4$ ,  $\lambda_2 = 3$  и  $\lambda_3 = 1$ . В этом случае

$$s_{\lambda}(x, y, z) = \frac{\begin{vmatrix} x^{4+2} & y^{4+2} & z^{4+2} \\ x^{2+1} & y^{2+1} & z^{2+1} \\ x^{1+0} & y^{1+0} & z^{1+0} \end{vmatrix}}{\begin{vmatrix} x^2 & y^2 & z^2 \\ x^1 & y^1 & z^1 \\ x^0 & y^0 & z^0 \end{vmatrix}}.$$

Достаточно доказать, что  $s_{\lambda}(x_1, \dots, x_{n-1}, 1) = \sum_{\mu} s_{\mu}(x_1, \dots, x_{n-1})$ . Действительно, если мы возьмём многочлен  $s_{\lambda}(x_1, \dots, x_{n-1}, 1)$  и домножим в нём каждый моном на  $x_n^{|\lambda|-k}$ , где  $k$  — степень этого монома, то получим

$s_\lambda(x_1, \dots, x_n)$ . Поэтому положим  $z = 1$ . Затем преобразуем оба определителя следующим образом. Сначала вычтем последний столбец из всех остальных. Затем поделим первый столбец на  $x - 1$ , а второй — на  $y - 1$ . В результате получим

$$\left| \begin{array}{ccc|ccc} x^5 + x^4 + \dots + x + 1 & y^5 + y^4 + \dots + y + 1 & 1 & x + 1 & y + 1 & 1 \\ & x^2 + x + 1 & y^2 + y + 1 & 1 & 1 & 1 \\ & 1 & 1 & 0 & 0 & 1 \end{array} \right|.$$

Вычтем теперь в каждом определителе вторую строку из первой и третью строку из второй. В результате получим

$$\left| \begin{array}{cc|cc} x^5 + x^4 + x^3 & y^5 + y^4 + x^3 & x^1 & y^1 \\ x^2 + x^1 & y^2 + y^1 & x^0 & y^0 \end{array} \right|.$$

Определитель линейно зависит от каждой строки, поэтому мы приходим к требуемому выражению.

В общем случае в числителе стоит определитель, у которого элемент  $i$ -й строки является суммой мономов  $x_j^{\mu_i + (n-1) - i}$ , где  $\mu_i + (n-1) - i$  изменяется от  $\lambda_i + n - i - 1 = \lambda_i + (n-1) - i$  до  $\lambda_{i+1} + n - (i+1) = \lambda_{i+1} + (n-1) - i$ , т.е.  $\lambda_i \geq \mu_i \geq \lambda_{i+1}$ .  $\square$

Фиксируем диаграмму Юнга  $\lambda$  и будем строить таблицы формы  $\lambda$ , пользуясь леммой. Пусть  $\lambda^{(0)} = \lambda$  и  $\lambda^{(1)} = \mu$ , где  $\mu$  — одно из разбиений, по которым ведётся суммирование. Всем клеткам диаграммы  $\lambda$ , расположенным вне  $\mu$ , сопоставим число  $n$ . Затем повторим то же самое с диаграммой  $\lambda^{(1)}$ , сопоставляя соответствующим клеткам число  $n - 1$ , и т.д. Ясно, что числа в каждой строке и в каждом столбце идут в неубывающем порядке. Покажем, что в одном столбце не может быть двух клеток с числами  $n$ . Числа  $n$  в первой строке стоят в столбцах  $\mu_1 + 1, \dots, \lambda_1$ ; во второй — в столбцах  $\mu_2 + 1, \dots, \lambda_2$ ; ..., в  $n$ -й — в столбцах  $1, \dots, \lambda_n$ . Поэтому числа  $n$  стоят в столбцах  $1, \dots, \lambda_n; \mu_{n-1} + 1, \dots, \lambda_{n-1}; \dots; \mu_1 + 1, \dots, \lambda_1$ . Остаётся заметить, что  $\lambda_n < \mu_{n-1} + 1, \dots, \lambda_2 < \mu_1 + 1$ . Для чисел  $n - 1, \dots, 1$  доказательство аналогично. Ясно также, что эта конструкция даёт в точности все таблицы формы  $\lambda$ .  $\square$

## Решения

**1.1.** По условию  $A^T = -A$ . Поэтому в случае нечётного  $n$  получаем  $|A^T| = (-1)^n |A| = -|A|$ . С другой стороны, для любой квадратной матрицы  $|A^T| = |A|$ . Следовательно,  $|A| = -|A|$ , т.е.  $|A| = 0$ .



**1.2.** Пусть  $A$  — кососимметрическая матрица чётного порядка. Тогда

$$\begin{pmatrix} 0 & 1 & \dots & 1 \\ -1 & & & \\ \vdots & & A & \\ -1 & & & \end{pmatrix} \text{ — кососимметрическая матрица нечётного порядка. Согласно задаче 1.1 её определитель равен нулю. Следовательно,}$$

$$|A| = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -x & & & \\ \vdots & & A & \\ -x & & & \end{pmatrix} \begin{pmatrix} 0 & 1 & \dots & 1 \\ -x & & & \\ \vdots & & A & \\ -x & & & \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ -x & & & \\ \vdots & & A & \\ -x & & & \end{pmatrix}.$$

Вычитая первый столбец последней матрицы из всех остальных, получаем требуемое.

**1.3.** Если из первой строки матрицы  $A_n$  вычесть вторую строку, то получится строка  $(1, 1, 0, \dots, 0)$ . Поэтому если к строкам с номерами  $3, \dots, 2n$  прибавить первую строку и вычесть вторую, то в результате получим матрицу

$$\begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ -1 & 0 & 1 & \dots & 1 \\ 0 & 0 & & & \\ \vdots & \vdots & & A_{n-1} & \\ 0 & 0 & & & \end{pmatrix}.$$

Следовательно,  $|A_n| = |A_{n-1}|$ . Остаётся заметить, что  $|A_1| = 1$ .

**1.4.** Предположим, что все члены разложения определителя  $|a_{ij}|_n^n$ , где  $n \geq 3$ , положительны. Тогда  $\text{sign}(a_{11}a_{22}) = -\text{sign}(a_{12}a_{21})$ ,  $\text{sign}(a_{12}a_{23}) = -\text{sign}(a_{13}a_{22})$  и  $\text{sign}(a_{21}a_{13}) = -\text{sign}(a_{11}a_{23})$ . Перемножив эти три равенства, приходим к противоречию.

**1.5.** Для всех  $i \geq 2$  вычтем из  $i$ -й строки  $(i-1)$ -ю строку, умноженную на  $a$ . В результате получим верхнюю треугольную матрицу с диагональными элементами  $a_{11} = 1$  и  $a_{ii} = 1 - a^2$  при  $i > 1$ . Определитель этой матрицы равен  $(1 - a^2)^{n-1}$ .

**1.6.** Раскладывая определитель  $\Delta_{n+1}$  по последнему столбцу, получаем  $\Delta_{n+1} = x\Delta_n + h\Delta_n = (x+h)\Delta_n$ . Ясно также, что  $\Delta_1 = x+h$ .

**1.7.** Докажем индукцией по  $n$ , что искомым определитель равен

$$\prod_{i=1}^n (x_i - a_i b_i) \left( 1 + \sum_{i=1}^n \frac{a_i b_i}{x_i - a_i b_i} \right).$$

При  $n = 2$  это легко проверяется. Доказательство шага индукции проведём для матриц порядка 3 (в общем случае доказательство аналогично). Определитель

является линейной функцией первого столбца, поэтому

$$\begin{vmatrix} x_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & x_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & x_3 \end{vmatrix} = \begin{vmatrix} x_1 - a_1 b_1 & a_1 b_2 & a_1 b_3 \\ 0 & x_2 & a_2 b_3 \\ 0 & a_3 b_2 & x_3 \end{vmatrix} + \begin{vmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & x_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & x_3 \end{vmatrix}.$$

Первый определитель мы можем вычислить согласно предположению индукции, а для вычисления второго определителя нужно для всех  $i \geq 2$  вычесть из  $i$ -й строки первую строку, умноженную на  $a_i$ , и вынести из первой строки множитель  $a_1$ .

**1.8.** Определитель матрицы  $I - A$  представляет собой сумму двух слагаемых, а именно,  $\det(I - A) = 1 - c$ . Матрица  $A$  является матрицей преобразования  $Ae_i = c_{i-1}e_{i-1}$ , поэтому  $A^n = c_1 \cdots c_n I = cI$ . Следовательно,  $(I + A + A^2 + \cdots + A^{n-1})(I - A) = I - A^n = (1 - c)I$ , а значит,  $(1 - c) \det(I + A + A^2 + \cdots + A^{n-1}) = (1 - c)^n$ . При  $c \neq 1$  это равенство можно сократить на  $1 - c$  и получить требуемое. Ясно также, что определитель рассматриваемой матрицы непрерывно зависит от  $c_1, \dots, c_n$ , поэтому равенство остаётся справедливым и при  $c = 1$ .

**1.9.** Равенство  $\frac{1}{1 - x_i y_j} = \frac{1}{y_j} \frac{1}{\frac{1}{y_j} - x_i}$  показывает, что  $|a_{ij}|_1^n = \sigma |b_{ij}|_1^n$ , где  $\sigma = \frac{1}{y_1 \cdots y_n}$  и  $b_{ij} = \left(\frac{1}{y_j} - x_i\right)^{-1}$ , т.е.  $|b_{ij}|_1^n$  — определитель Коши. Поэтому

$$\begin{aligned} |b_{ij}|_1^n &= \prod_{i>j} \left(\frac{1}{y_i} - \frac{1}{y_j}\right) (x_j - x_i) \bigg/ \prod_{i,j} \left(\frac{1}{y_j} - x_i\right) = \\ &= \sigma^{n-1} \prod_{i>j} (y_j - y_i)(x_j - x_i) \bigg/ \sigma^n \prod_{i,j} (1 - x_i y_j). \end{aligned}$$

В итоге получаем, что

$$|a_{ij}|_1^n = \frac{\prod_{i>j} (y_j - y_i)(x_j - x_i)}{\prod_{i,j} (1 - x_i y_j)}.$$

**1.10.** Пусть  $A_n = |a_{ij}|_0^m$ . Фиксируем  $m$  и применим индукцию по  $n$ . При  $n = 0$  получаем матрицу  $A_0$ , у которой на диагонали стоят единицы, а выше диагонали — нули. Поэтому  $\det A_0 = 1$ . Рассмотрим вспомогательную матрицу  $B = \|b_{ij}\|_0^m$ , где  $b_{ii} = 1$  при  $i \leq m$ ,  $b_{i,i+1} = 1$  при  $i \leq m - 1$ , а все остальные элементы равны нулю. Тожество  $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$  показывает, что  $A_n B = A_{n+1}$ . Ясно также, что  $\det B = 1$ . Поэтому  $\det A_{n+1} = \det A_n$ .

**1.11.** Рассмотрим точки  $A, B, \dots, F$  с координатами  $(a^2, a), (b^2, b), \dots, (f^2, f)$ . Эти точки лежат на параболе, поэтому согласно теореме Паскаля<sup>3</sup> точки пересечения прямых  $AB$  и  $DE$ ,  $BC$  и  $EF$ ,  $CD$  и  $FA$  лежат на одной прямой. Нетрудно проверить, что точка пересечения прямых  $AB$  и  $DE$  имеет координаты

$$\left( \frac{(a+b)de - (d+e)ab}{d+e-a-b}, \frac{de-ab}{d+e-a-b} \right).$$

<sup>3</sup>Доказательство теоремы Паскаля можно найти, например, в [П1]

Остаётся заметить, что если точки  $(x_1, y_1)$ ,  $(x_2, y_2)$  и  $(x_3, y_3)$  лежат на одной пря-

мой, то 
$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

**1.12.** Пусть  $a = a_{11}a_{22}a_{33}$ ,  $b = a_{21}a_{32}a_{13}$ ,  $c = a_{12}a_{23}a_{31}$ ,  $a' = a_{31}a_{22}a_{13}$ ,  $b' = a_{21}a_{12}a_{33}$  и  $c' = a_{32}a_{23}a_{11}$ . По условию  $a + b + c = a' + b' + c'$ , а значит,

$$\bar{a} + \bar{b} + \bar{c} = \bar{a}' + \bar{b}' + \bar{c}'. \quad (1)$$

Кроме того,

$$abc = a'b'c', \quad (2)$$

поскольку оба эти произведения равны  $\prod a_{ij}$ .

Умножив (1) на (2), получим

$$bc + ac + ab = b'c' + a'c' + a'b'.$$

Таким образом  $\sigma_i(a, b, c) = \sigma_i(a', b', c')$  для  $i = 1, 2, 3$  ( $\sigma_i$  — элементарная симметрическая функция). Это означает, что числа  $a, b, c$  равны числам  $a', b', c'$  в каком-то порядке. Например, если  $a = a'$ ,  $b = b'$  и  $c = c'$ , то первая строка пропорциональна третьей. Остальные пять вариантов равенств дают пропорциональность других строк или столбцов.

**1.13.** Применим индукцию по  $n$ . При  $n = 1$  доказывать нечего, поэтому рассмотрим случай  $n = 2$ . Можно выбрать целые числа  $p$  и  $q$  так, что  $a_1p - a_2q = d$ . Тогда матрица  $\begin{pmatrix} a_1 & a_2 \\ q & p \end{pmatrix}$  искома. Предположим теперь, что утверждение уже доказано для любых  $n - 1$  чисел, где  $n \geq 3$ . Пусть  $D_{n-1}$  — целочисленная матрица с первой строкой  $(a_1, \dots, a_{n-1})$ , определитель которой равен наибольшему общему делителю  $d_{n-1}$  чисел  $a_1, \dots, a_{n-1}$ . Ясно, что  $d$  — наибольший общий делитель чисел  $d_{n-1}$  и  $a_n$ , поэтому можно выбрать целые числа  $p$  и  $q$  так, что  $d_{n-1}p - a_nq = d$ . Рассмотрим целочисленную матрицу

$$D = \begin{pmatrix} D_{n-1} & a_n \\ \dots & 0 \\ \dots & 0 \\ \frac{qa_1}{d_{n-1}} & \frac{qa_2}{d_{n-1}} & \dots & \frac{qa_{n-1}}{d_{n-1}} & p \end{pmatrix}.$$

Раскладывая определитель по последнему столбцу, получаем  $\det D = d_{n-1}p + (-1)^n a_n \det E_{n-1}$ , где матрица  $E_{n-1}$  получается из матрицы  $D$  вычёркиванием первой строки и последнего столбца. Если последнюю строку матрицы  $E_{n-1}$  умножить на  $\frac{d_{n-1}}{q}$  и переставить её на место первой строки (совершив при этом  $n - 2$  транспозиции), то в результате получим матрицу  $D_{n-1}$ . Таким образом,  $\det E_{n-1} = (-1)^{n-2} \frac{q}{d_{n-1}} \det D_{n-1} = (-1)^{n-2} q$ , поэтому  $\det D = d_{n-1}p - a_nq = d$ , т.е.  $D$  — требуемая матрица.

**1.14.** Пусть  $s = x_1 + \dots + x_n$ . Тогда  $k$ -й элемент последнего столбца имеет вид  $(s - x_k)^{n-1} = (-x_k)^{n-1} + \sum_{i=0}^{n-2} p_i x_k^i$ . Следовательно, прибавляя к последнему столбцу линейную комбинацию остальных столбцов с коэффициентами  $-p_0, \dots, -p_{n-2}$ , приходим к определителю

$$\begin{vmatrix} 1 & x_1 & \dots & x_1^{n-2} & (-x_1)^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-2} & (-x_n)^{n-1} \end{vmatrix}.$$

**1.15.** Пусть  $\Delta$  — искомый определитель. Умножая первую строку на  $x_1, \dots, n$ -ю строку на  $x_n$ , получаем

$$\sigma \Delta = \begin{vmatrix} x_1 & x_1^2 & \dots & x_1^{n-2} & \sigma \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^2 & \dots & x_n^{n-2} & \sigma \end{vmatrix},$$

где  $\sigma = x_1 \dots x_n$ . Переставляя последний столбец нового определителя на первое место, получаем  $\sigma \Delta = (-1)^{n-1} \sigma V(x_1, \dots, x_n)$ , т.е.  $\Delta = (-1)^{n-1} V(x_1, \dots, x_n)$  при  $\sigma \neq 0$ . Определитель  $\Delta$  непрерывно зависит от  $x_1, \dots, x_n$ , поэтому равенство остаётся справедливым и при  $\sigma = 0$ .

**1.16.** Так как  $\lambda_i^{n-k}(1 + \lambda_i^2)^k = \lambda_i^n ((1 + \lambda_i^2)/\lambda_i)^k$ , то  $|a_{ik}|_0^n = \sigma^n V(\mu_0, \dots, \mu_n)$ , где  $\mu_i = \lambda_i + \frac{1}{\lambda_i}$  и  $\sigma = \lambda_0 \dots \lambda_n$ . Учитывая, что  $\mu_i - \mu_j = (\lambda_j - \lambda_i)(1 - \lambda_i \lambda_j)/\lambda_i \lambda_j$ , получаем  $|a_{ik}|_0^n = \prod_{i>j} (\lambda_j - \lambda_i)(1 - \lambda_i \lambda_j)$ .

**1.17.** Добавим к матрице  $V$  сначала  $(n+1)$ -й столбец, состоящий из  $n$ -х степеней, а потом добавим первую строку  $(1, -x, x^2, \dots, (-x)^n)$ . Полученная матрица  $W$  снова является матрицей Вандермонда, поэтому

$$\begin{aligned} \det W &= (x + x_1) \dots (x + x_n) \det V = \\ &= (\sigma_n + \sigma_{n-1}x + \dots + \sigma_1 x^{n-1} + x^n) \det V. \end{aligned}$$

С другой стороны, разложив определитель матрицы  $W$  по первой строке, получим

$$\det W = \det V_0 + x \det V_1 + \dots + x^{n-1} \det V_{n-1} + x^n \det V.$$

При фиксированных  $x_1, \dots, x_n$  получаем равенство двух многочленов от  $x$ . Их коэффициенты должны быть равны, поэтому  $\det V_k = \sigma_{n-k}(x_1, \dots, x_n) \det V$ .

**1.18.** Точно такие же рассуждения, как и при втором доказательстве теоремы 1.1.5.1, показывают, что искомый определитель равен  $a_1 \dots a_n \prod_{i>j} (x_i - x_j)$ .

**1.19.** а) Ясно, что

$$\begin{aligned} f'(x_i) &= (x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n) = \\ &= (-1)^{n-i} (x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n) = \\ &= (-1)^{n-i} \frac{\prod_{j>k} (x_j - x_k)}{\prod_{j>k, j,k \neq i} (x_j - x_k)} = \\ &= (-1)^{n-i} \frac{V(x_1, \dots, x_n)}{V(x_1, \dots, \hat{x}_i, \dots, x_n)}. \end{aligned}$$

б) Согласно а) получаем

$$\begin{aligned} \sum_{i=1}^n \frac{x_i^k}{f'(x_i)} &= \sum_{i=1}^n (-1)^{n-i} x_i^k \frac{V(x_1, \dots, x_n)}{V(x_1, \dots, \hat{x}_i, \dots, x_n)} = \\ &= \frac{(-1)^{n-1}}{V(x_1, \dots, x_n)} \sum_{i=1}^n (-1)^{i-1} x_i^k V(x_1, \dots, \hat{x}_i, \dots, x_n) = \\ &= \frac{(-1)^{n-1}}{V(x_1, \dots, x_n)} (-1)^{n-1} \begin{vmatrix} 1 & x_1 & \dots & x_1^{n-2} & x_1^k \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-2} & x_n^k \end{vmatrix}. \end{aligned}$$

При записи последнего равенства мы воспользовались разложением определителя по последнему столбцу.

в) Воспользуемся формулой б). Если  $0 \leq k \leq n-2$ , то мы получаем определитель с двумя одинаковыми строками. При  $k = n-1$  получаем определитель Вандермонда.

**1.20.** Ясно, что  $\binom{x}{k} = \frac{x^k}{k!} + \sum_{i=0}^{k-1} c_i x^i$ . Поэтому исходный определитель равен

$$\begin{vmatrix} 1 & x_1 & x_1^2/2! & \dots & x_1^{n-1}/(n-1)! \\ 1 & x_2 & x_2^2/2! & \dots & x_2^{n-1}/(n-1)! \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2/2! & \dots & x_n^{n-1}/(n-1)! \end{vmatrix} = \frac{1}{2!3! \dots (n-1)!} \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Кроме того,  $2!3! \dots (n-1)! = \prod_{1 \leq i < j \leq n} (j-i)$ .

**1.21.** Пусть  $x_i = in$ . Тогда  $a_{i1} = x_i$ ,  $a_{i2} = \frac{x_i(x_i-1)}{2}$ , ...,  $a_{ir} = \frac{x_i(x_i-1) \dots (x_i-r+1)}{r!}$ , т.е. в  $k$ -м столбце стоят одинаковые многочлены  $k$ -й степени от  $x$ . Так как определитель не изменяется при прибавлении к его столбцам линейных комбинаций других столбцов, то его можно привести к виду  $|b_{ik}|_1^r$ , где  $b_{ik} = x_i^k/k! = (n^k/k!)i^k$ . Поделив  $k$ -ю строку на  $k \frac{n^k}{k!}$  для всех  $k = 1, \dots, r$ , получим определитель Вандермонда. Следовательно,  $|a_{ik}|_1^r = |b_{ik}|_1^r = n^{\frac{n^2}{2!}} \dots \frac{n^r}{r!} r! V(1, 2, \dots, r) = n^{r(r+1)/2}$ , так как  $\prod_{1 \leq j < i \leq r} (i-j) = 2!3! \dots (r-1)!$ .

**Замечание.** По поводу другого решения этой задачи см. следствие 1 теоремы 1.1.14.1.

**1.22.** Как видно из решения задачи 1.20, рассматриваемое произведение рав-

но  $\begin{vmatrix} 1 & \binom{x_1}{1} & \dots & \binom{x_1}{n-1} \\ 1 & \binom{x_2}{1} & \dots & \binom{x_2}{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & \binom{x_n}{1} & \dots & \binom{x_n}{n-1} \end{vmatrix}$ , где  $\binom{x}{k} = \frac{x(x-1) \dots (x-k+1)}{k!}$ . Индукцией по  $k$  легко дока-

зать, что число  $\binom{x}{k}$  целое для всех целых  $x$ . Действительно,  $\binom{x}{1} = x$ . Кроме того,  $\binom{x+1}{k+1} - \binom{x}{k+1} = \binom{x}{k}$  и  $\binom{0}{k+1} = 0$ . Ясно также, что определитель матрицы с целочисленными элементами является целым числом.

**1.23.** Для  $i = 1, \dots, n$  умножим  $i$ -ю строку определителя  $|a_{ij}|_1^n$  на  $m_i!$ , где  $m_i = k_i + m - i$ . Получим определитель  $|b_{ij}|_1^n$ , где

$$b_{ij} = \frac{(k_i + n - i)!}{(k_i + j - i)!} = m_i(m_i - 1) \dots (m_i + j + 1 - n).$$

Элементы  $j$ -го столбца определителя  $|b_{ij}|_1^n$  являются одинаковыми многочленами степени  $n - j$  от  $m_i$ , причём коэффициенты при старших членах этих многочленов равны 1. Поэтому, вычитая из каждого столбца линейные комбинации предыдущих столбцов, определитель  $|b_{ij}|_1^n$  можно свести к определителю со строками  $(m_i^{n-1}, m_i^{n-2}, \dots, 1)$ . Этот определитель равен  $\prod_{i < j} (m_i - m_j)$ . Ясно также, что  $|a_{ij}|_1^n = |b_{ij}|_1^n / m_1! m_2! \dots m_n!$ .

**1.24.** При  $n = 3$  легко проверить, что

$$\|a_{ij}\|_0^2 = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{pmatrix} \begin{pmatrix} p_1 & p_1 x_1 & p_1 x_1^2 \\ p_2 & p_2 x_2 & p_2 x_2^2 \\ p_3 & p_3 x_3 & p_3 x_3^2 \end{pmatrix}.$$

В общем случае справедливо аналогичное равенство.

**1.25.** Искомый определитель можно представить в виде произведения двух определителей

$$\begin{vmatrix} 1 & \dots & 1 & 1 \\ x_1 & \dots & x_n & y \\ x_1^2 & \dots & x_n^2 & y^2 \\ \dots & \dots & \dots & \dots \\ x_1^n & \dots & x_n^n & y^n \end{vmatrix} \cdot \begin{vmatrix} 1 & x_1 & \dots & x_1^{n-1} & 0 \\ 1 & x_2 & \dots & x_2^{n-1} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_1 & \dots & x_1^{n-1} & 0 \\ 0 & 0 & \dots & 0 & 1 \end{vmatrix},$$

поэтому он равен  $\prod_i (y - x_i) \prod_{i > j} (x_i - x_j)^2$ .

**1.26.** При  $n = 2$  легко проверить, что

$$\|a_{ij}\|_0^2 = \begin{pmatrix} 1 & 2x_0 & x_0^2 \\ 1 & 2x_1 & x_1^2 \\ 1 & 2x_2 & x_2^2 \end{pmatrix} \begin{pmatrix} y_0^2 & y_1^2 & y_2^2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{pmatrix};$$

в общем случае элементами первой матрицы являются числа  $\binom{n}{k} x_i^k$ .

**1.27.** Предположим, что есть ненулевое решение, причём количество попарно различных чисел  $\lambda_i$  равно  $r$ . Объединив равные числа  $\lambda_i$  в  $r$  групп, получим  $m_1 \lambda_1^k + \dots + m_r \lambda_r^k = 0$  при  $k = 1, \dots, n$ . Пусть  $x_1 = m_1 \lambda_1, \dots, x_r = m_r \lambda_r$ . Тогда  $\lambda_1^{k-1} x_1 + \dots + \lambda_r^{k-1} x_r = 0$  при  $k = 1, \dots, n$ . Взяв первые  $r$  из этих уравнений, получим систему линейных уравнений от  $x_1, \dots, x_r$  причём определителем этой системы является определитель Вандермонда  $V(\lambda_1, \dots, \lambda_r) \neq 0$ . Следовательно,  $x_1 = \dots = x_r = 0$ , а значит,  $\lambda_1 = \dots = \lambda_r = 0$ . Получено противоречие, поэтому решение только нулевое.

**1.28.** Доказательство проведём индукцией по  $n$ . При  $n = 1$  утверждение очевидно. Вычитая первый столбец матрицы  $\|a_{ij}\|_0^n$  из всех остальных, получим матрицу  $\|b_{ij}\|_0^n$  где  $b_{ij} = \sigma_i(\hat{x}_j) - \sigma_i(\hat{x}_0)$  при  $j \geq 1$ . Докажем теперь, что

$$\sigma_k(\hat{x}_i) - \sigma_k(\hat{x}_j) = (x_j - x_i)\sigma_{k-1}(\hat{x}_i, \hat{x}_j).$$

В самом деле,

$$\begin{aligned}\sigma_k(x_1, \dots, x_n) &= \sigma_k(\hat{x}_i) + x_i\sigma_{k-1}(\hat{x}_i) = \\ &= \sigma_k(\hat{x}_i) + x_i\sigma_{k-1}(\hat{x}_i, \hat{x}_j) + x_ix_j\sigma_{k-2}(\hat{x}_i, \hat{x}_j),\end{aligned}$$

поэтому

$$\begin{aligned}\sigma_k(\hat{x}_i) + x_i\sigma_{k-1}(\hat{x}_i, \hat{x}_j) &= \sigma_k(x_1, \dots, x_n) - x_ix_j\sigma_{k-2}(\hat{x}_i, \hat{x}_j) = \\ &= \sigma_k(\hat{x}_j) + x_j\sigma_{k-1}(\hat{x}_i, \hat{x}_j).\end{aligned}$$

Следовательно,

$$|b_{ij}|_0^n = (x_0 - x_1) \cdots (x_0 - x_n)|c_{ij}|_0^n,$$

где  $c_{ij} = \sigma_{k-1}(\hat{x}_i, \hat{x}_j)$ .

**1.29.** а) Пусть  $k = [n/2]$ . Умножим на  $-1$  строки матрицы  $\|b_{ij}\|_1^n$  с номерами  $2, 4, \dots, 2k$ , а затем умножим на  $-1$  столбцы с номерами  $2, 4, \dots, 2k$ . В итоге получим матрицу  $\|a_{ij}\|_1^n$ .

б) Сначала умножим строки матрицы  $\|b_{ij}\|_1^n$  на  $(-1)^{m_i}$ , где  $i$  — номер строки, а затем умножим столбцы на  $(-1)^{n_j}$ , где  $j$  — номер столбца. В результате получим матрицу  $\|a_{ij}\|_1^n$ . При указанных операциях определитель умножается на  $(-1)^{M+N}$ .

**1.30.** Первое решение. Матрица  $\|b_{ij}\|_1^n$  получается из матрицы  $\|a_{ij}\|_1^n$  следующими операциями. Сначала переставим строку с номером  $n$  на первое место, затем строку с номером  $n-1$  переставим на второе место и т.д. Затем аналогичные операции нужно проделать со столбцами. При указанной перестановке строк определитель умножается на  $(-1)^{n(n-1)/2}$ , поэтому при перестановке строк и столбцов определитель умножается на  $((-1)^{n(n-1)/2})^2 = 1$ .

Второе решение. Рассмотрим матрицу  $J = \|\delta_{i,n+1-j}\|_1^n$ . Элемент  $c_{ij}$  матрицы  $C = JAJ$  равен  $\sum_{k,l} \delta_{i,n+1-k} a_{kl} \delta_{l,n+1-j} = \sum_l a_{n+1-i,l} \delta_{l,n+1-j} = a_{n+1-i,n+1-j}$ . Поэтому  $\det B = (\det J)^2 \det A = \det A$ , поскольку  $\det J = \pm 1$ .

**1.31.** Легко проверить, что оба выражения равны произведению определителей

$$\begin{vmatrix} a_1 & a_2 & 0 & 0 \\ a_3 & a_4 & 0 & 0 \\ 0 & 0 & b_1 & b_2 \\ 0 & 0 & b_3 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & 0 & c_2 & 0 \\ 0 & d_1 & 0 & d_2 \\ c_3 & 0 & c_4 & 0 \\ 0 & d_3 & 0 & d_4 \end{vmatrix}.$$

**1.32.** Оба определителя равны

$$\begin{aligned}
& a_1 a_2 a_3 \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{23} \end{vmatrix} + a_1 b_2 b_3 \begin{vmatrix} a_{11} & b_{12} & b_{13} \\ a_{21} & b_{22} & b_{23} \\ a_{31} & b_{32} & b_{23} \end{vmatrix} + b_1 a_2 b_3 \begin{vmatrix} b_{11} & a_{12} & b_{13} \\ b_{21} & a_{22} & b_{23} \\ b_{31} & a_{32} & b_{23} \end{vmatrix} - \\
& - a_1 a_2 b_3 \begin{vmatrix} a_{11} & a_{12} & b_{13} \\ a_{21} & a_{22} & b_{23} \\ a_{31} & a_{32} & b_{23} \end{vmatrix} - b_1 a_2 a_3 \begin{vmatrix} b_{11} & a_{12} & a_{13} \\ b_{21} & a_{22} & a_{23} \\ b_{31} & a_{32} & a_{23} \end{vmatrix} - b_1 b_2 b_3 \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{23} \end{vmatrix}.
\end{aligned}$$

**1.33.** Для определителей матриц порядка  $n + 1$  легко проверить следующие равенства:

$$\begin{aligned}
& \begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} & 0 \\ \dots & \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} & 0 \\ -1 & \dots & -1 & 1 \end{vmatrix} = \begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} & (n-1)s_1 \\ \dots & \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} & (n-1)s_n \\ -1 & \dots & -1 & 1-n \end{vmatrix} = \\
& = (n-1) \begin{vmatrix} s_1 - a_{11} & \dots & s_1 - a_{1n} & s_1 \\ \dots & \dots & \dots & \dots \\ s_n - a_{n1} & \dots & s_n - a_{nn} & s_n \\ -1 & \dots & -1 & -1 \end{vmatrix} = \\
& = (n-1) \begin{vmatrix} -a_{11} & \dots & -a_{1n} & s_1 \\ \dots & \dots & \dots & \dots \\ -a_{n1} & \dots & -a_{nn} & s_n \\ 0 & \dots & 0 & -1 \end{vmatrix}.
\end{aligned}$$

Чтобы получить первое равенство, нужно к последнему столбцу прибавить сумму всех остальных, а чтобы получить последнее равенство, нужно последний столбец вычесть из всех остальных.

**1.34.** Воспользуемся тождеством  $\binom{p}{q} + \binom{p}{q-1} = \binom{p+1}{q}$ . Складывая соответствующим образом столбцы, от матрицы со строками вида  $\left(\binom{n}{m}, \binom{n}{m-1}, \dots, \binom{n}{m-k}\right)$  можно перейти к матрице со строками  $\left(\binom{n}{m}, \binom{n+1}{m}, \binom{n+1}{m-1}, \dots, \binom{n+1}{m-k+1}\right)$ . Проледывая аналогичную операцию для столбцов с номерами  $2, 3, \dots, n$ , получим матрицу со строками  $\left(\binom{n}{m}, \binom{n+1}{m}, \binom{n+2}{m}, \binom{n+2}{m-1}, \dots, \binom{n+2}{m-k+2}\right)$ . После нескольких таких операций в итоге получим матрицу со строками  $\left(\binom{n}{m}, \binom{n+1}{m}, \dots, \binom{n+k}{m}\right)$ .

**1.35.** Вычтем в определителе  $\Delta_n(k)$  из строки с номером  $i + 1$  строку с номером  $i$  для  $i = n, n-1, \dots, 1$ . В результате получим  $\Delta_n(k) = \Delta'_{n-1}(k)$ , где  $\Delta'_m(k) = |a'_{ij}|_0^m$ ,  $a'_{ij} = \binom{k+i}{2j+1}$ . А так как  $\binom{k+i}{2j+1} = \frac{k+i}{2j+1} \binom{k-1+i}{2j}$ , то

$$\Delta'_{n-1}(k) = \frac{k(k+1) \dots (k+n-1)}{1 \cdot 3 \dots (2n-1)} \Delta_{n-1}(k-1).$$



**1.36.** [Ca2] Согласно задаче 1.34  $D_n = D'_n = |a'_{ij}|_0^n$ , где  $a'_{ij} = \binom{n+1+i}{2j}$ , т.е. в обозначениях задачи 1.35 получаем

$$D_n = \Delta_n(n+1) = \frac{(n+1)(n+2)\dots 2n}{1\cdot 3\dots (2n-1)} \Delta_{n-1}(n) = 2^n D_{n-1},$$

так как  $(n+1)(n+2)\dots 2n = (2n)!/n!$ ,  $1\cdot 3\dots (2n-1) = (2n)!/2\cdot 4\dots 2n$  и  $\Delta_{n-1}(n) = D_{n-1}$ . А так как  $D_0 = 1$ , то  $D_n = 2^k$ , где  $k = n + (n-1) + \dots + 1 = n(n+1)/2$ .

**1.37.** Доказательство проведём при  $n = 2$ . Согласно задаче 1.29  $|a_{ij}|_0^2 = |a'_{ij}|_0^2$ , где  $a'_{ij} = (-1)^{i+j} a_{ij}$ . Прибавим к последнему столбцу матрицы  $|a'_{ij}|_0^2$  её предпоследний столбец, а затем к последней строке полученной матрицы прибавим её предпоследнюю строку. В результате получим матрицу

$$\begin{pmatrix} a_0 & -a_1 & -\Delta^1 a_1 \\ -a_1 & a_2 & \Delta^1 a_2 \\ -\Delta^1 a_1 & \Delta^1 a_2 & \Delta^2 a_2 \end{pmatrix}, \text{ где } \Delta^1 a_k = a_k - a_{k+1}, \quad \Delta^{n+1} a_k = \Delta^1(\Delta^n a_k). \text{ Затем}$$

прибавим ко 2-й строке 1-ю, а к 3-й прибавим 2-ю строку полученной матрицы; такую же операцию проделаем для столбцов. В результате получим матрицу

$$\begin{pmatrix} a_0 & \Delta^1 a_0 & \Delta^2 a_0 \\ \Delta^1 a_0 & \Delta^2 a_0 & \Delta^3 a_0 \\ \Delta^2 a_0 & \Delta^3 a_0 & \Delta^4 a_0 \end{pmatrix}, \text{ Индукцией по } k \text{ легко проверить, что } b_k = \Delta^k a_0. \text{ В об-}$$

щем случае доказательство аналогично.

**1.38.** Матрицы  $A$  и  $B$  можно представить в виде

$$A = \begin{pmatrix} P & PX \\ YP & YPX \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} WQV & WQ \\ QV & Q \end{pmatrix},$$

где  $P = A_{11}$  и  $Q = B_{22}$  (см. § 2.2.3). Используя теорему 1.1.13.1, получим

$$\begin{aligned} \begin{pmatrix} P + WQV & PX + WQ \\ YP + QV & YPX + Q \end{pmatrix} &= \begin{pmatrix} P & WQ \\ YP & Q \end{pmatrix} \cdot \begin{pmatrix} I & \\ & V \end{pmatrix} = \\ &= \frac{1}{|P| \cdot |Q|} \begin{pmatrix} P & WQ \\ YP & Q \end{pmatrix} \cdot \begin{pmatrix} P & PX \\ QV & Q \end{pmatrix}. \end{aligned}$$

**1.39.** Раскладывая определитель матрицы

$$C = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} & b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & a_{nn} & b_{n1} & \dots & b_{nn} \\ a_{11} & 0 & \dots & 0 & b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & 0 & \dots & 0 & b_{n1} & \dots & b_{nn} \end{pmatrix}$$

по первым  $n$  строкам (теорема 1.2.4.1), получаем

$$\begin{aligned} C &= \sum_{k=1}^n (-1)^{\varepsilon_k} \begin{pmatrix} a_{12} & \dots & a_{1n} & a_{1k} \\ \dots & \dots & \dots & \dots \\ a_{n2} & \dots & a_{nn} & a_{nk} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & b_{11} & \dots & \hat{b}_{1k} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & b_{n1} & \dots & \hat{b}_{nk} & \dots & b_{nn} \end{pmatrix} = \\ &= \sum_{k=1}^n (-1)^{\varepsilon_k + \alpha_k + \beta_k} \begin{pmatrix} b_{1k} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_{nk} & a_{n2} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \dots & a_{11} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ b_{n1} & \dots & a_{n1} & \dots & b_{nn} \end{pmatrix}, \end{aligned}$$

где  $\varepsilon_k = (1 + 2 + \dots + n) + (2 + \dots + n + (k + n)) \equiv k + n + 1 \pmod{2}$ ,  $\alpha_k = n - 1$  и  $\beta_k = k - 1$ , т.е.  $\varepsilon_k + \alpha_k + \beta_k \equiv 1 \pmod{2}$ . С другой стороны, вычитая из  $i$ -й строки матрицы  $C$   $(i + n)$ -ю строку для  $i = 1, \dots, n$ , получаем  $|C| = -|A| \cdot |B|$ .

**1.40.** Доказательство проведём индукцией по  $p$ . При  $p = 1$  равенство очевидно. Тожество из задачи ?? можно в нашем случае записать в виде

$$|AB| \cdot |BC| = \sum_{k=1}^p (-1)^{p+k+1} |\hat{A}_1 BC_k| \cdot |A_1 B \hat{C}_k|. \quad (1)$$

Разложим определитель  $\Delta = |\hat{A}_i BC_j|$  по первой строке:

$$\Delta = \sum_{k=1}^p (-1)^{k+1} |\hat{A}_1 BC_k| \cdot \Delta_k. \quad (2)$$

Все элементы определителя  $\Delta_k$  имеют вид  $|\hat{A}_i BC_k|$ , где  $i > 1$ ; присоединим в них столбец  $A_1$  к матрице  $B$  (для этого нужно совершить  $p - 2$  транспозиции). К полученному определителю можно применить предположение индукции; в результате получим

$$\Delta_k = (-1)^p |AB|^{p-2} |A_1 B \hat{C}_k|$$

(в определителе  $|AB|$  столбец  $A_1$  снова переставлен на своё прежнее место). Подставляя это выражение для  $\Delta_k$  в (2) и учитывая (1), получаем требуемое.

**1.41.** Рекуррентное соотношение для чисел Фибоначчи позволяет определить их для всех целых  $n$ ; в частности,  $F_0 = 0$ . Поэтому при  $n = 1$  требуемые равенства выполняются. Кроме того,

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} F_{n+1} + F_n & F_{n+1} \\ F_n + F_{n-1} & F_n \end{pmatrix} = \begin{pmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{pmatrix}, \\ \begin{pmatrix} F_{n+1} & -F_n \\ -F_n & F_{n-1} \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -F_{n+1} - F_n & F_{n+1} \\ F_n + F_{n-1} & -F_n \end{pmatrix} = \begin{pmatrix} -F_{n+2} & F_{n+1} \\ F_{n+1} & -F_n \end{pmatrix}.$$

**1.42.** Обозначим рассматриваемый определитель  $\Delta_n$ . Для определителей такого вида выполняется соотношение  $\Delta_n = x\Delta_{n-1} + a^2\Delta_{n-2}$  (см. §1.1.11). Пусть  $P_n(x)$  — многочлен, указанный в качестве предполагаемого ответа. Легко видеть, что  $\Delta_1 = x = P_1(x)$  и  $\Delta_2 = x^2 + a^2 = P_2(x)$ . Поэтому достаточно доказать, что  $P_n = xP_{n-1} + a^2P_{n-2}$ . Но это равенство эквивалентно известному тождеству для биномиальных коэффициентов:  $\binom{n-k}{k} = \binom{n-k-1}{k} + \binom{n-k-1}{k-1}$ .

**1.43.** Рассмотрим определитель из задачи 1.42 для  $x = 1$  и  $a = 1$ . Рекуррентное соотношение для этого определителя показывает, что  $F_{n+1} = \Delta_n$ . Определитель  $\Delta_n$  в задаче 1.42 представлен в виде многочлена, указанного явным образом. При  $x = 1$  и  $a = 1$  этот многочлен обращается в  $1 + \binom{n-1}{1} + \binom{n-2}{2} + \binom{n-3}{3} + \dots$

**2.1.** Коэффициент при  $\lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_m}$  в определителе матрицы  $A + \text{diag}(\lambda_1, \dots, \lambda_n)$  равен минору, полученному из матрицы  $A$  вычёркиванием строк и столбцов с номерами  $i_1, \dots, i_m$ .

**2.2.** С точностью до знака множитель при  $a_{ij}$  в определителе  $|a_{ij}|_1^n$  равен множителю при  $x_i y_j$  в рассматриваемом определителе. Для сравнения знаков этих множителей нужно сравнить знаки перестановок  $\begin{pmatrix} i & n+1 & \dots \\ j & n+1 & \dots \end{pmatrix}$  и  $\begin{pmatrix} i & n+1 & \dots \\ n+1 & j & \dots \end{pmatrix}$ . Эти перестановки отличаются на транспозицию  $(j, n+1)$ , поэтому множители имеют разные знаки.

**2.3.** Согласно задаче 2.2 определитель, стоящий в левой части, равен

$$-\sum_{i,j=1}^p A_i(x) A_j(x) = -\sum_{i,j=1}^p \sum_{s,t=1}^n A_{ij} a_{is} x_s a_{jt} x_t,$$

где  $A_{ij}$  — алгебраическое дополнение элемента  $a_{ij}$  в матрице  $\|a_{ij}\|_1^p$ . В правой части стоит величина  $-\sum_{s,t=1}^n x_s x_t a_{st} \Delta$ , поэтому достаточно проверить, что  $\sum_{i,j=1}^p A_{ij} a_{is} a_{jt} = a_{st} \Delta$ . Так как  $\sum_{i=1}^p A_{ij} a_{is} = \delta_{js} \Delta$ , то

$$\sum_{j=1}^p a_{jt} \sum_{i=1}^p A_{ij} a_{is} = \sum_{j=1}^p a_{jt} \delta_{js} \Delta = a_{st} \Delta.$$

**2.4.** Пусть  $B = A^T A$ . Тогда

$$B \begin{pmatrix} i_1 & \dots & i_k \\ i_1 & \dots & i_k \end{pmatrix} = \begin{vmatrix} b_{i_1 i_1} & \dots & b_{i_1 i_k} \\ \dots & \dots & \dots \\ b_{i_k i_1} & \dots & b_{i_k i_k} \end{vmatrix} = \det \left[ \begin{pmatrix} a_{1 i_1} & \dots & a_{n i_1} \\ \dots & \dots & \dots \\ a_{1 i_k} & \dots & a_{n i_k} \end{pmatrix} \cdot \begin{pmatrix} a_{1 i_1} & \dots & a_{1 i_k} \\ \dots & \dots & \dots \\ a_{n i_1} & \dots & a_{n i_k} \end{pmatrix} \right].$$

Поэтому, воспользовавшись формулой Бине—Коши, получаем

$$B \begin{pmatrix} i_1 & \dots & i_k \\ i_1 & \dots & i_k \end{pmatrix} = \sum_{1 \leq j_1 < \dots < j_k \leq n} A^2 \begin{pmatrix} i_1 & \dots & i_k \\ j_1 & \dots & j_k \end{pmatrix}.$$

**2.5.** Коэффициент при  $u_1$  в сумме определителей, стоящих в левой части, равен  $a_{11} A_{11} + \dots + a_{n1} A_{n1} = |A|$ . Для  $u_2, \dots, u_n$  доказательство аналогично.

**2.6.** Пусть точка  $A$  имеет координаты  $(a_1, a_2)$ ; для остальных точек обозначения аналогичны. Тогда, например,

$$S_{ABC} = \frac{1}{2} \begin{vmatrix} 1 & 1 & 1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}.$$

Рассмотрим определитель матрицы

$$\frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & -1 & 0 & 0 \\ a_1 & b_1 & c_1 & -d_1 & 0 & 0 \\ a_2 & b_2 & c_2 & -d_2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ a_1 & b_1 & c_1 & 0 & e_1 & f_1 \\ a_2 & b_2 & c_2 & 0 & e_2 & f_2 \end{pmatrix}.$$

С одной стороны, если вычесть строку с номером  $i$  из строки с номером  $i + 3$  для  $i = 1, 2, 3$ , то получим, что этот определитель равен  $S_{DEF}S_{ABC}$ . С другой стороны, если записать разложение Лапласа для определителя этой матрицы, фиксируя первые три строки, то получим, что этот определитель равен  $S_{AEF}S_{DBC} + S_{BEF}S_{DCA} + S_{CEF}S_{DAB}$ .

2.7. Ответ:  $\begin{pmatrix} I & -A & AB \\ 0 & I & -B \\ 0 & 0 & I \end{pmatrix}.$

2.8. Если  $i < j$ , то, вычёркивая  $i$ -ю строку и  $j$ -й столбец верхней треугольной матрицы, получим верхнюю треугольную матрицу с нулями на диагонали на местах от  $i$  до  $j - 1$ .

2.9. Легко проверить, что

$$(I_n - BA)^{-1} = I_n + B(I_m - AB)^{-1}A.$$

Действительно,

$$\begin{aligned} (I_n + B(I_m - AB)^{-1}A)(I_n - BA) - I_n &= \\ &= B(I_m - AB)^{-1}A - BA - B(I_m - AB)^{-1}ABA = \\ &= B((I_m - AB)^{-1} - I_m - (I_m - AB)^{-1}AB)A = \\ &= B((I_m - AB)^{-1}(I_m - AB) - I)A = 0. \end{aligned}$$

2.10. Рассмотрим единичную матрицу порядка  $n - 1$ . Вставим между  $(i - 1)$ -м и  $i$ -м столбцом этой матрицы столбец из нулей, а затем между  $(j - 1)$ -й и  $j$ -й строкой полученной матрицы вставим строку из нулей. Минор  $M_{ji}$  полученной матрицы равен 1, а все остальные миноры равны 0, так как, вычёркивая любую другую строку или другой столбец, мы вычёркиваем хотя бы одну единицу, и получается вырожденная матрица.

2.11. По определению  $A_{ij} = (-1)^{i+j} \det B$ , где  $B$  — матрица порядка  $n - 1$ . Так как  $A^T = -A$ , то  $A_{ji} = (-1)^{i+j} \det(-B) = (-1)^{n-1} A_{ij}$ .

2.12. Ответ зависит от чётности  $n$ . Согласно задаче 1.3  $|A_{2k}| = 1$ , поэтому  $\text{adj } A_{2k} = A_{2k}^{-1}$ . При  $n = 4$  легко проверить, что

$$\begin{pmatrix} 0 & -1 & -1 & -1 \\ 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 0 \end{pmatrix} = I.$$

Аналогичное равенство справедливо и при любом чётном  $n$ .

Вычислим теперь  $\text{adj } A_{2k+1}$ . Так как  $|A_{2k}| = 1$ , то  $\text{rk } A_{2k+1} = 2k$ . Ясно также, что  $A_{2k+1}v = 0$ , если  $v$  — столбец  $(1, -1, 1, -1, \dots)$ . Следовательно, столбцы матрицы  $\text{adj } A_{2k+1}$  имеют вид  $\lambda v$ . Кроме того,  $b_{11} = |A_{2k}| = 1$  и матрица  $B$  симметрична (см. задачу 2.11). Поэтому

$$B = \begin{pmatrix} 1 & -1 & 1 & \dots \\ -1 & 1 & -1 & \dots \\ 1 & -1 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

**2.13.** а) Так как  $[\text{adj}(A - \lambda I)](A - \lambda I) = |A - \lambda I| \cdot I$  — диагональная матрица, то  $\left(\sum_{k=0}^{n-1} \lambda^k A_k\right)(A - \lambda I) = \sum_{k=0}^{n-1} \lambda^k A_k A - \sum_{k=1}^n \lambda^k A_{k-1} = A_0 A - \lambda^n A_{n-1} + \sum_{k=1}^{n-1} \lambda^k (A_k A - A_{k-1})$  — диагональная матрица.

б) Матрица  $A_{n-1}$  равна  $\pm I$ , поэтому при  $s = 1$  утверждение верно. Кроме того,  $A_{n-s-1} = dI + A_{n-s}A$ , поэтому из справедливости утверждения для  $s$  следует его справедливость для  $s + 1$ .

**Замечание.** В частности, матрица  $\text{adj } A = A_0$  выражается как многочлен степени  $n - 1$  от  $A$ .

**2.14.** Матрица  $A = \|a_{ij}\|_1^n$ , в каждой строке и каждом столбце которой стоит ровно один положительный элемент, обладает требуемым свойством; при этом  $A^{-1} = \|b_{ij}\|_1^n$ , где  $b_{ij} = 1/a_{ij}$ , если  $a_{ij} \neq 0$ .

Докажем, что другие матрицы не могут обладать требуемым свойством. Предположим, что матрица  $A = \|a_{ij}\|_1^n$  обладает требуемым свойством и  $a_{ir}, a_{is} > 0$ . Пусть  $A^{-1} = \|b_{ij}\|_1^n$ . Тогда при  $i \neq j$  имеет место равенство  $0 = \sum_{k=1}^n a_{ik} b_{kj} = a_{ir} b_{rj} + a_{is} b_{sj} + \dots$ . Поэтому  $b_{rj} = b_{sj} = 0$  при  $j \neq i$ , т.е. строки  $r$  и  $s$  имеют единственные ненулевые элементы  $b_{ri}$  и  $b_{si}$ . Следовательно, строки  $r$  и  $s$  пропорциональны и матрица  $A^{-1}$  вырождена. Получено противоречие.

**2.15.** Площадь треугольника с вершинами в точках  $(x_i, y_i)$ ,  $i = 1, 2, 3$ , равна  $\pm \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$ . Точка пересечения прямых  $a_i x + b_i y = c_i$  и  $a_j x + b_j y = c_j$  имеет координаты

$$\left( \frac{\begin{vmatrix} a_i & c_i \\ a_j & c_j \end{vmatrix}}{\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix}}, \frac{\begin{vmatrix} c_i & b_i \\ c_j & b_j \end{vmatrix}}{\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix}} \right).$$

Из формулы  $A \text{adj } A = |A| \cdot I$  следует, что  $|\text{adj } A| = |A|^{n-1}$ , где  $n$  — порядок

матрицы  $A$ . Поэтому

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}^2 = \begin{vmatrix} |b_2 & c_2| & -|a_2 & c_2| & |a_2 & b_2| \\ |b_3 & c_3| & -|a_3 & c_3| & |a_3 & b_3| \\ |b_1 & c_1| & -|a_1 & c_1| & |a_1 & b_1| \end{vmatrix}.$$

В последнем определителе знаки можно убрать, воспользовавшись задачей 1.29.

**2.16.** Докажем, что  $A^{-1} = \|b_{ij}\|_1^n$ , где  $b_{ij} = \varepsilon^{-ij}/n$ . В самом деле,

$$\sum_{k=1}^n a_{ik} b_{kj} = \left( \sum_{k=1}^n \varepsilon^{ik} \varepsilon^{-kj} \right) / n = \left( \sum_{k=1}^n \varepsilon^{k(i-j)} \right) / n = \delta_{ij}.$$

**2.17.** Первое решение. [K12] Пусть  $\sigma_{n-k}^i = \sigma_{n-k}(x_1, \dots, \hat{x}_i, \dots, x_n)$ . Используя результат задачи 1.17, легко проверить, что  $(\text{adj } V)^T = \|b_{ij}\|_1^n$ , где  $b_{ij} = (-1)^{i+j} \sigma_{n-j}^i V(x_1, \dots, \hat{x}_i, \dots, x_n)$ .

Второе решение. Пусть  $a_{ij} = x_i^{j-1}$ . Тогда  $\sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n b_{kj} x_i^{k-1}$ . Пусть

$$\frac{(x_1 - x) \dots (x_n - x)}{x_j - x} = c_{1j} + c_{2j}x + \dots + c_{nj}x^{n-1},$$

т.е.  $c_{pj} = \frac{(-1)^{p+1} \sum_{1 \leq k_1 < k_2 < \dots < k_{n-p} \leq n} x_{k_1} \dots x_{k_{n-p}}}{\prod_{p \neq j} (x_p - x_j)}$ . Положим  $b_{kj} = \frac{c_{kj}}{\prod_{p \neq j} (x_p - x_j)}$ . Тогда

$$\sum_{k=1}^n a_{ik} b_{kj} = \frac{(x_1 - x_i) \dots (x_n - x_i)}{x_j - x_i} \cdot \frac{1}{\prod_{p \neq j} (x_p - x_j)} = \delta_{ij}.$$

**2.18.** Покажем, что матрица  $\|b_{ij}\|_1^n$ , где

$$b_{ij} = \frac{\prod_k (x_j + y_k)(x_k + y_i)}{(x_j + y_i) \prod_{k \neq j} (x_j - x_k) \prod_{k \neq i} (y_i - y_k)},$$

является обратной к матрице Коши.

Ясно, что

$$\sum_s a_{is} b_{sj} = \frac{\prod_{k \neq s} (x_j + y_k) \prod_{k \neq i} (x_k + y_s)}{\prod_{k \neq j} (x_j - x_k) \prod_{k \neq s} (y_s - y_k)}.$$

Рассмотрим многочлен

$$P(x) = \frac{\prod_{k \neq s} (x_j + y_k - x) \prod_{k \neq i} (x_k + y_s)}{\prod_{k \neq j} (x_j - x_k) \prod_{k \neq s} (y_s - y_k)}.$$

Нужно доказать, что  $P(0) = \delta_{ij}$ . Рассмотрим для этого второй многочлен

$$Q(x) = \frac{\prod_{k \neq i} (x_j - x_k - x)}{\prod_{k \neq j} (x_j - x_k)}.$$

Степени обоих рассматриваемых многочленов не превосходят  $n-1$ . Легко также проверить, что  $P(x_j + y_i) = Q(x_j + y_i)$  при  $t = 1, \dots, n$ . Следовательно,  $P(x) = Q(x)$ . В частности,

$$P(0) = Q(0) = \frac{\prod_{k \neq i} (x_j - x_k)}{\prod_{k \neq j} (x_j - x_k)} = \delta_{ij}.$$

**2.19.** Ответ:  $(0 \ 1)$ . При вычёркивании последней строки и последнего столбца матрицы  $A$  получается единичная матрица, имеющая определитель 1. Ясно также, что ненулевой минор мы можем получить только при вычёркивании последней строки. Поэтому все столбцы матрицы  $\text{adj } A$ , кроме последнего, нулевые; в последнем столбце стоят числа  $x_1, x_2, \dots, x_{n-1}, 1$ . Теперь из равенства  $A(\text{adj } A) = 0$  получаем  $x_1 = 1, x_2 = 1, \dots, x_{n-1} = 1$ .

**2.20.** Пусть  $n$  — порядок матрицы  $A$ , а  $A_{n-1}$  — матрица, образованная первыми  $n-1$  столбцами матрицы  $A$ . Условие, что сумма элементов каждой строки матрицы  $A$  равна 0, влечёт, что  $A = (A_{n-1} \ 0) \begin{pmatrix} I_n & -\mathbf{1} \\ 0 & 0 \end{pmatrix}$ , где  $-\mathbf{1}$  — столбец чисел  $-1$ . Согласно задаче 2.19  $\text{adj} \begin{pmatrix} I_n & -\mathbf{1} \\ 0 & 0 \end{pmatrix} = (0 \ \mathbf{1})$ , поэтому  $\text{adj } A = (0 \ \mathbf{1}) \text{adj}(A_{n-1} \ 0)$ . Для матрицы  $(A_{n-1} \ 0)$  ненулевой минор может получиться только при вычёркивании последнего столбца, поэтому

$$\text{adj } A = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_1 & a_2 & \dots & a_n \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_n \end{pmatrix}.$$

**3.1.** Согласно теореме 1.3.1.3

$$\begin{vmatrix} I & A \\ A^T & I \end{vmatrix} = |I - A^T A| = (-1)^n |A^T A - I|.$$

Остаётся воспользоваться результатом задачи 2.1 при  $\lambda = -1$ , а затем результатом задачи 2.4.

**3.2.** Рассмотрим матрицу  $X = \begin{pmatrix} I_n & A \\ B & I_m \end{pmatrix}$ . Согласно теореме 1.3.1.1  $|X| = |I_n| \cdot |I_m - B I_n^{-1} A| = |I_m - BA|$ , а согласно теореме 1.3.1.2  $|X| = |I_n - AB|$ .

**3.3.** Ясно, что

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_m & A'_{12} \\ 0 & A'_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ A_{21} & I_n \end{pmatrix},$$

поэтому  $\det A \det A'_{22} = \det A_{11}$ .

**3.4.** Поменяем местами первый и последний столбец матрицы  $A_n$ , затем второй и предпоследний и т.д. В результате получим матрицу  $A'_n$ , причём  $|A_n| =$

$\varepsilon|A'_n|$ , где  $\varepsilon = (-1)^{n(n-1)/2}$ . Преобразуем матрицу  $A'_n$  следующим образом. Вычтем 1-й столбец из 2-го, ...,  $(n-1)$ -й из  $n$ -го; затем вычтем 1-ю строку из 2-й, ...,  $(n-1)$ -ю из  $n$ -й. В результате получим матрицу вида  $\begin{pmatrix} a & -u \\ v & B \end{pmatrix}$ , где  $a = (n^2 - n + 2)/2$ ,  $u$  — строка  $(n-1, n-2, \dots, 3, 2, 1)$ ;  $v$  — столбец  $(n, n-1, \dots, 4, 3, 2)$ ;  $B$  — кососимметрическая матрица с элементами  $-1$  над главной диагональю. Поэтому согласно теореме 1.3.1.3  $A_n = \varepsilon(a \det B + u(\operatorname{adj} B)v)$ . Разберем отдельно случаи чётного и нечётного  $n$ .

Пусть  $n = 2k$ . Тогда  $\varepsilon = (-1)^k$  и  $\det B = 0$  (см. задачу 1.1). Матрица  $\operatorname{adj} B$  вычислена в задаче 2.12. Легко проверить, что  $u(\operatorname{adj} B) = k(1, -1, 1, -1, \dots)$ , поэтому  $u(\operatorname{adj} B)v = k(k+1)$ . Следовательно,  $|A_{2k}| = (-1)^k k(k+1)$ .

Пусть теперь  $n = 2k+1$ . Тогда  $\varepsilon = (-1)^k$ ,  $a = 2k^2 + k + 1$  и  $\det B = 1$  (см. задачу 1.3). Легко проверить, что  $u(\operatorname{adj} B) = (-k, k+1, -k, k+1, \dots) = -k(1, -1, 1, -1, \dots) + (0, 1, 0, 1, \dots)$ , поэтому  $u(\operatorname{adj} B)v = -k^2 + k(k+1) = k$ . Следовательно,  $|A_{2k+1}| = (-1)^k (2k^2 + 2k + 1)$ .



## Глава 2.

# Линейные пространства

### Историческая справка

Понятие векторного пространства появилось значительно позже, чем понятие определителя. Лейбницу принадлежат серьезные заслуги и в создании этого понятия. Его не устраивало то, что на языке алгебры в геометрии можно было описывать лишь различные величины, но не положения точек и не направления прямых. Лейбниц начал рассматривать наборы точек  $A_1 \dots A_n$  и считал, что  $A_1 \dots A_n = X_1 \dots X_n$ , если длины отрезков  $A_i A_j$  и  $X_i X_j$  равны для всех  $i$  и  $j$  (он использовал, конечно, несколько иные обозначения). На этом языке уравнение  $AB = AY$  задает сферу радиуса  $AB$  с центром  $A$ , а уравнение  $AY = BY = CY$  задает прямую, перпендикулярную плоскости  $ABC$ .

Хотя Лейбниц и рассматривал пары точек, эти пары вовсе не соответствовали векторам: учитывались лишь длины отрезков, но не их направления; пары  $AB$  и  $BA$  не различались.

Эти работы Лейбница оставались неопубликованными более 100 лет после его смерти. Они были опубликованы в 1833 году, и за развитие этих идей была назначена премия. В 1845 году об этой премии Мёбиус сообщил Грассману, через год Грассман представил свою работу, и премия была присуждена ему. Книгу Грассмана напечатали, но она никого не заинтересовала.

Важным шагом в создании понятия векторного пространства было геометрическое представление комплексных чисел. Вычисления с комплексными числами настоятельно требовали обоснования их применения и создания достаточно строгой их теории. Еще в XVII веке Джон Валлис пытался представить комплексные числа геометрически,

но это ему не удалось. В 1799–1831 годах шесть математиков независимо опубликовали работы, содержащие геометрическую интерпретацию комплексных чисел. Наибольшее влияние на математиков оказала работа Гаусса, опубликованная в 1831 году. Сам Гаусс не считал геометрическую интерпретацию достаточно убедительным обоснованием существования комплексных чисел, потому что он в то время уже пришёл к созданию неевклидовой геометрии.

Решающий вклад в создание понятия  $n$ -мерного пространства сделали сразу два математика — Гамильтон и Грассман. Их подходы были принципиально различны, разным было и влияние их работ на развитие математики. Работы Грассмана содержали глубокие идеи, которые оказали большое влияние на развитие алгебры и алгебраической геометрии. Но его книги были трудны для понимания, и значение его идей было осознано далеко не сразу. Развитие линейной алгебры пошло в основном по пути, указанному Гамильтоном.

### Гамильтон (1805–1865)

Ирландский математик и астроном сэр Уильям Роуэн Гамильтон, член многих академий, в том числе и Петербургской, родился в 1805 году в Дублине. С трёх лет он воспитывался у своего дяди, священника, и к 13 годам изучил 13 языков, а в 16 лет он прочитал «Небесную механику» Лапласа.

В 1823 году Гамильтон поступил в Тринити колледж в Дублине, по окончании которого ему предложили стать профессором астрономии в Дублинском университете и Королевским астрономом Ирландии. Большую известность Гамильтону принесло теоретическое предсказание двух ранее неизвестных явлений в оптике, вскоре после этого обнаруженных экспериментально. В 1837 году он стал президентом Ирландской академии; в том же году он опубликовал свои работы, в которых комплексные числа вводились как пары действительных чисел. Этому открытию поначалу не придали большого значения. Всех математиков, кроме Гаусса и Бойяи, геометрическая интерпретация комплексных чисел вполне устраивала. И лишь после того как неевклидова геометрия стала достаточно широко известна, математики заинтересовались интерпретацией комплексных чисел как пар вещественных чисел.

Гамильтон вскоре осознал возможности, которые давало его открытие. В 1841 году он пришёл к рассмотрению наборов  $(a_1, \dots, a_n)$ , где  $a_i$  — вещественные числа. Именно на этой идее основан наиболее общепринятый подход к понятию линейного пространства.

С особым увлечением Гамильтон занимался тройками действительных чисел; ему хотелось получить трёхмерный аналог комплексных чисел. Его возбужденное состояние передалось его детям. Вспоминая об этом, в 1865 г. Гамильтон писал своему сыну Арчибальду, что когда он спускался к завтраку, они кричали:

— Ну, папа, ты научился перемножать тройки?

На что он принужден был отвечать, грустно качая головой:

— Нет, я умею только складывать и вычитать их.

Гамильтон, конечно, имел в виду умножение, обладающее свойством  $\|ab\| = \|a\| \cdot \|b\|$ .

Эти напряженные занятия привели к тому, что 16 октября 1843 года, во время прогулки, Гамильтон почти воочию увидел символы  $i, j, k$  и соотношения  $i^2 = j^2 = k^2 = ijk = -1$ . Последние 25 лет своей жизни Гамильтон занимался почти исключительно кватернионами и их применениями в геометрии, механике, астрономии. Он забросил свои блестящие исследования по физике и изучал, например, возведение кватернионов в кватернионную степень. О кватернионах он напечатал 2 книги и более 100 статей. Занимаясь кватернионами, Гамильтон дал определение скалярного и векторного произведения векторов в трёхмерном пространстве.

### Грассман (1809-1877)

Внешняя сторона жизни Германа Грассмана была далеко не такой блестящей, как у Гамильтона. До конца своей жизни он оставался учителем гимназии в своём родном городе Штеттине. Несколько раз он пытался получить университетскую должность, но безуспешно. Гамильтон, прочитав книгу Грассмана, назвал его величайшим немецким гением. По поводу той же книги через 30 лет после её публикации издатель писал Грассману: «Вашей книги в продаже больше нет. Так как её почти никто не покупал, то 600 экземпляров были использованы в качестве макулатуры, а оставшиеся несколько экземпляров распроданы все, кроме одного, который остаётся в нашей библиотеке». Следующая книга Грассмана, как считал он сам, пользовалась ещё меньшим успехом. Идеи Грассмана получили распространение лишь к концу его жизни. Сам он в это время уже потерял контакты с математиками и утратил интерес к геометрии. Последние годы жизни Грассман занимался преимущественно санскритом. Он сделал перевод «Ригведы» (более 1000 страниц) и составил к ней словарь (около 2000 страниц). За эти труды Американское общество востоковедов избрало его своим членом. В современных исследованиях «Ригведы» их часто цитируют.

Грассмана можно, пожалуй, назвать самоучкой. Он окончил университет в Берлине, но изучал там филологию и теологию. Отец его был преподавателем математики в Штеттине, но книги, написанные им, Грассман прочитал только в университете. Впоследствии он говорил, что многие идеи он почерпнул из этих книг и только развил их.

В 1832 году Грассман фактически пришёл к векторной записи законов механики; это сильно упрощало многие расчёты. Он заметил коммутативность и ассоциативность сложения векторов и выделил эти свойства в явном виде. Впоследствии Грассман излагал свою теорию в весьма общем виде, для произвольных систем, обладающих определёнными свойствами. Это сильно затрудняло понимание его книг; почти никто кроме него ещё не осознавал значения коммутативности, ассоциативности и дистрибутивности в алгебре.

Грассман определил геометрическое произведение двух векторов как параллелограмм, натянутый на эти векторы. Равновеликие и одинаково ориентированные параллелограммы, параллельные одной плоскости, он считал эквивалентными. Впоследствии он по аналогии ввел геометрическое произведение  $r$  векторов в  $n$ -мерном пространстве. Это произведение он рассматривал как геометрический объект, координатами которого являются миноры порядка  $r$  матрицы размера  $r \times n$ , составленной из координат данных векторов.

В работах Грассмана было фактически построено понятие линейного пространства со всеми его атрибутами: он дал определение подпространства и линейной зависимости векторов.

В 40-е годы XIX века математики оказались неподготовленными к восприятию идей Грассмана. Свою первую книгу он послал Гауссу и получил в ответ записку, в которой Гаусс благодарил его и писал, что подобными вещами он занимался полвека тому назад и кое-что недавно опубликовал на эту тему. Мёбиус, в ответ на просьбу Грассмана написать рецензию на его книгу, сообщил ему, что не понимает философской части книги и поэтому не может прочитать её до конца. Впоследствии Мёбиус говорил, что он знает лишь одного математика, который прочитал книгу Грассмана до конца, — это Бретшнейдер. Получив премию за развитие идей Лейбница, Грассман обратился к министру культуры с просьбой дать ему университетскую должность, и его работы послали на рецензию Куммеру. В рецензии было сказано, что в них недостаёт ясности. На просьбу Грассмана ответили отказом.

В 60-е и 70-е годы XIX века многие математики своими путями приходят к идеям, аналогичным идеям Грассмана. Его работы получают высокую оценку Кремоны, Ганкеля, Клебша и Клейна, но самого Грассмана математика уже больше не интересует.

## 2.1. Двойственное пространство. Ортогональное дополнение

*Линейным (или векторным) пространством* над полем  $k$  называют множество  $V$ , в котором определено сложение элементов  $V$  и умножение элементов поля на элементы  $V$ , причём выполняются следующие условия:

- Относительно сложения  $V$  — абелева группа, т.е. сложение коммутативно ( $v + w = w + v$ ); в  $V$  есть нулевой элемент  $0$ , для которого  $0 + v = v$ ; для каждого  $v \in V$  существует противоположный элемент  $-v \in V$ , для которого  $v + (-v) = 0$ .
- $\lambda(v + w) = \lambda v + \lambda w$  для любых  $\lambda \in k, v, w \in V$ .
- $(\lambda + \mu)v = \lambda v + \mu v$  для любых  $\lambda, \mu \in k, v \in V$ .
- $(\lambda\mu)v = \lambda(\mu v)$  для любых  $\lambda, \mu \in k, v \in V$ .
- $1v = v$  для любого  $v \in V$ .

Элементы линейного пространства  $V$  называют *векторами*.

Примером линейного пространства над полем  $k$  служит множество строк  $(a_1, \dots, a_n)$  фиксированной длины  $n$ , состоящих из элементов поля  $k$ , в котором требуемые операции определены следующим образом:

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n),$$

$$\lambda(a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n).$$

Подмножество  $U \subset V$  в линейном пространстве называют *подпространством*, если  $U$  является подгруппой в  $V$  и  $\lambda v \in U$  для любых  $\lambda \in k$  и  $v \in U$ .

### 2.1.1. Линейно независимые векторы

Векторы  $v_1, \dots, v_k$  называют *линейно зависимыми*, если  $\lambda_1 v_1 + \dots + \lambda_k v_k = 0$  для некоторого набора чисел  $\lambda_1, \dots, \lambda_k$ , среди которых есть число, отличное от нуля. Выбрав число  $\lambda_i \neq 0$ , можно получить равенство

$$v_i = \mu_1 v_1 + \dots + \mu_{i-1} v_{i-1} + \mu_{i+1} v_{i+1} + \dots + \mu_k v_k,$$

где  $\mu_j = -\lambda_j/\lambda_i$ . Другими словами, векторы  $v_1, \dots, v_k$  линейно зависимы тогда и только тогда, когда один из них линейно выражается через другие.

Подпространство, состоящее из векторов вида  $\lambda_1 v_1 + \dots + \lambda_k v_k$ , где  $v_1, \dots, v_k$  — фиксированные векторы, а  $\lambda_1, \dots, \lambda_k$  пробегают все числа из поля  $K$ , называют подпространством, порождённым  $v_1, \dots, v_k$ ; мы будем обозначать это подпространство  $\langle v_1, \dots, v_k \rangle$ .

Пространство, порождённое конечным набором векторов, называют *конечномерным*. В этой книге рассматриваются только конечномерные векторные пространства.

Набор векторов  $e_1, \dots, e_n$  называют *базисом* пространства  $V$ , если любой вектор  $v \in V$  единственным образом представляется в виде  $v = \lambda_1 e_1 + \dots + \lambda_n e_n$ . В частности, базисные векторы линейно независимы.

**Теорема 2.1.1.1.** *Любое конечномерное линейное пространство  $V$  обладает базисом,*

**Доказательство.** Пусть  $V = \langle v_1, \dots, v_k \rangle$ . Если, например, вектор  $v_k$  линейно выражается через  $v_1, \dots, v_{k-1}$ , то  $V = \langle v_1, \dots, v_{k-1} \rangle$ . Поэтому можно считать, что  $V = \langle e_1, \dots, e_n \rangle$ , где векторы  $e_1, \dots, e_n$  линейно независимы. Но для линейно независимых векторов представление  $v = \lambda_1 e_1 + \dots + \lambda_n e_n$  единственно. Действительно, если  $\sum \lambda_i e_i = \sum \mu_i e_i$ , то  $\sum (\lambda_i - \mu_i) e_i = 0$ , поэтому  $\lambda_i = \mu_i$  для всех  $i$ .  $\square$

**Теорема 2.1.1.2.** *Все базисы конечномерного линейного пространства  $V$  содержат одно и то же число векторов.*

**Доказательство.** Достаточно проверить, что если  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_m$  — два базиса пространства  $V$ , то  $m \leq n$ . Предположим, что  $m > n$ . Каждый вектор  $\varepsilon_j$  можно выразить через  $e_1, \dots, e_n$ :  $\varepsilon_j = \mu_{1j} e_1 + \dots + \mu_{nj} e_n$ . Рассмотрим матрицу  $\begin{pmatrix} \mu_{11} & \dots & \mu_{1m} \\ \dots & \dots & \dots \\ \mu_{n1} & \dots & \mu_{nm} \end{pmatrix}$ . Число строк этой матрицы равно  $n$ , поэтому её ранг не превосходит  $n$  (см. §1.2.2). Число столбцов равно  $m > n$ , поэтому столбцы линейно зависимы, т.е. существует набор  $(\lambda_1, \dots, \lambda_m) \neq (0, \dots, 0)$ , для которого  $\lambda_1 \mu_{1i} + \dots + \lambda_m \mu_{mi} = 0$  для  $i = 1, \dots, n$ . Это означает, что

$$\lambda_1 \varepsilon_1 + \dots + \lambda_m \varepsilon_m = (\lambda_1 \mu_{11} + \dots + \lambda_m \mu_{1n}) e_1 + \dots + (\lambda_1 \mu_{n1} + \dots + \lambda_m \mu_{nm}) e_n = 0,$$

т.е. векторы  $\varepsilon_1, \dots, \varepsilon_m$  линейно зависимы. Приходим к противоречию.  $\square$

Количество векторов в базисе конечномерного векторного пространства  $V$  называют *размерностью*  $V$ .

Если пространство размерности  $n$  является подпространством в пространстве размерности  $m$ , то говорят, что *коразмерность* этого подпространства равна  $m - n$ .

### 2.1.2. Двойственное пространство

Линейному пространству  $V$  над полем  $K$  можно сопоставить линейное пространство  $V^*$ , элементами которого являются *линейные функции* на  $V$ , т.е. такие отображения  $f : V \rightarrow K$ , что  $f(\lambda_1 v_1 + \lambda_2 v_2) = f(\lambda_1 v_1) + f(\lambda_2 v_2)$  для любых  $\lambda_1, \lambda_2 \in K$  и  $v_1, v_2 \in V$ . Пространство  $V^*$  называют *двойственным* (или *сопряжённым*) к  $V$ .

Базису  $e_1, \dots, e_n$  пространства  $V$  можно сопоставить базис  $e_1^*, \dots, e_n^*$  пространства  $V^*$  следующим образом:  $e_i^*(e_j) = \delta_{ij}$ . Докажем, что  $e_1^*, \dots, e_n^*$  действительно базис. Если  $f \in V^*$  и  $v = \sum x_i e_i$ , то  $f(v) = f(\sum x_i e_i) = \sum x_i f(e_i) = \sum a_i x_i = \sum a_i e_i^*(v)$ . Поэтому  $f = \sum a_i e_i^*$ . Остаётся проверить линейную независимость векторов  $e_1^*, \dots, e_n^*$ . Пусть  $\sum a_i e_i^* = 0$ . Тогда  $0 = (\sum a_i e_i^*)(e_j) = a_j$ , т.е.  $a_1 = \dots = a_n = 0$ .

Итак, если в пространстве  $V$  выбран базис  $e_1, \dots, e_n$ , то можно построить изоморфизм  $g : V \rightarrow V^*$  следующим образом:  $g(e_i) = e_i^*$ . Но при выборе другого базиса получим, вообще говоря, другой базис (см. § 2.1.4), т.е. построенный изоморфизм не канонический.

Можно, однако, построить канонический изоморфизм пространств  $V$  и  $(V^*)^*$ , сопоставив вектору  $v \in V$  такой элемент  $v' \in (V^*)^*$ , что  $v'(f) = f(v)$  для любого  $f \in V^*$ . Линейность и взаимная однозначность этого отображения проверяются очевидным образом.

**Замечание.** Элементы пространства  $V^*$  иногда называют *ковекторами*. Употребима также следующая терминология: элементы  $V$  называют *контравариантными векторами*, а элементы  $V^*$  называют *ковариантными векторами*.

При переходе к двойственному пространству подпространства соответствуют факторпространствам (и наоборот) в следующем смысле. Подпространству  $U \subset V$  естественно сопоставить пространство линейных функций  $V$ , обращающихся в нуль на  $U$ . Но это в точности то же самое, что пространство линейных функций на  $V/U$ , т.е. пространство  $(V/U)^*$ .

### 2.1.3. Двойственный оператор

Линейному оператору  $A : V_1 \rightarrow V_2$  можно сопоставить *двойственный* (сопряжённый) оператор  $A^* : V_2^* \rightarrow V_1^*$  следующим образом:  $(A^* f_2)(v_1) = f_2(Av_1)$  для любых  $f_2 \in V_2^*$  и  $v_1 \in V_1$ .

Для величины  $f(v)$ , где  $v \in V$  и  $f \in V^*$ , удобно ввести более симметричное обозначение  $\langle f, v \rangle$ . Определение оператора  $A^*$  в этих обозначениях запишется следующим образом:  $\langle A^* f_2, v_1 \rangle = \langle f_2, Av_1 \rangle$ .

Если в пространстве  $V_1$  выбран базис  $\{e_\alpha\}$ , а в пространстве  $V_2$  выбран базис  $\{\varepsilon_\beta\}$ , то оператору  $A$  можно сопоставить матрицу  $\|a_{ij}\|$ , где  $Ae_j = \sum_i a_{ij} \varepsilon_i$ . Аналогичным образом оператору  $A^*$  можно сопоставить матрицу  $\|a_{ij}^*\|$  относительно базисов  $\{e_\alpha^*\}$  и  $\{\varepsilon_\beta^*\}$ . Докажем, что  $\|a_{ij}^*\| = \|a_{ij}\|^T$ . В самом деле, с одной стороны,  $\langle \varepsilon_k^*, Ae_j \rangle = \sum_i a_{ij} \langle \varepsilon_k^*, \varepsilon_i \rangle = a_{kj}$ . С другой стороны,  $\langle \varepsilon_k^*, Ae_j \rangle = \langle A^* \varepsilon_k^*, e_j \rangle = \sum_p a_{pk}^* \langle \varepsilon_k^*, e_j \rangle = a_{jk}^*$ . Следовательно,  $a_{jk}^* = a_{kj}$ .

### 2.1.4. Отождествление $V$ и $V^*$ при наличии метрики

Пусть в пространстве  $V$  выбраны базисы  $\{e_\alpha\}$  и  $\{\varepsilon_\beta\}$ ;  $A$  — матрица перехода от базиса  $\{e_\alpha\}$  к базису  $\{\varepsilon_\beta\}$ , т.е.  $\varepsilon_i = Ae_i$ . Пусть, далее,  $A^*$  — матрица перехода от базиса  $\{e_\alpha^*\}$  к базису  $\{\varepsilon_\beta^*\}$ . Докажем, что  $A^* = (A^T)^{-1}$ . В самом деле,  $\langle e_i^*, e_j \rangle = \delta_{ij} = \langle \varepsilon_i^*, \varepsilon_j \rangle = \langle A^* e_i^*, Ae_j \rangle = \langle e_i^*, (A^*)^T Ae_j \rangle$ , поэтому  $(A^*)^T A = I$ , т.е.  $A^* = (A^T)^{-1}$ .

Теперь мы готовы ответить на вопрос, затронутый в § 2.1.2: когда совпадают изоморфизмы  $V \rightarrow V^*$ , построенные с помощью базисов  $\{e_\alpha\}$  и  $\{\varepsilon_\beta\}$ ? Пусть отображения  $f, g : V \rightarrow V^*$  заданы формулами  $f(e_i) = e_i^*$  и  $g(\varepsilon_i) = \varepsilon_i^*$ ;  $A$  — матрица перехода от базиса  $\{e_\alpha\}$  к  $\{\varepsilon_\beta\}$ . Тогда  $g(e_i) = g(A^{-1} \varepsilon_i) = A^{-1} g(\varepsilon_i) = A^{-1} \varepsilon_i^* = A^{-1} (A^T)^{-1} e_i^*$ . Поэтому отображения  $f$  и  $g$  совпадают тогда и только тогда, когда  $A^T = A^{-1}$ , т.е.  $A$  — ортогональная матрица.

Скалярное произведение в пространстве  $V$  над полем действительных чисел позволяет выделить множество ортонормированных реперов, а матрицы перехода между ортонормированными реперами ортогональны. Поэтому скалярное произведение позволяет построить однозначно определённый изоморфизм  $V \rightarrow V^*$ . При этом изоморфизме вектору  $v \in V$  сопоставляется функция  $v^*$ :  $v^*(x) = (v, x)$  для любого  $x \in V$ .



### 2.1.5. Системы линейных уравнений

Рассмотрим систему линейных уравнений

$$f_1(x) = b_1, \dots, f_m(x) = b_m. \quad (2.1)$$

Можно считать, что ковекторы  $f_1, \dots, f_k$  линейно независимы и  $f_i = \sum_{j=1}^k \lambda_{ij} f_j$  при  $i > k$ . Если  $x_0$  — решение системы (2.1), то  $f_i(x_0) = \sum_{j=1}^k \lambda_{ij} f_j(x_0)$  при  $i > k$ , т.е.

$$b_i = \sum_{j=1}^k \lambda_{ij} b_j \quad \text{при } i > k. \quad (2.2)$$

Докажем, что если выполнены условия (2.2), то система (2.1) совместна. Дополним ковекторы  $f_1, \dots, f_k$  до базиса и рассмотрим двойственный базис  $e_1, \dots, e_n$ . В качестве решения можно взять  $x_0 = b_1 e_1 + \dots + b_k e_k$ ; общее решение системы (2.1) имеет вид  $x_0 + t_1 e_{k+1} + \dots + t_{n-k} e_n$ , где  $t_1, \dots, t_{n-k}$  — любые числа.

**Теорема 2.1.5.1.** Если система (2.1) совместна, то она имеет решение  $x = (x_1, \dots, x_n)$ , где  $x_i = \sum_{j=1}^k c_{ij} b_j$ , причём числа  $c_{ij}$  не зависят от  $b_j$ .

Для доказательства достаточно рассмотреть координаты вектора  $x_0 = b_1 e_1 + \dots + b_k e_k$  относительно нужного базиса.

**Теорема 2.1.5.2.** Если  $f_i(x) = \sum_{j=1}^n a_{ij} x_j$ , где  $a_{ij} \in \mathbb{Q}$  и ковекторы  $f_1, \dots, f_m$  образуют базис (в частности,  $m = n$ ), то система (2.1) имеет решение  $x_i = \sum_{j=1}^n c_{ij} b_j$ , где числа  $c_{ij}$  рациональны и не зависят от  $b_j$ ; это решение единственно.

**Доказательство.** Так как

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

то

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Если элементы матрицы  $A$  рациональны, то элементы матрицы  $A^{-1}$  тоже рациональны.  $\square$

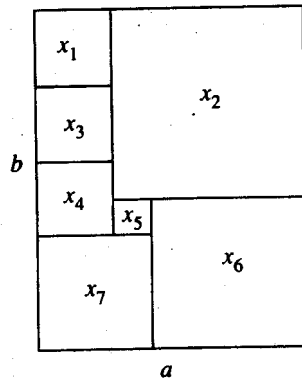


Рис. 2.1.

### 2.1.6. Разрезание прямоугольника на квадраты

Полученные в §2.1.5 результаты имеют несколько неожиданное применение.

**Теорема 2.1.6.1.** *Если прямоугольник со сторонами  $a$  и  $b$  произвольным образом разрезан на квадраты со сторонами  $x_1, \dots, x_n$ , то  $x_i/a \in \mathbb{Q}$  и  $x_i/b \in \mathbb{Q}$  для всех  $i$ .*

**Доказательство.** Например, для рис. 2.1 можно написать следующую систему уравнений:

$$\begin{aligned}
 x_1 + x_2 &= a, & x_1 + x_3 + x_4 + x_7 &= b, \\
 x_3 + x_2 &= a, & x_2 + x_5 + x_7 &= b, \\
 x_4 + x_2 &= a, & x_2 + x_6 &= b, \\
 x_4 + x_5 + x_6 &= a, \\
 x_6 + x_7 &= a.
 \end{aligned} \tag{2.1}$$

Аналогичную систему уравнений можно написать и для любого другого разбиения прямоугольника на квадраты. Заметим также, что если система, соответствующая некоторому рисунку, имеет еще одно положительное решение, то и этому решению можно сопоставить разбиение прямоугольника на квадраты, и для любого разбиения справедливо равенство  $x_1^2 + \dots + x_n^2 = ab$  (равенство площадей).

Предположим сначала, что система (2.1) имеет единственное решение. Тогда  $x_i = \lambda_i a + \mu_i b$ , причём  $\lambda_i, \mu_i \in \mathbb{Q}$ . Подставив эти значения во все уравнения системы (2.1), получим равенства вида  $p_j a + q_j b = 0$ , где  $p_j, q_j \in \mathbb{Q}$ . Если  $p_j = q_j = 0$  для всех  $j$ , то система (2.1) совместна для всех  $a$  и  $b$ . Поэтому при любом достаточно малом изменении чисел  $a$  и  $b$  система (2.1) имеет положительное решение  $x_i = \lambda_i a + \mu_i b$ , а значит, существует соответствующее разбиение прямоугольника. Следовательно, для всех  $a$  и  $b$ , лежащих в некоторых интервалах, выполняется равенство

$$\left(\sum \lambda_i^2\right) a^2 + 2 \left(\sum \lambda_i \mu_i\right) ab + \left(\sum \mu_i^2\right) b^2 = ab.$$

Поэтому  $\sum \lambda_i^2 = \sum \mu_i^2 = 0$ , а значит,  $\lambda_i = \mu_i = 0$  для всех  $i$ . Получено противоречие, поэтому в одном из равенств  $p_j a + q_j b = 0$  одно из чисел  $p_j$  и  $q_j$  отлично от нуля. Следовательно,  $b = ra$ , где  $r \in \mathbb{Q}$ , и  $x_i = (\lambda_i + r\mu_i)a$ , где  $\lambda_i + r\mu_i \in \mathbb{Q}$ .

Докажем теперь, что размерность пространства решений системы (2.1) не может быть больше нуля. Решения системы (2.1) имеют вид  $x_i = \lambda_i a + \mu_i b + \alpha_{1i} t_1 + \dots + \alpha_{ki} t_k$ , где  $t_1, \dots, t_k$  принимают произвольные значения. Поэтому равенство

$$\sum_i (\lambda_i a + \mu_i b + \alpha_{1i} t_1 + \dots + \alpha_{ki} t_k)^2 = ab \quad (2.2)$$

должно выполняться для всех  $t_1, \dots, t_k$ , лежащих в некоторых интервалах. Левая часть равенства (2.2) является квадратичной функцией от  $t_1, \dots, t_k$ ; эта функция имеет вид  $\sum_{p,i} \alpha_{pi}^2 t_i^2 + \dots$ , поэтому она не может оставаться постоянной при всех малых изменениях чисел  $t_1, \dots, t_k$ .  $\square$

### 2.1.7. Количество треугольников среди частей, на которые прямые разрезают плоскость

Приведем ещё одно довольно неожиданное приложение свойств систем линейных уравнений.

**Теорема 2.1.7.1 ([Sh1]).** *Предположим, что на плоскости дано  $n \geq 3$  попарно не параллельных прямых, никакие три из которых не пересекаются в одной точке. Тогда среди фигур, на которые эти прямые разрезают плоскость, есть по крайней мере  $n - 2$  треугольника.*

**Доказательство.** [B2], [Gr5]. Прежде всего заметим, что среди указанных фигур есть хотя бы один треугольник. При  $n = 3$  это очевидно. Шаг индукции тоже очевиден: дополнительная прямая либо не пересекает треугольник, либо отсекает от него меньший треугольник.

Пусть прямые  $l_1, \dots, l_n$  заданы уравнениями  $a_i x + b_i y = c_i$ ,  $i = 1, \dots, n$ . Наряду с этим набором прямых рассмотрим набор прямых  $l_1(t), \dots, l_n(t)$ , заданных уравнениями  $a_i x + b_i y = c_i + d_i t$ ; здесь  $t$  — вещественный параметр, а  $d_1, \dots, d_n$  — вещественные числа, которые будут выбраны по ходу доказательства. Задача 2.15 показывает, что площадь треугольника, образованного прямыми  $l_i(t)$ ,  $l_j(t)$  и  $l_k(t)$ , с точностью до множителя, не зависящего от  $t$ , равна

$$\left( \begin{vmatrix} a_i & b_i & d_i \\ a_j & b_j & d_j \\ a_k & b_k & d_k \end{vmatrix} t + \begin{vmatrix} a_i & b_i & c_i \\ a_j & b_j & c_j \\ a_k & b_k & c_k \end{vmatrix} \right)^2.$$

Таким образом, площадь этого треугольника не зависит от  $t$  тогда и только тогда, когда числа  $d_i$ ,  $d_j$  и  $d_k$  удовлетворяют линейному однородному уравнению

$$\begin{vmatrix} a_i & b_i & d_i \\ a_j & b_j & d_j \\ a_k & b_k & d_k \end{vmatrix} = 0. \quad (2.1)$$

Предположим, что среди частей, на которые прямые  $l_1, \dots, l_n$  разрезают плоскость, менее  $n - 2$  треугольников. Для каждой тройки прямых, ограничивающих один из этих треугольников, запишем уравнение (2.1). В результате для  $n$  неизвестных  $d_1, \dots, d_n$  мы получим не более  $n - 3$  уравнений. Добавим к этим уравнениям еще два уравнения:  $d_1 = 0$  и  $d_2 = 0$ . Полученная система уравнений имеет ненулевое решение. В дальнейшем будем считать, что  $d_1, \dots, d_n$  — именно это решение.

Равенства  $d_1 = 0$  и  $d_2 = 0$  означают, что прямые  $l_1(t)$  и  $l_2(t)$  остаются неподвижными, т.е.  $l_1(t) = l_1$  и  $l_2(t) = l_2$  для всех  $t$ . По условию есть хотя бы одно число  $d_i \neq 0$ . Поэтому для некоторого  $t$  прямая  $l_i(t)$  проходит через точку пересечения прямых  $l_1$  и  $l_2$ . Таким образом, в конфигурации прямых  $l_1(t), \dots, l_n(t)$  для каких-то значений параметра  $t$  обязательно возникают вырождения, когда через одну точку проходят по крайней мере три прямые. Пусть  $t_0$  — наименьший по абсолютной величине момент вырождения; можно считать, что  $t_0 > 0$ .

По построению в любой момент  $t$ , где  $0 \leq t \leq t_0$ , среди частей, на которые прямые  $l_1(t), \dots, l_n(t)$  разрезают плоскость, число треугольников постоянно, причём площади этих треугольников тоже постоянны. С другой стороны, при  $t \rightarrow t_0$  в конфигурации прямых, пересекающихся в точке вырождения, есть треугольник, площадь которого стремится к 0. Получено противоречие.  $\square$

### 2.1.8. Аннулятор

Как уже говорилось, не существует канонического изоморфизма между  $V$  и  $V^*$ . Есть, однако, каноническое взаимно однозначное соответствие между множеством  $k$ -мерных подпространств  $V$  и множеством  $(n - k)$ -мерных подпространств  $V^*$ . Подпространству  $W \subset V$  можно сопоставить множество

$$W^\perp = \{f \in V^* \mid \langle f, w \rangle = 0 \quad \forall w \in W\}.$$

Это множество называют *аннулятором* (или *ортогональным дополнением*) пространства  $W$ . Достаточно очевидно, что  $W^\perp$  — линейное подпространство  $V^*$ . Докажем, что  $\dim W + \dim W^\perp = \dim V$ . Пусть  $e_1, \dots, e_k$  — базис  $W$ . Дополним его до базиса  $e_1, \dots, e_n$  пространства  $V$ . Ясно, что  $e_{k+1}^*, \dots, e_n^* \in W^\perp$ . Кроме того, если  $f = \alpha_1 e_1^* + \dots + \alpha_k e_k^*$  и, например,  $\alpha_1 \neq 0$ , то  $\langle f, e_1 \rangle = \alpha_1 \neq 0$ . Поэтому  $e_{k+1}^*, \dots, e_n^*$  — базис пространства  $W^\perp$ .

Легко проверить, что если  $W_1 \subset W_2$ , то  $W_2^\perp \subset W_1^\perp$ . Ортогональное дополнение обладает также следующими свойствами:

- а)  $(W^\perp)^\perp = W$ ;
- б)  $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$  и  $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$ ;
- в) если  $V = W_1 \oplus W_2$ , то  $V^* = W_1^\perp \oplus W_2^\perp$ .

Докажем эти свойства. а) Легко проверить, что  $W \subset (W^\perp)^\perp$  и размерности этих пространств равны.

б) Пусть  $e$  — базис  $W_1 \cap W_2$ ,  $e \cup w_1$  — базис  $W_1$ ,  $e \cup w_2$  — базис  $W_2$  и  $e \cup w_1 \cup w_2 \cup v$  — базис  $V$ . Тогда  $e \cup w_1 \cup w_2$  — базис  $W_1 + W_2$ , поэтому  $v^*$  — базис  $(W_1 + W_2)^\perp$ . С другой стороны,  $w_2^* \cup v^*$  и  $w_1^* \cup v^*$  — базисы  $W_1^\perp$  и  $W_2^\perp$ , поэтому  $v^*$  — базис  $W_1^\perp \cap W_2^\perp$ . Аналогично доказывается, что  $w_1^* \cup w_2^* \cup v^*$  — базис как пространства  $(W_1 \cap W_2)^\perp$ , так и пространства  $W_1^\perp + W_2^\perp$ .

в) Согласно свойству б) имеем  $W_1^\perp + W_2^\perp = (W_1 \cap W_2)^\perp = V^*$  и  $W_1^\perp \cap W_2^\perp = (W_1 + W_2)^\perp = 0$ .

Подпространство  $W^\perp$  определено инвариантно, поэтому линейная оболочка векторов  $e_{k+1}^*, \dots, e_n^*$  не зависит от выбора базиса пространства  $V$ , а зависит лишь от самого подпространства  $W$ . А вот линейная оболочка векторов  $e_1^*, \dots, e_k^*$  зависит от выбора базиса  $e_1^*, \dots, e_n^*$ : она может быть любым  $k$ -мерным подпространством  $V^*$ , пересекающимся с  $W^\perp$  только по нулю. В самом деле, пусть  $W_1$  —  $k$ -мерное подпространство  $V^*$  и  $W_1 \cap W^\perp = 0$ . Тогда  $(W_1)^\perp$  является  $(n - k)$ -мерным подпространством  $V$ , пересекающимся с  $W$  только по нулю. Пусть  $e_{k+1}, \dots, e_n$  — базис  $(W_1)^\perp$ . Дополним его базисом  $e_1, \dots, e_k$  пространства  $W$  до базиса  $e_1, \dots, e_n$ . Тогда  $e_1^*, \dots, e_n^*$  — базис  $W_1$ .

**Теорема 2.1.8.1.** Если  $A : V \rightarrow V$  — линейный оператор и  $AW \subset W$ , то  $A^*W^\perp \subset W^\perp$ .

**Доказательство.** Пусть  $x \in W$  и  $f \in W^\perp$ . Тогда  $\langle A^*f, x \rangle = \langle f, Ax \rangle = 0$ , так как  $Ax \in W$ . Следовательно,  $A^*f \in W^\perp$ .  $\square$

### 2.1.9. Скалярное произведение в пространстве матриц

Если в пространстве  $V$  над полем действительных чисел задано скалярное произведение, то пространства  $V$  и  $V^*$  можно отождествить, сопоставив вектору  $v \in V$  линейную функцию  $f(x) = (v, x)$  (см. §2.1.4). При таком отождествлении ортогональным дополнением пространства  $W \subset V$  будет подпространство  $W^\perp$ , состоящее из векторов, ортогональных всем векторам из  $W$ .

В пространстве матриц размера  $m \times n$  можно ввести естественное скалярное произведение. Это скалярное произведение можно записать в виде  $\text{tr}(XY^T) = \sum_{i,j} x_{ij}y_{ij}$ .

**Теорема 2.1.9.1.** Пусть  $A$  — матрица размера  $m \times n$ . Если для любой матрицы  $X$  размера  $n \times m$  справедливо равенство  $\text{tr}(AX) = 0$ , то  $A = 0$ .

**Доказательство.** Пусть  $W$  — подпространство матриц вида  $\lambda A$ , где  $\lambda \in \mathbb{R}$ . Если  $\text{tr}(AX) = 0$ , то  $X^T \in W^\perp$ , поэтому  $W^\perp = V$ , а значит,  $W = 0$ .

Для доказательства можно также воспользоваться равенством  $\text{tr}(AA^T) = \sum_{ij} a_{ij}^2$ .  $\square$

### 2.1.10. Аффинное пространство

Пусть  $V$  — линейное пространство. Аффинным пространством, ассоциированным с пространством  $V$ , называют множество  $S$ , для которого задана операция сложения  $S \times V \rightarrow V$ , обладающая следующими свойствами:

- 1)  $A + (v + w) = (A + v) + w$  для любых  $A \in S$ ,  $v, w \in V$ ;
- 2)  $A + 0 = A$  для нулевого вектора  $0$ ;
- 3) для любых  $A, B \in S$  существует единственный вектор  $v$ , для которого  $A + v = B$ .

Элементы  $S$  называют *точками*. Размерность пространства  $V$  называют *размерностью* аффинного пространства  $S$ .

Вектор  $v$ , для которого  $A + v = B$ , обозначают  $\overrightarrow{AB}$ . Ясно, что  $A + \overrightarrow{AC} = (A + \overrightarrow{AB}) + \overrightarrow{BC}$  (оба эти выражения задают точку  $C$ ), поэтому из условия 1) следует, что  $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ .

Линейное пространство  $V$  можно рассматривать как аффинное пространство, считая векторы точками и определив сложение точки и вектора как сложение двух векторов. При этом  $v + \vec{v}w = w$ , т.е.  $\vec{v}w = w - v$ .

Наоборот, аффинное пространство  $S$ , в котором фиксирована некоторая точка  $O$ , можно рассматривать как линейное пространство, сопоставив каждой точке  $A$  вектор  $\vec{OA}$ . При этом равенство  $\vec{OA} + \vec{AB} = \vec{OB}$  показывает, что сложение точки и вектора превращается в сложение двух векторов.

*Аффинным подпространством* в  $S$  называют множество точек вида  $A + U$ , где  $A$  — точка в  $S$ , а  $U$  — подпространство в  $V$ .

Отображение  $F : S \rightarrow S$  называют *аффинным*, если для фиксированной точки  $A \in S$  и любого вектора  $v \in V$  имеем  $F(A + v) = F(A) + f(v)$ , где  $f$  — линейное отображение. Легко видеть, что тогда  $F(B + v) = F(B) + f(v)$  для любой точки  $B$ . Действительно, пусть  $B = A + v_0$ . Тогда  $F(B + v) = F(A + v_0 + v) = F(A) + f(v_0 + v) = F(A) + f(v_0) + f(v) = F(A + v_0) + f(v) = F(B) + f(v)$ .

Линейное отображение переводит любое линейное подпространство в некоторое линейное подпространство, поэтому аффинное отображение переводит любое аффинное подпространство в некоторое аффинное подпространство.

Если в аффинном пространстве  $S$  фиксирована точка  $O$ , то для любых  $n$  чисел  $\lambda_1, \dots, \lambda_n$  и любых  $n$  точек  $A_1, \dots, A_n$  можно определить точку  $\lambda_1 A_1 + \dots + \lambda_n A_n$  как точку  $O + \sum \lambda_i \vec{OA}_i$ . Несложно проверить, что эта точка не зависит от выбора точки  $O$  тогда и только тогда, когда  $\sum \lambda_i = 1$ . Действительно, если мы выберем другую точку  $O'$ , то получим точку  $O' + \sum \lambda_i \vec{O'A}_i = O + \vec{OO'} + (\sum \lambda_i) \vec{O'O} + \sum \lambda_i \vec{OA}_i = O + (1 - \sum \lambda_i) \vec{O'O} + \sum \lambda_i \vec{OA}_i$ . Вектор  $(1 - \sum \lambda_i) \vec{O'O}$  должен быть нулевым для всех  $O$ ; это эквивалентно тому, что  $\sum \lambda_i = 1$ .

Используя эту конструкцию, можно определить *центр масс* точек  $A_1, \dots, A_n$  с массами  $x_1, \dots, x_n$  соответственно. А именно, положим  $\lambda_i = \frac{x_i}{x_1 + \dots + x_n}$  и назовём центром масс точку  $\lambda_1 A_1 + \dots + \lambda_n A_n$ . При этом не предполагается, что массы положительны; предполагается лишь, что  $x_1 + \dots + x_n \neq 0$ . (Случай, когда  $x_1 + \dots + x_n = 0$ , соответствует бесконечно удалённой точке: см. задачу 1.8).

Пусть точки  $A_0, \dots, A_n$  в  $n$ -мерном аффинном пространстве не лежат в одной гиперплоскости. Если точка  $X$  является центром масс точек  $A_0, \dots, A_n$  с массами  $x_0, \dots, x_n$ , то набор  $(x_0 : \dots : x_n)$  называют *барицентрическими координатами* точки  $X$ . Барицентрические координаты определены однозначно с точностью до пропорциональности. Действительно, векторы  $e_i = \vec{A_0 A_i}$ ,  $i = 1, \dots, n$ , образуют базис, поэтому  $\vec{A_0 X} = \sum_{i=1}^n \lambda_i e_i$ ,

где числа  $\lambda_1, \dots, \lambda_n$  определены однозначно. Точка  $X$  является центром масс точек  $A_0, \dots, A_n$  с массами  $1 - \sum \lambda_i, \lambda_1, \dots, \lambda_n$  (или с массами, пропорциональными им). Барицентрические координаты, сумма которых равна 1, называют *абсолютными*. Для абсолютных барицентрических координат мы будем использовать обозначение  $(x_0, \dots, x_n)$ .

**Теорема 2.1.10.1.** Точки  $B_0, \dots, B_n$  лежат в одной гиперплоскости тогда и только тогда, когда матрица  $\begin{pmatrix} x_0^0 & x_1^0 & \dots & x_n^0 \\ \dots & \dots & \dots & \dots \\ x_0^n & x_1^n & \dots & x_n^n \end{pmatrix}$ , составленная из их барицентрических координат, вырожденная.

**Доказательство.** Можно считать, что барицентрические координаты абсолютные. Определитель матрицы не изменится, если к первому столбцу прибавить сумму всех остальных. В результате мы получим матрицу

$\begin{pmatrix} 1 & x_1^0 & \dots & x_n^0 \\ \dots & \dots & \dots & \dots \\ 1 & x_1^n & \dots & x_n^n \end{pmatrix}$ . Вычтем теперь первую строку из всех остальных.

Определитель при этом не изменится, а матрица примет вид

$\begin{pmatrix} 1 & x_1^0 & \dots & x_n^0 \\ 0 & x_1^1 - x_1^0 & \dots & x_n^1 - x_n^0 \\ \dots & \dots & \dots & \dots \\ 0 & x_1^n - x_1^0 & \dots & x_n^n - x_n^0 \end{pmatrix}$ . Её определитель равен определителю матрицы

$\begin{pmatrix} x_1^1 - x_1^0 & \dots & x_n^1 - x_n^0 \\ \dots & \dots & \dots \\ x_1^n - x_1^0 & \dots & x_n^n - x_n^0 \end{pmatrix}$ . Эта матрица состоит из координат векторов  $\overrightarrow{B_0B_1}, \dots, \overrightarrow{B_0B_n}$  относительно базиса  $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}$ . Её невырожденность эквивалентна линейной независимости этих векторов, что, в свою очередь, эквивалентно тому, что точки  $B_0, \dots, B_n$  не лежат в одной гиперплоскости.  $\square$

## Задачи

**1.1.** Матрица  $A$  порядка  $n$  такова, что для любой матрицы  $X$  порядка  $n$  с нулевым следом  $\text{tr } AX = 0$ . Докажите, что  $A = \lambda I$ .

**1.2.** Пусть  $A$  и  $B$  — матрицы размера  $m \times n$  и  $k \times n$  соответственно, причём если  $AX = 0$  для некоторого столбца  $X$ , то  $BX = 0$ . Докажите, что  $B = CA$ , где  $C$  — матрица размера  $k \times m$ .

**1.3.** Все координаты вектора  $v \in \mathbb{R}^n$  ненулевые. Докажите, что ортогональное дополнение  $v$  содержит векторы всех ортантов, кроме ортантов, содержащих  $v$  и  $-v$ .



**1.4.** Через вершины треугольника  $ABC$  проведены попарно не параллельные прямые  $a$ ,  $b$  и  $c$ , задающиеся ко векторами  $a^\perp$ ,  $b^\perp$  и  $c^\perp$ . Эти ко векторы определены с точностью до пропорциональности, поэтому будем считать, что  $a^\perp + b^\perp + c^\perp = 0$ . Докажите, что прямые  $a$ ,  $b$  и  $c$  пересекаются в одной точке тогда и только тогда, когда для любой точки  $O$

$$a^\perp(\overrightarrow{OA}) + b^\perp(\overrightarrow{OB}) + c^\perp(\overrightarrow{OC}) = 0.$$

**1.5.** Пусть изоморфизм  $V \rightarrow V^*$  ( $x \mapsto x^*$ ) обладает тем свойством, что равенства  $x^*(y) = 0$  и  $y^*(x) = 0$  эквивалентны. Докажите, что  $x^*(y) = B(x, y)$ , где  $B$  — симметрическая или кососимметрическая билинейная функция.

### Барицентрические координаты

**1.6.** Точка  $P$  имеет барицентрические координаты  $(x_0 : x_1 : \dots : x_n)$  относительно точек  $A_0, A_1, \dots, A_n$ . Найдите барицентрические координаты точки пересечения прямой  $PA_0$  с гиперплоскостью  $A_1 \dots A_n$ .

**1.7.** Точки  $A$  и  $B$  имеют абсолютные барицентрические координаты  $(a_0, \dots, a_n)$  и  $(b_0, \dots, b_n)$ . Докажите, что любая точка прямой  $AB$  имеет абсолютные барицентрические координаты вида  $(\lambda a_0 + (1 - \lambda)b_0, \dots, \lambda a_n + (1 - \lambda)b_n)$ .

**1.8.** Докажите, что бесконечно удалённая точка имеет барицентрические координаты, сумма которых равна 0.

**1.9.** Точки  $X, B_0, \dots, B_n$  имеют барицентрические координаты  $(x_0 : \dots : x_n)$ ,  $(y_0^0 : \dots : y_n^0)$ ,  $\dots$ ,  $(y_0^n : \dots : y_n^n)$  относительно точек  $A_0, \dots, A_n$ . Найдите барицентрические координаты точки  $X$  относительно точек  $B_0, \dots, B_n$ .

**1.10.** Точки  $P_0, \dots, P_n$  имеют абсолютные барицентрические координаты  $(x_0^0, \dots, x_n^0), \dots, (x_0^n, \dots, x_n^n)$  относительно точек  $A_0, \dots, A_n$ . Докажите, что отношение ориентированных объёмов симплексов  $P_0 \dots P_n$  и  $A_0 \dots A_n$  равно

$$\frac{\begin{vmatrix} x_0^0 & \dots & x_n^0 \\ \dots & \dots & \dots \\ x_0^n & \dots & x_n^n \end{vmatrix}}{\begin{vmatrix} x_0^0 & \dots & x_n^0 \\ \dots & \dots & \dots \\ x_0^n & \dots & x_n^n \end{vmatrix}}.$$

**1.11.** Через вершины симплекса  $A_0 \dots A_n$ , имеющего объём  $V$ , проведены параллельные прямые, пересекающие гиперплоскости граней в точках  $B_0, \dots, B_n$ .

а) Докажите, что объём симплекса  $B_0 \dots B_n$  равен  $nV$ .

б) Докажите, что объём симплекса  $A_0 A_1 \dots A_{r-1} B_r B_{r+1} \dots B_n$ , где  $0 \leq r \leq n+1$ , равен  $|n-r|V$ .

## 2.2. Ядро и образ оператора. Факторпространство

### 2.2.1. Ядро и образ оператора

Для линейного отображения  $A : V \rightarrow W$  можно рассмотреть два множества:

$\text{Ker } A = \{v \in V \mid Av = 0\}$  — *ядро* отображения;

$\text{Im } A = \{w \in W \mid \exists v : Av = w\}$  — *образ* отображения.

Легко проверить, что  $\text{Ker } A$  — линейное подпространство в  $V$ , а  $\text{Im } A$  — подпространство в  $W$ . Пусть  $e_1, \dots, e_k$  — базис пространства  $\text{Ker } A$ , а  $e_1, \dots, e_k, e_{k+1}, \dots, e_n$  — расширение этого базиса до базиса пространства  $V$ . Тогда  $Ae_{k+1}, \dots, Ae_n$  — базис пространства  $\text{Im } A$ , поэтому

$$\dim \text{Ker } A + \dim \text{Im } A = \dim V.$$

Выберем в подпространствах  $V$  и  $W$  базисы и рассмотрим матрицу оператора  $A$  относительно этих базисов. Пространство  $\text{Im } A$  порождено столбцами матрицы  $A$ , поэтому  $\dim \text{Im } A = \text{rk } A$ . Видно, в частности, что ранг матрицы оператора  $A$  не зависит от выбора базисов, т.е. ранг оператора определён однозначно. Если даны отображения  $A : U \rightarrow V$  и  $B : V \rightarrow W$ , то  $\text{Im } A$  и  $\text{Ker } B$  могут пересекаться. Для вычисления размерности пересечения этих подпространств можно воспользоваться следующей формулой.

**Теорема 2.2.1.1.** *Имеют место равенства  $\dim(\text{Im } A \cap \text{Ker } B) = \dim \text{Im } A - \dim \text{Im } BA = \dim \text{Ker } BA - \dim \text{Ker } A$ .*

**Доказательство.** Пусть  $C$  — ограничение отображения  $B$  на  $\text{Im } A$ . Тогда  $\dim \text{Ker } C + \dim \text{Im } C = \dim \text{Im } A$ , т.е.  $\dim(\text{Im } A \cap \text{Ker } B) + \dim \text{Im } BA = \dim \text{Im } A$ . Для доказательства второго из требуемых равенств достаточно заметить, что  $\dim \text{Im } BA = \dim V - \dim \text{Ker } BA$  и  $\dim \text{Im } A = \dim V - \dim \text{Ker } A$ .  $\square$

### 2.2.2. Альтернатива Фредгольма

**Теорема 2.2.2.1.** *Ядро и образ оператора  $A$  и сопряжённого к нему оператора  $A^*$  связаны следующими соотношениями:  $\text{Ker } A^* = (\text{Im } A)^\perp$  и  $\text{Im } A^* = (\text{Ker } A)^\perp$ .*

**Доказательство.** Равенство  $A^*f = 0$  означает, что  $f(Ax) = A^*f(x) = 0$  для любого  $x \in V$ , т.е.  $f \in (\operatorname{Im} A)^\perp$ . Поэтому  $\operatorname{Ker} A^* = (\operatorname{Im} A)^\perp$ , а так как  $(A^*)^* = A$ , то  $\operatorname{Ker} A = (\operatorname{Im} A^*)^\perp$ . Следовательно,  $(\operatorname{Ker} A)^\perp = ((\operatorname{Im} A^*)^\perp)^\perp = \operatorname{Im} A^*$ .  $\square$

**Следствие.**  $\operatorname{rk} A = \operatorname{rk} A^*$ .

**Доказательство.**  $\operatorname{rk} A^* = \dim \operatorname{Im} A^* = \dim(\operatorname{Ker} A)^\perp = \dim V - \dim \operatorname{Ker} A = \dim \operatorname{Im} A = \operatorname{rk} A$ .  $\square$

**Замечание 1.** Из равенства  $\operatorname{rk} A = \operatorname{rk} A^*$  легко получить другое доказательство того, что ранг матрицы по строкам равен рангу по столбцам (см. §1.2.2).

**Замечание 2.** Если  $V$  — пространство со скалярным произведением, то  $V^*$  можно отождествить с  $V$  и тогда  $V = \operatorname{Im} A \oplus (\operatorname{Im} A)^\perp = \operatorname{Im} A \oplus \operatorname{Ker} A^*$ . Аналогично  $V = \operatorname{Im} A^* \oplus \operatorname{Ker} A$ . (В матричной записи  $A^* = A^T$ .)

**Теорема 2.2.2.2 (альтернатива Фредгольма).** Пусть  $A : V \rightarrow V$  — линейный оператор. Рассмотрим 4 уравнения:

$$\begin{array}{ll} (1) & Ax = y, \quad x, y \in V; \\ (2) & A^*f = g, \quad f, g \in V^*; \end{array} \quad \begin{array}{ll} (3) & Ax = 0; \\ (4) & A^*f = 0. \end{array}$$

Тогда либо уравнения (1) и (2) разрешимы при любых правых частях, причём в этом случае решение единственно, либо уравнения (3) и (4) имеют одинаковое количество линейно независимых решений  $x_1, \dots, x_k$  и  $f_1, \dots, f_k$ , причём в этом случае уравнение (1) (соответственно (2)) разрешимо тогда и только тогда, когда  $f_1(y) = \dots = f_k(y) = 0$  (соответственно  $g(x_1) = \dots = g(x_k) = 0$ ).

**Доказательство.** Альтернатива Фредгольма является, по сути дела, переформулировкой теоремы 2.2.2.1. Разрешимость уравнений (1) и (2) при любых правых частях означает, что  $\operatorname{Im} A = V$  и  $\operatorname{Im} A^* = V$ , т.е.  $(\operatorname{Ker} A^*)^\perp = V$  и  $(\operatorname{Ker} A)^\perp = V$ , а значит,  $\operatorname{Ker} A^* = 0$  и  $\operatorname{Ker} A = 0$ . Эти равенства эквивалентны, так как  $\operatorname{rk} A = \operatorname{rk} A^*$ .

Если же  $\operatorname{Ker} A \neq 0$ , то  $\dim \operatorname{Ker} A^* = \dim \operatorname{Ker} A$  и  $y \in \operatorname{Im} A$  тогда и только тогда, когда  $y \in (\operatorname{Ker} A^*)^\perp$ , т.е.  $f_1(y) = \dots = f_k(y) = 0$ . Аналогично  $g \in \operatorname{Im} A^*$  тогда и только тогда, когда  $g(x_1) = \dots = g(x_k) = 0$ .  $\square$

### 2.2.3. Теорема Кронекера—Капелли

Образ линейного отображения  $A$  связан с разрешимостью линейного уравнения

$$Ax = b; \quad (1)$$

это уравнение разрешимо тогда и только тогда, когда  $b \in \text{Im } A$ . В том случае, когда отображение задано матрицей, есть простой критерий разрешимости уравнения (1).

**Теорема 2.2.3.1 (Кронекер—Капелли).** Пусть  $A$  — матрица,  $x$  и  $b$  — столбцы, причём их размеры, таковы, что (1) имеет смысл. Уравнение (1) разрешимо тогда и только тогда, когда  $\text{rk } A = \text{rk}(A, b)$ , где  $(A, b)$  — матрица, полученная из матрицы  $A$  приписыванием столбца  $b$ .

**Доказательство.** Пусть  $A_1, \dots, A_n$  — столбцы матрицы  $A$ . Уравнение (1) можно переписать в виде  $x_1 A_1 + \dots + x_n A_n = b$ . Это равенство означает, что столбец  $b$  является линейной комбинацией столбцов  $A_1, \dots, A_n$ , т.е.  $\text{rk } A = \text{rk}(A, b)$ .  $\square$

Линейные отображения можно задавать разными способами; например, выражение  $f(X) = A_1 X B_1 + \dots + A_n X B_n$ , где размеры матриц  $X$ ,  $A_i$  и  $B_i$  таковы, что оно имеет смысл, задаёт линейное отображение одного пространства матриц в другое. Мы ограничимся исследованием уравнения

$$C = AXB. \quad (2)$$

Приведем сначала это уравнение к более простому виду.

**Теорема 2.2.3.2.** Пусть  $a = \text{rk } A$ . Тогда существуют такие невырожденные матрицы  $L_A$  и  $R_A$ , что  $L_A A R_A = I_a$  — единичная матрица порядка  $a$ , дополненная нулями до размеров матрицы  $A$ .

**Доказательство.** Рассмотрим отображение  $A : V^n \rightarrow V^m$ , соответствующее матрице  $A$  (в пространствах  $V^n$  и  $V^m$  фиксированы базисы  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_m$ ). Пусть  $y_{a+1}, \dots, y_n$  — базис  $\text{Ker } A$ , векторы  $y_1, \dots, y_a$  дополняют этот базис до базиса  $V^n$ . Зададим отображение  $R_A : V^n \rightarrow V^n$  формулой  $R_A(e_i) = y_i$ . Тогда  $A R_A(e_i) = A y_i$  при  $i \leq a$  и  $A R_A(e_i) = 0$  при  $i > a$ . Векторы  $x_1 = A y_1, \dots, x_a = A y_a$  образуют базис пространства  $\text{Im } A$ ; дополним их векторами  $x_{a+1}, \dots, x_m$  до базиса пространства  $V^m$ . Зададим отображение  $L_A : V^m \rightarrow V^m$  формулой  $L_A x_i = \varepsilon_i$ . Тогда  $L_A A R_A(e_i) = \varepsilon_i$  при  $1 \leq i \leq a$  и  $L_A A R_A(e_i) = 0$  при  $i > a$ . Поэтому матрицы операторов  $L_A$  и  $R_A$  относительно базисов  $e$  и  $\varepsilon$  соответственно являются искомыми.  $\square$

**Теорема 2.2.3.3.** Уравнение (2) разрешимо тогда и только тогда, когда выполнено одно из следующих эквивалентных условий:

- а) существуют такие матрицы  $Y$  и  $Z$ , что  $AY = ZB$ ;  
 б)  $\operatorname{rk} A = \operatorname{rk}(A, C)$  и  $\operatorname{rk} \begin{pmatrix} B \\ C \end{pmatrix} = \operatorname{rk} \begin{pmatrix} A & C \end{pmatrix}$ , где матрица  $(A, C)$  составлена из столбцов матриц  $A$  и  $C$ , а другая матрица составлена из строк матриц  $B$  и  $C$ .

**Доказательство.** Эквивалентность условий а) и б) доказывается аналогично теореме 2.2.3.1. Ясно также, что если  $C = AXB$ , то можно положить  $Y = XB$  и  $Z = AX$ . Предположим теперь, что  $C = AY$  и  $Z = AX$ . Используя теорему 2.2.3.2, уравнение (2) можно переписать в виде  $D = I_a W I_b$ , где  $D = L_A C R_B$  и  $W = R_A^{-1} X L_B^{-1}$ . Условия  $C = AY$  и  $C = ZB$  запишутся при этом в виде  $D = I_a (R_A^{-1} Y R_B)$  и  $D = (L_A Z L_B^{-1}) I_b$ . Первое из этих равенств означает, что последние  $m$  — строк матрицы  $D$  нулевые, а второе равенство означает, что последние  $n - b$  столбцов матрицы  $D$  нулевые. Поэтому в качестве  $W$  можно взять матрицу  $D$ ; более того, в матрице  $D$  элементы последних  $m$  — строк и последних  $n - b$  столбцов можно заменить любыми числами, т.е. пространство решений имеет размерность  $mn - ab$ . По матрице  $W$  матрица  $X$  определяется однозначно:  $X = R_A W L_B$ .  $\square$

## 2.2.4. Факторпространство

Если  $W$  — подпространство в  $V$ , то  $V$  можно разбить на подмножества  $M_v = \{x \in V \mid v \in W\}$ . Легко проверить, что  $M_v = M_{v'}$  тогда и только тогда, когда  $v - v' \in W$ . На множестве  $V/W = \{M_v \mid v \in V\}$  можно ввести структуру линейного пространства, полагая  $\lambda M_v = M_{\lambda v}$  и  $M_v + M_{v'} = M_{v+v'}$ ; легко проверить, что  $M_{\lambda v}$  и  $M_{v+v'}$  не зависят от выбора  $v$  и  $v'$ , а зависят лишь от самих множеств  $M_v$  и  $M_{v'}$ . Пространство  $V/W$  называется *факторпространством* пространства  $V$  по подпространству  $W$ ; класс  $M_v$  удобно обозначать  $v + W$ .

Существует каноническое отображение  $p: V \rightarrow V/W$ , где  $p(v) = M_v$ ; отображение  $p$  называют *проекцией*. Ясно, что  $\operatorname{Ker} p = W$  и  $\operatorname{Im} p = V/W$ . Если  $e_1, \dots, e_k$  — базис  $W$  и  $e_1, \dots, e_k, e_{k+1}, \dots, e_n$  — базис  $V$ , то  $p(e_1) = \dots = p(e_k) = 0$ , а  $(e_{k+1}), \dots, (e_n)$  — базис  $V/W$ . Поэтому  $\dim(V/W) = \dim V - \dim W$ .

**Теорема 2.2.4.1.** Имеют место следующие канонические изоморфизмы:

- а)  $(U/W)/(V/W) \cong U/V$ , если  $W \subset V \subset U$ ;  
 б)  $V/V \cap W \cong (V + W)/W$ , если  $V, W \subset U$ .

**Доказательство.** Пусть  $u_1, u_2 \in U$ . Классы  $u_1 + W$  и  $u_2 + W$  задают один класс по модулю  $V/W$  тогда и только тогда, когда  $[(u_1 + W) - (u_2 + W)] \in W$ , т.е.  $u_1 - u_2 \in V + W = V$ , а значит, элементы  $u_1$  и  $u_2$  задают один класс по модулю  $V$ .

б) Элементы  $v_1, v_2 \in V$  задают один класс по модулю  $V \cap W$  тогда и только тогда, когда  $v_1 - v_2 \in W$ , а значит, классы  $v_1 + W$  и  $v_2 + W$  совпадают.  $\square$

### 2.2.5. Точные последовательности

Последовательность отображений  $\dots \rightarrow U \xrightarrow{A} V \xrightarrow{B} W \rightarrow \dots$  называют *точной* в члене  $V$ , если  $\text{Im } A = \text{Ker } B$ .

Точность последовательности  $U \xrightarrow{A} V \rightarrow 0$  означает, что  $A$  — *эпиморфизм* (отображение на всё пространство  $V$ ), а точность последовательности  $0 \rightarrow U \xrightarrow{A} V$  означает, что  $A$  — *мономорфизм* (отображение с нулевым ядром).

Для ядра отображения  $A : U \rightarrow V$  мы имеем точную последовательность  $0 \rightarrow \text{Ker } A \rightarrow U \xrightarrow{A} V$ . *Коядром* отображения  $A : U \rightarrow V$  называют пространство  $\text{Coker } A = V / \text{Im } A$ , для которого точна последовательность  $U \xrightarrow{A} V \rightarrow \text{Coker } A \rightarrow 0$ . Соединяя эти две точные последовательности, мы получаем точную последовательность

$$0 \rightarrow \text{Ker } A \rightarrow U \xrightarrow{A} V \rightarrow \text{Coker } A \rightarrow 0.$$

**Теорема 2.2.5.1.** При переходе к двойственным пространствам и двойственным отображениям точность последовательностей сохраняется.

**Доказательство.** Рассмотрим последовательность отображений  $U \xrightarrow{A} V \xrightarrow{B} W$  и двойственную ей последовательность отображений  $U^* \xleftarrow{A^*} V^* \xleftarrow{B^*} W^*$ . Требуется доказать, что если  $\text{Ker } B = \text{Im } A$ , то  $\text{Im } B^* = \text{Ker } A^*$ . По определению

$$\begin{aligned} \text{Ker } A^* &= \{f \in V^* \mid f(Au) = 0 \quad \forall u \in U\}, \\ \text{Im } B^* &= \{f \in V^* \mid \exists g \in W^* : f(v) = g(Bv) \quad \forall v \in V\}. \end{aligned}$$

Легко видеть, что  $\text{Im } B^* \subset \text{Ker } A^*$ . Действительно, если  $f(v) = g(Bv)$ , то  $f(Au) = g(BAu) = 0$ , поскольку  $BA = 0$ . Далее, согласно теореме 2.2.2.1  $\dim \text{Ker } A^* = \dim V - \dim \text{Im } A = \dim V - \dim \text{Ker } B = \dim \text{Im } B^*$ . Поэтому  $\text{Im } B^* = \text{Ker } A^*$ .  $\square$

При переходе к двойственным пространствам точная последовательность

$$0 \rightarrow \operatorname{Ker} A \rightarrow U \xrightarrow{A} V \rightarrow \operatorname{Coker} A \rightarrow 0$$

переходит в точную последовательность

$$0 \leftarrow (\operatorname{Ker} A)^* \leftarrow U^* \xleftarrow{A^*} V^* \leftarrow (\operatorname{Coker} A)^* \leftarrow 0,$$

поэтому ядро и коядро являются двойственными понятиями.

### Задачи

2.1. Пусть  $A$  — вещественная матрица. Докажите, что  $\operatorname{rk} A^T A = \operatorname{rk} A$ .

2.2. Пусть  $A$  — линейный оператор. Докажите, что

$$\dim \operatorname{Ker} A^{n+1} = \dim \operatorname{Ker} A + \sum_{k=1}^n \dim (\operatorname{Im} A^k \cap \operatorname{Ker} A)$$

и

$$\dim \operatorname{Im} A = \dim \operatorname{Im} A^{n+1} + \sum_{k=1}^n \dim (\operatorname{Im} A^k \cap \operatorname{Ker} A).$$

2.3. Докажите, что линейную функцию  $f$  можно представить в виде линейной комбинации линейных функций  $f_1, \dots, f_k$  тогда и только тогда, когда  $\bigcap_{i=1}^k \operatorname{Ker} f_i \subset \operatorname{Ker} f$ .

2.4. Пусть  $A : U \rightarrow V$  и  $B : V \rightarrow W$  — линейные отображения, для которых  $BA = 0$ . Тогда определено факторпространство  $\operatorname{Ker} B / \operatorname{Im} A$ . Докажите, что  $\operatorname{Ker} A^* / \operatorname{Im} B^*$  — двойственное пространство для  $\operatorname{Ker} B / \operatorname{Im} A$ .

## 2.3. Базисы. Линейная независимость

### 2.3.1. Характеристический многочлен оператора

Различным объектам — линейным отображениям, квадратичным формам, кососимметрическим формам — можно сопоставить матрицу, выбрав предварительно базис. При переходе к другому базису эта матрица может измениться. Для наглядности мы, забегаая несколько вперед, соберём вместе законы преобразований этих матриц при замене базиса.

Рассмотрим линейное отображение  $f : V \rightarrow W$ . Пусть  $x$  — столбец  $(x_1, \dots, x_n)$ ,  $e$  — строка  $(e_1, \dots, e_n)$  и  $\varepsilon$  — строка  $(\varepsilon_1, \dots, \varepsilon_m)$ , причём строки  $e$  и  $\varepsilon$  составлены из базисных векторов. Тогда  $A$  будет матрицей отображения  $f$  относительно базисов  $e$  и  $\varepsilon$ , если  $f(ex) = \varepsilon Ax$ . Пусть

$e' = eP$  и  $\varepsilon' = \varepsilon Q$  — другие базисы. Тогда  $f(e'x) = f(ePx) = \varepsilon APx = \varepsilon' Q^{-1} APx$ , поэтому  $A' = Q^{-1} AP$  — матрица отображения  $f$  относительно базисов  $e'$  и  $\varepsilon'$ . Особенно важен тот случай, когда  $V = W$ , т.е.  $f$  — линейный оператор, и  $Q = P$ . В этом случае  $A' = P^{-1} AP$ .

**Теорема 2.3.1.1.** Пусть  $A$  — матрица линейного оператора. Многочлен  $|A - \lambda I| = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0$  не зависит от выбора базиса, а зависит лишь от самого оператора.

**Доказательство.**  $|P^{-1} AP - \lambda I| = |P^{-1}(A - \lambda I)P| = |P^{-1}| \cdot |P| \cdot |(A - \lambda I)| = |A - \lambda I|$ .  $\square$

В частности, инвариантами оператора являются  $|A| = a_0$  и  $\text{tr } A = (-1)^{n-1} a_{n-1}$ . Многочлен  $|A - \lambda I|$  называют *характеристическим многочленом* оператора, а его корни — *характеристическими* или *собственными числами (значениями)* оператора.

Пусть  $e_1, \dots, e_n$  — базис  $V$ ;  $e_1^*, \dots, e_n^*$  — двойственный базис;  $e$  — строка  $(e_1, \dots, e_n)$ ;  $e^*$  — столбец  $(e_1^*, \dots, e_n^*)$ . Тогда  $e^* e = I_n$  — единичная матрица порядка  $n$ . Вектору  $v$  можно сопоставить столбец  $(x_1, \dots, x_n)$ , а ковектору  $f$  — строку  $y = (y_1, \dots, y_n)$  так, что  $v = ex$  и  $f = ye^*$ . Столбец  $x$  и строку  $y$  называют *координатами*  $v$  и  $f$  относительно базиса  $e$ . Пусть  $x'$  и  $y'$  — координаты  $v$  и  $f$  относительно базиса  $e' = eP$ . Тогда  $ex = v = e'x' = ePx'$ , поэтому  $x = Px'$ . Далее,  $ye^* = f = y'(e')^*$ , поэтому  $y' = (e')^* y = e^* y = ye^* = yP$ . Итак, если  $x = Px'$ , то  $y' = yP$  и  $y = Px'$ , т.е. координаты ковектора преобразуются точно так же, как преобразуются базис (*ковариантность*), а координаты вектора преобразуются не так (*контравариантность*).

Рассмотрим теперь квадратичную форму  $B(x) = x^T A x$ . При переходе к базису  $e' = eP$  получим  $x = Px'$ , поэтому  $(x')^T A' x' = x^T A x = (x')^T P^T A P x$ , а значит,  $A' = P^T A P$ . Видно, в частности, что форму  $x_1^2 + \dots + x_n^2$  сохраняют преобразования, для которых  $I = P^T I P = P^T P$ , т.е. ортогональные преобразования.

### 2.3.2. Ориентация

Два базиса  $e$  и  $e'$  в линейном пространстве над  $\mathbb{R}$  называют *одинаково ориентированными* если  $e' = eP$ , где  $\det P > 0$ . В противном случае базисы  $e$  и  $e'$  называют *противоположно ориентированными*. Таким образом, все базисы в линейном пространстве над  $\mathbb{R}$  разбиваются на два класса эквивалентности. Базисы из одного класса одинаково ориентированы,



а базисы из разных классов противоположно ориентированы. Заданием *ориентации* в пространстве называют выбор одного из этих классов эквивалентности.

Если поменять местами два вектора в базисе, то ориентация при этом меняется. Действительно, определитель матрицы перехода в этом случае равен  $(-1)^\sigma$ , где  $\sigma$  — транспозиция. Если же происходит некоторая перестановка векторов базиса, то для чётной перестановки ориентация сохраняется, а для нечётной изменяется.

Пусть  $V$  — линейное пространство над  $\mathbb{R}$ . Говорят, что линейный оператор  $A : V \rightarrow V$  *сохраняет ориентацию*, если  $\det A > 0$ . Если же  $\det A < 0$ , то говорят, что оператор  $A$  *изменяет ориентацию*. Оператор, сохраняющий ориентацию, переводит каждый базис в одинаково ориентированный с ним базис.

### 2.3.3. Решётки

Пусть  $V$  —  $n$ -мерное пространство над  $\mathbb{R}$ . Множество  $\Gamma \subset V$  называют *решёткой*, если оно состоит из векторов  $m_1 e_1 + \dots + m_n e_n$ , где  $m_1, \dots, m_n \in \mathbb{Z}$ , а  $e_1, \dots, e_n$  — некоторый фиксированный базис.

**Теорема 2.3.3.1.** *Подгруппа (относительно сложения)  $\Gamma \subset V$  является решёткой тогда и только тогда, когда она дискретна (т.е. в некоторой окрестности нуля нет других элементов  $\Gamma$ ) и не содержится ни в каком подпространстве размерности  $n - 1$ .*

**Доказательство.** Ясно, что решётка дискретна и не содержится в подпространстве размерности  $n - 1$ . Докажем индукцией по  $n$ , что если подгруппа  $\Gamma$  дискретна и не содержится в подпространстве размерности  $n - 1$ , то можно выбрать базис  $e_1, \dots, e_n$  требуемым образом. База индукции  $n = 0$  очевидна. Докажем шаг индукции. По условию  $\Gamma$  не содержится в подпространстве размерности  $n - 1$ , поэтому можно выбрать линейно независимые векторы  $v_1, \dots, v_n \in \Gamma$ . Пусть  $V_1 = \langle v_1, \dots, v_{n-1} \rangle$  и  $\Gamma_1 = V_1 \cap \Gamma$ . К  $\Gamma_1$  можно применить предположение индукции и выбрать линейно независимые векторы  $e_1, \dots, e_{n-1}$  так, что  $\Gamma_1$  состоит из векторов  $m_1 e_1 + \dots + m_{n-1} e_{n-1}$ , где  $m_1, \dots, m_{n-1} \in \mathbb{Z}$ .

Возьмём базис  $e_1, \dots, e_{n-1}, v_n$  и рассмотрим параллелепипед

$$P = \{t_1 e_1 + \dots + t_{n-1} e_{n-1} + t_n v_n \mid 0 \leq t_i \leq 1\}.$$

Параллелепипед  $P$  компактен, поэтому множество  $P \cap \Gamma$  состоит из конечного числа векторов. Следовательно, можно выбрать  $e_n \in (P \cap \Gamma) \setminus V_1$  так, что координата  $t_n$  в разложении  $e_n = t_1 e_1 + \dots + t_{n-1} e_{n-1} + t_n v_n$

минимальна. Покажем, что  $e_1, \dots, e_n$  — требуемый базис. Пусть  $v \in \Gamma$ . Выберем целые числа  $m_1, \dots, m_n$  так, что  $v - \sum_{j=1}^n m_j e_j = \sum_{j=1}^{n-1} s_j e_j + s v_n$ , где  $0 \leq s_j < 1$  и  $0 \leq s < t_n$ . Вектор  $v - \sum_{j=1}^n m_j e_j$  принадлежит  $P \cap \Gamma$ , поэтому из минимальности  $t_n$  следует, что  $s = 0$ . Значит,  $v - \sum_{j=1}^n m_j e_j \in \Gamma \cap V_1 = \Gamma_1$ , поэтому  $s_j = 0$  для всех  $j = 1, \dots, n-1$ . Таким образом,  $v = \sum_{j=1}^n m_j e_j$ , что и требовалось.  $\square$

### 2.3.4. Свободные абелевы группы

*Абелева группа*  $A$  — это множество, в котором для любых двух элементов  $a$  и  $b$  определена их сумма  $a + b$ ; эта операция обладает следующими свойствами:

- $a + b = b + a$  (коммутативность);
- $(a + b) + c = a + (b + c)$  (ассоциативность);
- в  $A$  существует нулевой элемент  $0$ , для которого  $a + 0 = a$  для любого  $a \in A$ ;
- для любого  $a \in A$  существует противоположный элемент  $-a \in A$ , для которого  $a + (-a) = 0$ .

Для каждого натурального числа  $n$  можно определить элемент  $na = \underbrace{a + \dots + a}_n$ . Если положить  $(-n)a = n(-a)$ , то это определение переносится на все целые числа  $n$ .

Примерами абелевых групп служат группа целых чисел  $\mathbb{Z}$  и группа  $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$  (группа вычетов по модулю  $n$ ).

Для элементов  $a_1, \dots, a_k \in A$  можно рассмотреть подгруппу  $\langle a_1, \dots, a_k \rangle$ , состоящую из элементов вида  $n_1 a_1 + \dots + n_k a_k$ , где  $n_1, \dots, n_k$  пробегает все целые числа. Абелеву группу  $A$  называют *конечно порождённой*, если  $A = \langle a_1, \dots, a_k \rangle$  для некоторого натурального числа  $k$ .

Элементы  $a_1, \dots, a_k \in A$  называют *линейно независимыми*, если равенство  $n_1 a_1 + \dots + n_k a_k = 0$  выполняется только при  $n_1 = \dots = n_k = 0$ . *Базисом* абелевой группы  $A$  называют набор линейно независимых элементов  $a_1, \dots, a_k$ , для которого  $A = \langle a_1, \dots, a_k \rangle$ .

Не любая конечно порождённая абелева группа обладает базисом. Например, группа  $\mathbb{Z}_n$  порождена одним элементом, но для любого элемента  $a \in \mathbb{Z}_n$  выполняется равенство  $na = 0$ .

Конечно порождённую абелеву группу, обладающую базисом, называют *свободной*.

Все базисы свободной абелевой группы имеют одно и то же число элементов. Это доказывается точно так же, как аналогичное утверждение для конечномерных векторных пространств. Число элементов базиса свободной абелевой группы называют её *рангом*.

**Теорема 2.3.4.1.** *Любая подгруппа свободной абелевой группы ранга  $n$  является свободной абелевой группой ранга не более  $n$ .*

**Первое доказательство.** отождествим свободную абелеву группу ранга  $n$  с решёткой векторов в  $\mathbb{R}^n$ , имеющих целые координаты. Векторы, принадлежащие подгруппе, порождают подпространство в  $\mathbb{R}^n$ ; это подпространство имеет размерность  $m \leq n$ . Рассматриваемая подгруппа является решёткой в этом пространстве, поэтому она является свободной абелевой группой ранга  $m$ .  $\square$

**Второе доказательство.** Применим индукцию по  $n$ . При  $n = 0$  утверждение очевидно. Пусть  $n > 0$ ,  $a_1, \dots, a_n$  — базис абелевой группы  $A$  и  $H \subset A$  — данная подгруппа. Рассмотрим подгруппу  $A_{n-1} = \langle a_1, \dots, a_{n-1} \rangle$ . По предположению индукции группа  $H_1 = A_1 \cap H$  является свободной абелевой группой ранга  $m \leq n - 1$ . Пусть  $b_1, \dots, b_m$  — её базис.

Рассмотрим последние координаты всех элементов  $H$  относительно базиса  $a_1, \dots, a_n$  группы  $A$ . Они образуют подгруппу в  $\mathbb{Z}$ ; эта подгруппа имеет вид  $k\mathbb{Z}$ , где  $k$  — некоторое натуральное число. Если  $k = 0$ , то  $H = H_1$ . Если  $k \neq 0$ , то выберем элемент  $b_{m+1} \in H$ , у которого последняя координата равна  $k$ . Тогда  $b_1, \dots, b_{m+1}$  — базис  $H$ .  $\square$

### 2.3.5. Две теоремы об обмене векторов базисов

Большинство общих утверждений о базисах достаточно очевидно. Есть, однако, несколько неочевидных теорем о получении базисов путем обмена векторов двух систем линейно независимых векторов.

**Теорема 2.3.5.1.** *Пусть  $x_1, \dots, x_n$  и  $y_1, \dots, y_n$  — два базиса и  $1 \leq k \leq n$ . Тогда  $k$  из векторов  $y_1, \dots, y_n$  можно поменять с векторами  $x_1, \dots, x_k$  так, что снова получим два базиса.*

**Доказательство.** [Gr1] Векторы  $y_1, \dots, y_n$  возьмём в качестве базиса пространства  $V$ . Для любого набора  $n$  векторов  $z_1, \dots, z_n$  пространства  $V$  рассмотрим величину  $M(z_1, \dots, z_n)$  — определитель матрицы, строки которой являются координатами векторов  $z_1, \dots, z_n$  относительно базиса  $y_1, \dots, y_n$ . Векторы  $z_1, \dots, z_n$  образуют базис тогда и только тогда,

когда  $M(z_1, \dots, z_n) \neq 0$ . Формулу разложения определителя матрицы  $M(x_1, \dots, x_n)$  по первым  $k$  строкам можно записать в виде

$$M(x_1, \dots, x_n) = \sum_{A \subset Y} \pm M(x_1, \dots, x_k, A) M(Y \setminus A, x_{k+1}, \dots, x_n), \quad (2.1)$$

где суммирование производится по всем  $(n-k)$ -элементным подмножествам множества  $Y = \{y_1, \dots, y_n\}$ . Так как  $M(x_1, \dots, x_n) \neq 0$ , то в сумме (2.1) есть хотя бы один ненулевой член; соответствующее подмножество  $A$  определяет требуемый набор векторов базиса  $y_1, \dots, y_n$ .  $\square$

**Теорема 2.3.5.2.** Пусть  $x_1, \dots, x_n$  и  $y_1, \dots, y_n$  — два базиса. Тогда для любого  $k$  из  $x_1, \dots, x_n$  можно выбрать несколько векторов, меняя которые с некоторыми из векторов  $y_1, \dots, y_k$ , снова получим два базиса.

**Доказательство.** [Gr2] Вектор  $y_i$  можно разложить по базису  $x_1, \dots, x_n$ . Пусть  $A_i$  и  $A'_i$  — множества тех векторов из  $x_k, \dots, x_n$  и  $x_1, \dots, x_{k-1}$ , которые входят в это разложение с ненулевыми коэффициентами. Тогда  $y_i \in \langle A_i, A'_i \rangle$ , и если  $x_j \in A_i$ , то  $x_j \in \langle y_i, A_i \setminus x_j, A'_i \rangle$ .

Докажем, что в  $\{y_1, \dots, y_k\}$  существует такое непустое подмножество  $S$ , что

$$y_i \in S \Rightarrow A_i \not\subset \langle \bar{S} \rangle, \quad (2.2)$$

где  $\bar{S} = \{y_1, \dots, y_n\} \setminus S$ . Предположим противное. Тогда множество  $S = \{y_1, \dots, y_k\}$  не обладает свойством (2.2), поэтому найдётся такое  $i \in \{1, 2, \dots, k\}$ , что  $A_i \subset \langle \bar{S} \rangle = \langle y_{k+1}, \dots, y_n \rangle$ . Пусть, для определённости,  $i = 1$ . Множество  $\{y_2, \dots, y_k\}$  тоже не обладает свойством (2.2), поэтому найдётся такое  $i \in \{2, \dots, k\}$ , что  $A_i \subset \langle y_1, y_{k+1}, \dots, y_n \rangle$ ; для определённости положим  $i = 2$ . Продолжая такие рассуждения, дойдём до  $A_k \subset \langle y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_n \rangle$ . Рассмотрим подпространство  $W = \langle x_1, \dots, x_{k-1}, y_{k+1}, \dots, y_n \rangle$ . Все подмножества  $A'_i$  лежат в  $\langle x_1, \dots, x_{k-1} \rangle$ , поэтому они лежат в  $W$ . Так как  $A_1 \subset \langle y_{k+1}, \dots, y_n \rangle$ , то  $y_1 \in \langle A_1, A'_1 \rangle \subset W$ . Поэтому  $A_2 \subset \langle y_1, y_{k+1}, \dots, y_n \rangle \subset W$ , а значит,  $y_2 \in \langle A_2, A'_2 \rangle \subset W$  и т.д. до  $y_k \in W$ . В итоге получаем  $\langle y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_n \rangle \subset W$ . Но  $\dim W \leq (k-1) + (n-k) = n-1$ . Приходим к противоречию.

Рассмотрим минимальное множество  $S$ , обладающее свойством (2.2). Если  $S$  состоит из одного элемента, например  $y_1$ , то  $A_1 \not\subset \langle y_2, \dots, y_n \rangle$ , поэтому найдётся такой  $x_i \in A_1$ , что  $x_i \notin \langle y_2, \dots, y_n \rangle$ , т.е.  $x_i, y_2, \dots, y_n$  — базис. С другой стороны,  $x_i \in \langle y_1, A_1 \setminus x_i, A'_1 \rangle$ , поэтому  $y_1, x_1, \dots, \hat{x}_i, \dots, x_n$  — тоже базис. Предположим теперь, что  $S$  состоит более чем из одного элемента. В силу минимальности  $S$  для всех  $y_i \in S$  множество  $S \setminus y_i$  не обладает свойством (2.2), т.е.

найдётся такой  $y_j \in S \setminus y_i$ , что  $A_j \subset \langle \bar{S}, y_i \rangle$ . В результате каждому элементу  $y_i$  сопоставлен некоторый элемент  $y_j = f(y_i)$ . Это отображение  $f : S \rightarrow S$  — биекция без неподвижных точек. В самом деле, если  $f(y_i) = f(y_{i'}) = y_j$ , то  $A_j \subset \langle \bar{S}, y_i \rangle \cap \langle \bar{S}, y_{i'} \rangle = \langle \bar{S} \rangle$ , что противоречит свойству (2.2). Инъективное отображение конечного множества в себя является биекцией.

Сопоставим каждому  $y_i \in S$  элемент  $z_i \in \{x_k, \dots, x_n\}$  следующим образом. Пусть  $f(y_i) = y_j$ . Так как  $A_j \subset \langle \bar{S}, y_i \rangle$  и  $A_j \not\subset \langle \bar{S} \rangle$ , то найдётся такой  $z_i \in A_j \setminus \langle \bar{S} \rangle$ , что  $\langle \bar{S}, y_i \rangle = \langle \bar{S}, z_i \rangle$ . Если  $y_\alpha, y_\beta \in S$ , то  $A_\alpha \subset \langle \bar{S}, y_p \rangle$  и  $A_\beta \subset \langle \bar{S}, y_q \rangle$ , где  $y_p \neq y_q$ . Следовательно,  $A_\alpha \cap A_\beta \subset \langle \bar{S} \rangle$ , т.е. множества  $A_\alpha \setminus \langle \bar{S} \rangle$  и  $A_\beta \setminus \langle \bar{S} \rangle$  попарно не пересекаются, поэтому  $z_i$  не входит в  $A_j$  при  $f(y_i) \neq y_j$ . Пусть для определённости  $S = \{y_1, \dots, y_p\}$ . Докажем, что  $\{y_1, \dots, y_p\}$  можно поменять с  $\{z_1, \dots, z_p\}$ . В самом деле,  $y_i \in \langle \bar{S}, z_i \rangle = \langle z_i, y_{p+1}, \dots, y_n \rangle$  при  $i \leq p$ , поэтому  $y_1, \dots, y_p \in \langle z_1, \dots, z_p, y_{p+1}, \dots, y_n \rangle$  т.е.  $z_1, \dots, z_p, y_{p+1}, \dots, y_n$  — базис. С другой стороны,  $z_i \in \langle y_j, A_j \setminus z_i, A'_j \rangle$  и в  $A_j$  из всех элементов  $z_1, \dots, z_p$  входит только  $z_i$ . Поэтому  $z_i \in \langle y_j, A \rangle$ , где  $A = \{x_1, \dots, x_n\} \setminus \{z_1, \dots, z_p\}$ , а значит,  $\{y_1, \dots, y_p, A\}$  — базис.  $\square$

### 2.3.6. Разбиение $\mathbb{R}^n$ на многогранные углы

С обменом векторов в двух базисах связана также следующая задача. Пусть  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_n$  — два базиса, причём каждый набор векторов  $v_1, \dots, v_n$ , где  $v_i = e_i$  или  $\varepsilon_i$ , является базисом. Каждому такому базису можно сопоставить конус  $\text{Cone}(v_1, \dots, v_n)$ , состоящий из векторов  $\lambda_1 v_1 + \dots + \lambda_n v_n$ , где  $\lambda_1, \dots, \lambda_n$  — неотрицательные числа. При каких условиях эти  $2^n$  конусов покрывают  $\mathbb{R}^n$  (без перекрытий)?

Простейший пример, когда  $2^n$  конусов покрывают  $\mathbb{R}^n$ , — покрытие ортантами; в этом случае  $e_1, \dots, e_n$  — канонический базис (или произвольный базис), а  $\varepsilon_i = -e_i$  для  $i = 1, \dots, n$ .

**Теорема 2.3.6.1.** [Sa] Пусть для определённости базис  $e_1, \dots, e_n$  положительно ориентирован. Тогда рассматриваемые  $2^n$  конусов покрывают  $\mathbb{R}^n$  тогда и только тогда, когда знак ориентации базиса  $v_1, \dots, v_n$  равен  $(-1)^s$ , где  $s$  — количество векторов  $\varepsilon_i$  в этом базисе.

**Доказательство.** Легко видеть, что это утверждение о знаке ориентации базиса эквивалентно следующему: базисы  $v_1, \dots, e_i, \dots, v_n$  и  $v_1, \dots, \varepsilon_i, \dots, v_n$  всегда противоположно ориентированы (задача 3.6). Если конусы  $\text{Cone}(v_1, \dots, e_i, \dots, v_n)$  и  $\text{Cone}(v_1, \dots, \varepsilon_i, \dots, v_n)$  не перекрываются (т.е. их пересечение имеет размерность строго меньше  $n$ ), то векторы  $e_i$  и  $\varepsilon_i$  лежат по разные стороны от гиперплоскости  $\langle v_1, \dots, \hat{v}_i, \dots, v_n \rangle$ , поэтому

указанное условие является необходимым. Остаётся доказать, что оно является также и достаточным.

Заметим сначала, что если указанное условие выполняется, то рассматриваемые конусы покрывают  $\mathbb{R}^n$ . Действительно, если они не покрывают  $\mathbb{R}^n$ , то дополнение к их объединению — открытое множество, граница которого состоит из частей  $(n-1)$ -мерных граней рассматриваемых конусов. Но грань  $\text{Cone}(v_1, \dots, \hat{v}_i, \dots, v_n)$  не может попасть в дополнение, поскольку она лежит внутри объединения конусов  $\text{Cone}(v_1, \dots, e_i, \dots, v_n)$  и  $\text{Cone}(v_1, \dots, \varepsilon_i, \dots, v_n)$ ; это следует из того, что векторы  $e_i$  и  $\varepsilon_i$  лежат по разные стороны от гиперплоскости  $\langle v_1, \dots, \hat{v}_i, \dots, v_n \rangle$ .

Остаётся доказать, что рассматриваемые конусы не перекрываются. Для этого достаточно доказать, что если среди наборов векторов  $v_1, \dots, v_n$  и  $v'_1, \dots, v'_n$  (второй набор тоже один из рассматриваемых) есть ровно  $r$  общих, то пересечение конусов  $\text{Cone}(v_1, \dots, v_n)$  и  $\text{Cone}(v'_1, \dots, v'_n)$  — это конус, порождённый  $r$  общими векторами (если выполняется указанное условие). Применим индукцию по  $n-r$ . При  $n-r=0$  доказывать нечего. При  $n-r=1$  требуемое следует из того, что векторы  $e_i$  и  $\varepsilon_i$  лежат по разные стороны от гиперплоскости  $\langle v_1, \dots, \hat{v}_i, \dots, v_n \rangle$ . Предположим, что требуемое утверждение доказано для  $n-r=k$ , и рассмотрим два конуса  $L$  и  $M$ , для которых множества порождающих векторов имеют ровно  $r-1 = n-k-1$  общих векторов. Для простоты будем считать, что эти конусы порождены векторами  $e_1, \dots, e_n$  и  $e_1, \dots, e_{r-1}, \varepsilon_r, \dots, \varepsilon_n$ ; в других случаях рассуждения аналогичны. Предположим, что в пересечении этих конусов есть некоторый вектор  $w$ , не принадлежащий  $\text{Cone}(e_1, \dots, e_{r-1})$ . Заметим, что  $w$  не может принадлежать грани  $M$ , противоположной  $\varepsilon_i$ , где  $r \leq i \leq n$ . Действительно, иначе  $w$  принадлежит пересечению конусов  $L$  и  $\text{Cone}(e_1, \dots, e_{r-1}, \varepsilon_r, \dots, e_i, \dots, \varepsilon_n)$ , а по предположению индукции это пересечение имеет вид  $\text{Cone}(e_1, \dots, e_{r-1}, e_i)$ . Таким образом, вектор  $w$  линейно выражается как через  $e_1, \dots, e_{r-1}, e_i$ , так и через  $e_1, \dots, e_{r-1}, \varepsilon_r, \dots, \hat{\varepsilon}_i, \dots, \varepsilon_n$ . По условию векторы  $e_1, \dots, e_{r-1}, \varepsilon_r, \dots, e_i, \dots, \varepsilon_n$  линейно независимы, поэтому  $w \in \text{Cone}(e_1, \dots, e_{r-1})$ . Приходим к противоречию.

Можно считать, что вектор  $w$  лежит внутри  $M$ . Действительно, мы показали, что векторы  $\varepsilon_r, \dots, \varepsilon_n$  входят в  $w$  с ненулевыми коэффициентами. Поэтому, добавив при необходимости к  $w$  вектор из  $\text{Cone}(e_1, \dots, e_{r-1})$ , мы получим вектор, в который все векторы  $e_1, \dots, e_{r-1}, \varepsilon_r, \dots, \varepsilon_n$  входят с ненулевыми коэффициентами.

Представим вектор  $e_r$  в виде  $e_r = e + \varepsilon$ , где  $e \in \langle e_1, \dots, e_{r-1} \rangle$  и  $\varepsilon \in \langle \varepsilon_r, \dots, \varepsilon_n \rangle$ . Выберем в  $\text{Cone}(e_1, \dots, e_{r-1})$  вектор  $v$  так, что  $v + e$  лежит внутри  $\text{Cone}(e_1, \dots, e_{r-1})$ , и рассмотрим семейство векторов  $w_t =$

$w + t(e_r + v)$  для  $t \geq 0$ . Ясно, что  $w_t \in L$  и  $w \notin \text{Cone}(e_1, \dots, e_{r-1})$ . Кроме того, в вектор  $w_t = w + t(v + e) + t\varepsilon$  все векторы  $e_1, \dots, e_{r-1}$  входят с ненулевыми коэффициентами, поэтому он не может принадлежать грани  $M$ , противоположной  $e_i$ , где  $1 \leq i \leq r-1$ .

Вектор  $v$  лежит в  $\langle e_1, \dots, e_{r-1} \rangle$ , поэтому если  $t$  велико, то векторы  $w_t$  и  $\varepsilon_r$  лежат по разные стороны от гиперплоскости  $\langle e_1, \dots, e_{r-1}, \varepsilon_{r+1}, \dots, \varepsilon_n \rangle$  (как и векторы  $e_r$  и  $\varepsilon_r$ ). Значит, если  $t$  велико, то вектор  $w_t$  лежит вне конуса  $M$ . Поэтому можно выбрать  $t_0 > 0$  так, что вектор  $w_{t_0}$  принадлежит грани  $M$ . Как только что было показано, это должна быть грань, противоположная  $\varepsilon_i$ , где  $r \leq i \leq n$ . Но это приводит к противоречию с предположением индукции точно так же, как это делалось для вектора  $w$ .  $\square$

### 2.3.7. Линейная зависимость степеней оператора

**Теорема 2.3.7.1.** Пусть  $T$  — линейный оператор в пространстве  $V^n$ , причём для любого вектора  $\xi \in V$  векторы  $\xi, T\xi, \dots, T^n\xi$  линейно зависимы. Тогда операторы  $I, T, \dots, T^n$  линейно зависимы.

**Доказательство.** [Ау3] Можно считать, что  $n$  — минимальное из всех чисел, обладающих тем свойством, что векторы  $\xi_0, T\xi_0, \dots, T^n\xi_0$  линейно независимы и  $T^n\xi_0 \in \langle \xi_0, \dots, T^{n-1}\xi_0 \rangle$  для некоторого вектора  $\xi_0$ . Тогда существует такой многочлен  $p_0$  степени  $n$ , что  $p_0(T)\xi_0 = 0$ , и можно считать, что коэффициент при старшей степени многочлена  $p_0$  равен 1. Фиксируем вектор  $\eta \in V$  и докажем, что  $p_0(T)\eta = 0$ . Рассмотрим пространство  $W = \langle \xi_0, \dots, T^n\xi_0, \eta, \dots, T^n\eta \rangle$ . Легко проверить, что  $\dim W \leq 2n$  и  $T(W) \subset W$ . Для каждого  $\lambda \in \mathbb{C}$  рассмотрим векторы  $f_0(\lambda) = \xi_0 + \lambda\eta, f_1(\lambda) = Tf_0(\lambda), \dots, f_{n-1}(\lambda) = T^{n-1}f_0(\lambda), g(\lambda) = T^n f_0(\lambda)$ . Векторы  $f_0(0), \dots, f_{n-1}(0)$  линейно независимы, поэтому на  $W$  существуют такие линейные функции  $\varphi_0, \dots, \varphi_{n-1}$  что  $\varphi_i(f_j(0)) = \delta_{ij}$ . Пусть  $\Delta(\lambda) = |a_{ij}(\lambda)|_0^{n-1}$ , где  $a_{ij}(\lambda) = \varphi_i(f_j(\lambda))$ . Тогда  $\Delta(\lambda)$  — многочлен от  $\lambda$  степени не более  $n$ , причём  $\Delta(0) = 1$ . По условию теоремы для любого  $\lambda \in \mathbb{C}$  существуют такие комплексные числа  $a_0(\lambda), \dots, a_{n-1}(\lambda)$ , что

$$g(\lambda) = a_0(\lambda)f_0(\lambda) + \dots + a_{n-1}(\lambda)f_{n-1}(\lambda). \quad (2.1)$$

Следовательно,  $\varphi_i(g(\lambda)) = \sum_{k=0}^{n-1} a_k(\lambda)\varphi_i(f_k(\lambda))$  при  $i = 0, \dots, n-1$ . Если  $\Delta(\lambda) \neq 0$ , то полученную систему линейных уравнений относительно  $a_k(\lambda)$  можно решить с помощью правила Крамера. Поэтому  $a_k(\lambda)$  — рациональная функция для всех  $\lambda \in \mathbb{C} \setminus \Delta$ , где  $\Delta$  — (конечное) множество корней многочлена  $\Delta(\lambda)$

Равенство (2.1) можно записать в виде  $p_\lambda(T)f_0(\lambda) = 0$ , где  $p_\lambda(T) = T^n - \alpha_{n-1}(\lambda)T^{n-1} - \dots - a_0(\lambda)I$ . Пусть  $\beta_1(\lambda), \dots, \beta_n(\lambda)$  — корни многочлена  $p_\lambda$ . Тогда

$$(T - \beta_1(\lambda)I) \dots (T - \beta_n(\lambda)I)f_0(\lambda) = 0.$$

Если  $\lambda \notin \Delta$ , то векторы  $f_0(\lambda), \dots, f_{n-1}(\lambda)$  линейно независимы, т.е.  $h(T)f_0(\lambda) \neq 0$  для любого ненулевого многочлена  $h$  степени  $n - 1$ . Поэтому

$$w = (T - \beta_2(\lambda)I) \dots (T - \beta_n(\lambda)I)f_0(\lambda) \neq 0$$

и  $(T - \beta_1(\lambda)I)w = 0$ , т.е.  $\beta_1(\lambda)$  — собственное значение оператора  $T$ . Аналогично доказывается, что  $\beta_2(\lambda), \dots, \beta_n(\lambda)$  — собственные значения оператора  $T$ . Следовательно,  $|\beta_i(\lambda)| \leq \|T\|$  (см. §6.3.1). Рациональные функции  $\alpha_0(\lambda), \dots, \alpha_{n-1}(\lambda)$  являются элементарными симметрическими функциями от ограниченных на  $\mathbb{C} \setminus \Delta$  функций  $\beta_1(\lambda), \dots, \beta_n(\lambda)$ , поэтому они сами ограничены на  $\mathbb{C} \setminus \Delta$ . Следовательно, функции  $\alpha_0(\lambda), \dots, \alpha_{n-1}(\lambda)$  ограничены на  $\mathbb{C}$ , т.е. они постоянны:  $\alpha_i(\lambda) = \alpha_i$ . Пусть  $p(T) = T^n - \alpha_{n-1}T^{n-1} - \dots - \alpha_0I$ . Тогда  $p(T)f_0(\lambda) = 0$  для  $\lambda \in \mathbb{C} \setminus \Delta$ , а значит,  $p(T)f_0(\lambda) = 0$  для всех  $\lambda$ . В частности,  $p(T)\xi_0 = 0$ . Следовательно,  $p = p_0$  и  $p_0(T)\eta = 0$ .  $\square$

## Задачи

**3.1.** Векторы  $e_1, \dots, e_{m+1}$  в  $\mathbb{R}^n$  линейно независимы. Докажите, что среди их линейных комбинаций есть ненулевой вектор, у которого первые  $m$  координат нулевые.

**3.2.** В пространстве  $V^n$  даны векторы  $e_1, \dots, e_m$ . Докажите, что если  $m \geq n + 2$ , то существуют такие числа  $\alpha_1, \dots, \alpha_n$ , не все равные нулю, что  $\sum \alpha_i e_i = 0$  и  $\sum \alpha_i e_i = 0$ .

**3.3.** Выпуклой линейной комбинацией векторов  $v_1, \dots, v_m$  называют любой вектор  $x = t_1 v_1 + \dots + t_m v_m$ , где  $t_i \geq 0$  и  $\sum t_i = 1$ . Докажите, что в вещественном пространстве размерности  $n$  любая выпуклая линейная комбинация  $m$  векторов является также выпуклой линейной комбинацией не более чем  $n + 1$  из этих векторов.

**3.4.** Докажите, что если  $|a_{ii}| > \sum_{k \neq i} |a_{ik}|$  при  $i = 1, \dots, n$ , то матрица  $A = \|a_{ij}\|$  невырожденная.

**Замечание.** Матрицы такого вида называют *матрицами с доминирующей диагональю*.



**3.5.** Векторы  $x_1, \dots, x_k$  линейно независимы. Известно, что если в базисе  $e_1, \dots, e_n$  вектор  $e_i$  заменить на  $x_i$ , то снова получится базис (при любом  $i$ ). Обязательно ли  $x_1, \dots, x_k, e_{k+1}, \dots, e_n$  — базис?

**3.6.** Пусть  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_n$  — два базиса, причём первый базис положительно ориентирован и каждый набор векторов  $v_1, \dots, v_n$ , где  $v_i = e_i$  или  $\varepsilon_i$ , является базисом. Докажите, что следующие условия эквивалентны:

(1) базис  $v_1, \dots, v_n$  положительно ориентирован тогда и только тогда, когда среди этих векторов встречается чётное число векторов  $\varepsilon_i$ ;

(2) базисы  $v_1, \dots, e_i, \dots, v_n$  и  $v_1, \dots, \varepsilon_i, \dots, v_n$  всегда противоположно ориентированы.

**3.7.** а) В  $n$ -мерном евклидовом пространстве даны векторы  $e_1, \dots, e_{n+1}$ , причём  $(e_i, e_j) < 0$  при  $i \neq j$ . Докажите, что любые  $n$  из этих векторов образуют базис.

б) Докажите, что в  $n$ -мерном евклидовом пространстве  $n + 2$  вектора не могут образовывать попарно тупые углы.

**3.8.** В  $n$ -мерном евклидовом пространстве даны ненулевые векторы  $e_1, \dots, e_m$  так, что  $(e_i, e_j) \leq 0$  при  $i \neq j$ . Докажите, что  $m \leq 2n$ .

**3.9.** Для каждой пары  $i, j$ , где  $1 \leq i < j \leq n + 1$ , задано натуральное число  $a_{ij}$ . Докажите, что в  $\mathbb{R}^n$  найдутся векторы  $x_1, \dots, x_{n+1}$ , для которых  $(x_i, x_j) = a_{ij}$  при  $i < j$ , причём векторы  $x_1, \dots, x_n$  линейно независимы.

**3.10.** [Еп] Пусть  $M$  — выпуклый многогранник в  $\mathbb{R}^n$ , имеющий  $m$  граней размерности  $n - 1$ . Докажите, что в  $\mathbb{R}^{n+m-1} \supset \mathbb{R}^n$  существует (аффинный)  $m$ -гранный угол с попарно ортогональными  $(n + m - 2)$ -мерными гранями, сечение которого подпространством  $\mathbb{R}^n$  — это данный многогранник  $M$ .

## 2.4. Ранг матрицы

### 2.4.1. Неравенства для ранга матрицы

Столбцы матрицы  $A$  являются линейными комбинациями столбцов матрицы  $A$ , поэтому  $\text{rk } AB \leq \text{rk } A$ ; строки матрицы  $AB$  являются линейными комбинациями строк матрицы  $B$ , поэтому  $\text{rk } AB \leq \text{rk } B$ . Если матрица  $B$  обратима, то  $\text{rk } A = \text{rk}(AB)B^{-1} \leq \text{rk } A$ , поэтому  $\text{rk } A = \text{rk } AB$ .

Приведём ещё два неравенства для рангов произведений матриц.

**Теорема 2.4.1.1 (неравенство Фробениуса).**  $\text{rk } BC + \text{rk } AB \leq \text{rk } ABC + \text{rk } B$ .

**Доказательство.** Если  $U \subset V$  и  $X : V \rightarrow W$ , то  $\dim(\text{Ker } X|_U) \leq \dim \text{Ker } X = \dim V - \dim \text{Im } X$ . Запишем это неравенство для  $U = \text{Im } BC$ ,  $V = \text{Im } B$  и  $X = A$ :

$$\dim(\text{Ker } A|_{\text{Im } BC}) \leq \dim \text{Im } B - \dim \text{Im } AB.$$

Ясно также, что  $\dim(\text{Ker } A|_{\text{Im } BC}) = \dim \text{Im } BC - \dim \text{Im } ABC$ .  $\square$

**Теорема 2.4.1.2 (неравенство Сильвестра).**  $\text{rk } A + \text{rk } B \leq \text{rk } AB + n$ , где  $n$  — число столбцов матрицы  $A$  и число строк матрицы  $B$ .

**Доказательство.** Запишем неравенство Фробениуса для матриц  $A_1 = A$ ,  $B_1 = I$  и  $C_1 = B$ . В итоге получим  $\text{rk } B_1 C_1 + \text{rk } A_1 B_1 \leq \text{rk } A_1 B_1 C_1 + \text{rk } B_1$ , т.е.  $\text{rk } B + \text{rk } A \leq \text{rk } AB + n$ .  $\square$

### 2.4.2. Другое определение ранга

Для ранга матрицы можно дать другое определение: ранг матрицы  $A$  равен наименьшему из размеров матриц  $B$  и  $C$ , произведение которых равно  $A$ .

Докажем, что это определение эквивалентно обычному. Если  $A = BC$  и наименьший из размеров матриц  $B$  и  $C$  равен  $k$ , то  $\text{rk } A \leq \min(\text{rk } B, \text{rk } C) \leq k$ . Остаётся доказать, что если  $A$  — матрица размера  $m \times n$  и  $\text{rk } A = k$ , то матрицу  $A$  можно представить в виде произведения матриц размеров  $m \times k$  и  $k \times n$ . Выделим в матрице  $A$  линейно независимые столбцы  $A_1, \dots, A_k$ . Все остальные столбцы через них выражаются, поэтому

$$\begin{aligned} A &= (x_{11}A_1 + \dots + x_{k1}A_k, \dots, x_{1n}A_1 + \dots + x_{kn}A_k) = \\ &= (A_1, \dots, A_k) \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{k1} & \dots & x_{kn} \end{pmatrix}. \end{aligned}$$

### 2.4.3. Подпространства матриц ограниченного ранга

Пусть  $M_{m,n}$  — пространство матриц размера  $m \times n$ . В этом пространстве можно указать подпространство размерности  $mr$ , ранг элементов которого не превосходит  $r$ ; для этого достаточно взять матрицы, в последних  $m - r$  строках которых стоят нули. Докажем сначала, что если  $m \leq n$  и  $r$  — наибольший ранг элементов подпространства  $U \subset M_{m,n}$ , то  $\dim U \leq nr$ , а затем выясним, как устроены такие подпространства максимальной размерности.

**Теорема 2.4.3.1.** Пусть  $r \leq m \leq n$ ,  $U \subset M_{m,n}$  — линейное подпространство, причём наибольший ранг элементов  $U$  равен  $r$ . Тогда  $\dim U \leq nr$ .

**Доказательство.** [F11] Дополнив при необходимости матрицы блоком нулей, будем считать, что матрицы имеют размер  $n \times n$ . В пространстве  $U$  выберем матрицу  $A$  ранга  $r$ . Преобразованием  $X \mapsto PQX$ , где  $P$  и  $Q$  — невырожденные матрицы, матрицу  $A$  можно привести к виду  $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$  (см. теорему 2.2.3.2). Такое же преобразование произведём над всеми матрицами пространства  $U$  и запишем их в соответствующем блочном виде.

**Лемма 1.** Если  $B \in U$ , то  $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{pmatrix}$ , где  $B_{21}B_{12} = 0$ .

**Доказательство.** Пусть  $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \in U$ , причём матрица  $B_{21}$  состоит из строк  $u_1, \dots, u_{n-r}$ , а матрица  $B_{12}$  — из столбцов  $v_1, \dots, v_{n-r}$ . Любой минор порядка  $r+1$  матрицы  $tA + B$  равен нулю, поэтому

$$\Delta(t) = \begin{vmatrix} tI_r + B_{11} & v_j \\ u_i & b_{ij} \end{vmatrix} = 0.$$

Коэффициент при  $t^r$  равен  $b_{ij}$ , поэтому  $b_{ij} = 0$ . Следовательно,  $\Delta(t) = -u_i \operatorname{adj}(tI_r + B_{11})v_j$  (теорема 1.3.1.3). Так как  $\operatorname{adj}(tI_r + B_{11}) = t^r I_r + \dots$ , коэффициент при  $t^{r-1}$  многочлена  $\Delta(t)$  равен  $-u_i v_j$ . Поэтому  $u_i v_j = 0$ , а значит,  $B_{21}B_{12} = 0$ .  $\square$

**Лемма 2.** Если  $B, C \in U$ , то  $B_{21}C_{12} + C_{21}B_{12} = 0$ .

**Доказательство.** Применив лемму 1 к матрице  $B+C \in U$ , получим  $(B_{21} + C_{21})(B_{12} + C_{12}) = 0$ , т.е.  $B_{21}C_{12} + C_{21}B_{12} = 0$ .  $\square$

Перейдём непосредственно к доказательству теоремы 2.4.3.1. Рассмотрим отображение  $f : U \rightarrow M_{r,n}$ , заданное формулой  $f(C) = (C_{11}, C_{12})$ . Кег  $f$  состоит из матриц вида  $\begin{pmatrix} 0 & 0 \\ B_{21} & 0 \end{pmatrix}$ , причём согласно лемме 2  $B_{21}C_{12} = 0$  для всех матриц  $C \in U$ . Рассмотрим, далее, отображение  $g : \operatorname{Ker} f \rightarrow M_{r,n}^*$ ,  $g(B) : (X_{11}, X_{12}) \mapsto \operatorname{tr}(B_{21}X_{12})$ . Это отображение мономорфно (см. §2.1.9), поэтому пространство  $g(\operatorname{Ker} f) \subset M_{r,n}^*$  имеет размерность  $k = \dim \operatorname{Ker} f$ . Следовательно,  $(g(\operatorname{Ker} f))^\perp$  — подпространство размерности  $nr - k$  в  $M_{r,n}$ . Если  $C \in U$ , то  $B_{21}C_{12} = 0$ , а значит,  $\operatorname{tr} B_{21}C_{12} = 0$ . Поэтому  $f(U) \subset (g(\operatorname{Ker} f))^\perp$ , т.е.  $\dim f(U) \leq nr - k$ . Остаётся заметить, что  $\dim f(U) + k = \dim \operatorname{Im} f + \dim \operatorname{Ker} f = \dim U$ .  $\square$

Теорему 2.4.3.2 удобно сформулировать в инвариантной форме, поэтому вместо пространства  $M_{m,n}$  рассмотрим пространство  $\text{Hom}(N^n, M^n)$ , состоящее из линейных отображений  $N^n \rightarrow M^n$ .

**Теорема 2.4.3.2.** Пусть  $U \subset \text{Hom}(N^n, M^n)$  — линейное подпространство,  $r$  — наибольший ранг элементов  $U$  и  $\dim U = \max(mr, nr)$ . Тогда выполнено одно из следующих условий:

- 1)  $m \geq n$  и в  $N$  существует такое  $(n-r)$ -мерное подпространство  $N_1$ , что  $U = \{A \mid AN_1 = 0\}$ ;
- 2)  $m \leq n$  и в  $M$  существует такое  $r$ -мерное подпространство  $M_1$ , что  $U = \text{Hom}(N^n, M_1)$ .

**Доказательство.** [F11] Инвариантные обозначения понадобились нам лишь для формулировки. Добавив при необходимости к пространству  $M^m$  или  $N^n$  прямое слагаемое, будем считать, что  $m = n$ . Воспользуемся результатами и обозначениями теоремы 2.4.3.1. Матрицы  $C \in U$  имеют вид  $\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & 0 \end{pmatrix}$ . Требуется доказать, что либо  $C_{12} = 0$  для всех  $C \in U$ , либо  $C_{21} = 0$ .

Рассмотрим снова отображение  $f : U \rightarrow M_{r,n}$ , заданное формулой  $f(C) = (C_{11}, C_{12})$ . Пусть  $H \subset M_{r,n-r}$  — подпространство, порождённое всеми матрицами  $C_{12}$ . Тогда  $f(U) \subset M_{r,r} \oplus H$ . Пространства  $M_{r,n-r}$  и  $M_{n-r,r}$  можно отождествить операцией транспонирования; для  $B_{21} \in M_{n-r,r}$  и  $X_{12} \in M_{r,n-r}$  величину  $\text{tr}(B_{21}X_{12})$  можно рассматривать как скалярное произведение в  $M_{r,n-r}$ . Относительно этого скалярного произведения  $H \subset (\text{Ker } f)^\perp$ , поэтому  $\dim H \leq r(n-r) - \dim \text{Ker } f$ , а значит,  $\dim f(U) \leq r^2 + \dim H \leq nr - \dim \text{Ker } f$ . А так как  $nr = \dim \text{Im } f + \dim \text{Ker } f$ , то  $H = (\text{Ker } f)^\perp$ ,  $\dim H + \dim \text{Ker } f = r(n-r)$  и  $\dim f(U) = r^2 + \dim H$ , т.е.  $f(U) = M_{r,r} \oplus H$  и  $\text{Ker } f = H^\perp$ .

Выберем в  $M_{n-r,r}$  подпространство  $L$  так, что  $M_{n-r,r} = \text{Ker } f \oplus L$ . Существует единственное линейное отображение  $\lambda : f(U) \rightarrow L$ , делающее диаграмму

$$\begin{array}{ccc} U & \xrightarrow{f} & M_{r,r} \oplus H = f(U) \\ h \downarrow & & \downarrow \lambda \\ M_{n-r,r} & \xrightarrow{p} & L \end{array}$$

коммутативной; здесь  $h \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & 0 \end{pmatrix} = C_{21}$ ,  $p$  — естественная проекция  $M_{n-r,r} = \text{Ker } f \oplus L$  на  $L$ . В самом деле, если  $(C_{11}, C_{12}) \in f(U)$ , то в  $M_{n-r,r}$  найдётся такая матрица  $C_{21}$ , что  $\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & 0 \end{pmatrix} \in U$ , причём

матрица  $C_{21}$  определена с точностью до слагаемого из  $\text{Ker } f$ . Таким образом, мы получаем отображение  $\lambda : f(U) \rightarrow M_{n-r,r} / \text{Ker } f = L$ . Легко проверить, что это отображение линейно. Докажем, что  $\lambda = 0$ . Если  $C_{11} \in M_{r,r}$  и  $C_{12} \in H$ , то существует матрица  $\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & 0 \end{pmatrix} \in U$ . Тогда  $\lambda(C_{11}, C_{12}) = C_{21}$ , поэтому согласно лемме 1  $\lambda(C_{11}, C_{12})C_{12} = 0$ . Аналогично,  $\lambda(C_{11}, -C_{12})C_{12} = 0$ . Следовательно,  $\lambda(C_{11}, 0)C_{12} = \lambda(C_{11}, C_{12} - C_{12})C_{12} = 0$ , т.е.  $\lambda(C_{11}, 0) \in H^\perp = \text{Ker } f$ , а значит,  $\lambda(C_{11}, 0) \in L \cap \text{Ker } f = 0$ . Величина  $\lambda(C_{11}, C_{12}) = \lambda(C_{11}, 0) + \lambda(0, C_{12}) = \lambda(0, C_{12})$  зависит только от  $C_{12}$ ; обозначим её  $\lambda(C_{12})$ . Пусть  $C_{12} \in H$  — ненулевая матрица,  $v$  — её ненулевой столбец. Тогда если  $C_{11} \in M_{r,r}$  и  $u$  — строка  $\lambda(C_{12})$ , то

$$0 = \begin{vmatrix} C_{11} & v \\ u & 0 \end{vmatrix} = -u \cdot \text{adj } C_{11} \cdot v.$$

Матрицу  $C_{11}$  можно подобрать так, чтобы матрица  $\text{adj } C_{11}$  имела ровно один ненулевой элемент, причём этот элемент может стоять в любом заданном месте (см. задачу 2.10). Следовательно,  $u = 0$ , а значит,  $\lambda(C_{12}) = 0$ .

Итак, если  $\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & 0 \end{pmatrix} \in U$  и  $C_{12} \neq 0$ , то  $C_{21} = \lambda(C_{12}) = 0$ . Поэтому в любой матрице из  $U$  одна из матриц  $C_{12}$  и  $C_{21}$  нулевая. Ясно также, что матрицы  $\begin{pmatrix} B_{11} & B_{12} \\ 0 & 0 \end{pmatrix}$  и  $\begin{pmatrix} C_{11} & 0 \\ C_{21} & 0 \end{pmatrix}$  с ненулевыми матрицами  $B_{12}$  и  $C_{12}$  не могут одновременно входить в  $U$ , так как иначе их сумма не входила бы в  $U$ .

Соображения размерности показывают, что пространство, порождённое матрицами  $C_{12}$  или  $C_{21}$ , имеет максимальную размерность  $r(n-r)$ , т.е. оно совпадает с пространством  $M_{r,n-r}$  или  $M_{n-r,r}$ .  $\square$

### Задачи

**4.1.** Пусть  $a_{ij} = x_i + y_j$ . Докажите, что  $\text{rk } \|a_{ij}\| \leq 2$ .

**4.2.** Матрица  $A$  квадратная, причём  $\text{rk } A = 1$ . Докажите, что  $|A + I| = \text{tr } A + 1$ .

**4.3.** Докажите, что  $\text{rk}(A^*A) = \text{rk } A$  (определение оператора  $A^*$  см. в §2.7.4).

**4.4.** Пусть  $A$  — невырожденная матрица. Докажите, что ранг матрицы  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$  равен рангу матрицы  $A$  тогда и только тогда, когда  $D = CA^{-1}B$ .

**4.5.** Пусть  $B$  — билинейная функция на  $V \times W$ ;  $N_V = \{v \in V \mid B(v, w) = 0 \ \forall w \in W\}$  и  $N_W = \{w \in W \mid B(v, w) = 0 \ \forall v \in V\}$ . Докажите, что  $\dim V - \dim N_V = \dim W - \dim N_W$ .

**4.6.** Матрицы  $A_1$  и  $A_2$  имеют одинаковые размеры. Пространства  $V_1$  и  $V_2$  порождены их строками, пространства  $W_1$  и  $W_2$  — столбцами. Докажите, что следующие условия эквивалентны:

- 1)  $\text{rk}(A_1 + A_2) = \text{rk } A_1 + \text{rk } A_2$ ;
- 2)  $V_1 \cap V_2 = 0$ ;
- 3)  $W_1 \cap W_2 = 0$ .

**4.7.** Докажите, что если  $A$  и  $B$  — матрицы одного размера и  $B^T A = 0$ , то  $\text{rk}(A + B) = \text{rk } A + \text{rk } B$ .

**4.8.** Пусть  $A$  и  $B$  — квадратные матрицы нечётного порядка. Докажите, что если  $AB = 0$ , то хотя бы одна из матриц  $A + A^T$  и  $B + B^T$  вырожденная.

**4.9.** Точки  $X_1, \dots, X_n$  расположены на окружности в указанном порядке. Рассмотрим кососимметрическую матрицу  $A = \|a_{ij}\|$ , где  $a_{ij} = X_i X_j$  при  $i > j$ . Докажите, что  $\text{rk } A = 2$  (обобщённая теорема Птолемея).

## 2.5. Подпространства. Ортогонализация

### 2.5.1. Размерность пересечения двух пространств

Размерность пересечения двух подпространств и размерность пространства, порождённого ими, связаны следующим соотношением.

**Теорема 2.5.1.1.**  $\dim(V + W) + \dim(V \cap W) = \dim V + \dim W$ .

**Доказательство.** Пусть  $e_1, \dots, e_r$  — базис  $V \cap W$ ; его можно дополнить до базиса  $e_1, \dots, e_r, v_1, \dots, v_{n-r}$  пространства  $V^n$  и до базиса  $e_1, \dots, e_r, w_1, \dots, w_{m-r}$  пространства  $W^m$ . Тогда  $e_1, \dots, e_r, v_1, \dots, v_{n-r}, w_1, \dots, w_{m-r}$  — базис пространства  $V + W$ . Поэтому  $\dim(V + W) + \dim(V \cap W) = (r + n - r + m - r) + r = n + m = \dim V + \dim W$ .  $\square$

### 2.5.2. Ортогонализация Грама—Шмидта

Пусть  $V$  — пространство над  $\mathbb{R}$ . Скалярным произведением в  $V$  называют отображение, сопоставляющее паре векторов  $u, v \in V$  число  $(u, v) \in \mathbb{R}$  и обладающее следующими свойствами:

- 1)  $(u, v) = (v, u)$ ;
- 2)  $(\lambda u + \mu v, w) = \lambda(u, w) + \mu(v, w)$ ;

3)  $(u, u) > 0$  для любого  $u \neq 0$ ; величину  $\|u\| = \sqrt{(u, u)}$  называют длиной вектора  $u$ .

Базис  $e_1, \dots, e_n$  пространства  $V$  со скалярным произведением называют *ортонормированным* (соответственно *ортогональным*), если  $(e_i, e_j) = \delta_{ij}$  (соответственно  $(e_i, e_j) = 0$  при  $i \neq j$ ).

Матрицу  $A$  перехода от одного ортонормированного базиса к другому ортонормированному базису называют *ортогональной*. Столбцы такой матрицы образуют ортонормированную систему векторов, поэтому  $A^T A = I$ , а значит,  $A^T = A^{-1}$  и  $AA^T = I$ . Если  $A$  — ортогональная матрица, то  $(Ax, Ay) = (x, A^T Ay) = (x, y)$  (см. §§2.1.3 и 2.1.4).

Легко проверить, что любые векторы  $e_1, \dots, e_n$ , обладающие свойством  $(e_i, e_j) = \delta_{ij}$ , линейно независимы. В самом деле, если  $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$ , то  $\lambda_i = (\lambda_1 e_1 + \dots + \lambda_n e_n, e_i) = 0$ . Аналогично можно доказать, что ортогональная система ненулевых векторов линейно независима.

**Теорема 2.5.2.1 (ортогонализация Грама—Шмидта).** Пусть  $e_1, \dots, e_n$  — произвольный базис. Тогда существует ортогональный базис  $\varepsilon_1, \dots, \varepsilon_n$ , для которого  $\varepsilon_i \in \langle e_1, \dots, e_i \rangle$  для всех  $i = 1, \dots, n$ .

**Доказательство.** Применим индукцию по  $n$ . При  $n = 1$  утверждение очевидно. Предположим, что утверждение верно для  $n$  векторов. Рассмотрим базис  $e_1, \dots, e_{n+1}$ . По предположению индукции существует такой ортогональный базис  $\varepsilon_1, \dots, \varepsilon_n$ , что  $\varepsilon_i \in \langle e_1, \dots, e_i \rangle$  для  $i = 1, \dots, n$ . Рассмотрим вектор  $\varepsilon_{n+1} = \lambda_1 \varepsilon_1 + \dots + \lambda_n \varepsilon_n + e_{n+1}$ . Условие  $(\varepsilon_i, \varepsilon_{n+1}) = 0$  означает, что  $\lambda_i (\varepsilon_i, \varepsilon_i) + (e_{n+1}, \varepsilon_i) = 0$ , т.е.  $\lambda_i = -\frac{(e_{n+1}, \varepsilon_i)}{(\varepsilon_i, \varepsilon_i)}$ . Взяв такие  $\lambda_i$ , получим ортогональную систему векторов  $\varepsilon_1, \dots, \varepsilon_{n+1}$ , причём  $\varepsilon_{n+1} \neq 0$ , так как  $\varepsilon_{n+1} \notin \langle \varepsilon_1, \dots, \varepsilon_n \rangle = \langle e_1, \dots, e_n \rangle$ .  $\square$

**Замечание 1.** От ортогонального базиса  $\varepsilon_1, \dots, \varepsilon_n$  можно перейти к ортонормированному базису  $\varepsilon'_1, \dots, \varepsilon'_n$ , где  $\varepsilon'_i = \frac{1}{\sqrt{(\varepsilon_i, \varepsilon_i)}} \varepsilon_i$ .

**Замечание 2.** Процесс ортогонализации имеет достаточно простой геометрический смысл: из вектора  $e_{n+1}$  вычитается его ортогональная проекция на подпространство  $W = \langle e_1, \dots, e_i \rangle$  и в результате получается вектор  $\varepsilon_{n+1}$ , ортогональный  $W$ .

Доказательство теоремы 2.5.2.1 даёт рекуррентное построение ортогонального базиса  $\varepsilon_1, \dots, \varepsilon_n$ . Явное выражение вектора  $\varepsilon_k$  через  $e_1, \dots, e_k$  выглядит следующим образом.

**Теорема 2.5.2.2.** Пусть  $e_1, \dots, e_n$  — произвольный базис,  $g_{ij} = (e_i, e_j)$ . То-

гда векторы  $\varepsilon_1 = e_1$ ,  $\varepsilon_k = \begin{pmatrix} g_{11} & \dots & g_{1k} \\ \dots & \dots & \dots \\ g_{k-1,1} & \dots & g_{k-1,k} \\ e_1 & \dots & e_k \end{pmatrix}$ ,  $k = 2, \dots, n$ , образуют

ортogonalный базис, причём квадрат длины вектора  $\varepsilon_k$  равен  $G_{k-1}G_k$ , где

$G_k = \begin{pmatrix} g_{11} & \dots & g_{1k} \\ \dots & \dots & \dots \\ g_{k1} & \dots & g_{kk} \end{pmatrix}$  при  $k = 1, \dots, n$  и  $G_0 = 1$ .

**Доказательство.** Ясно, что  $(e_i, \varepsilon_k) = \begin{pmatrix} g_{11} & \dots & g_{1k} \\ \dots & \dots & \dots \\ g_{k-1,1} & \dots & g_{k-1,k} \\ g_{i1} & \dots & g_{ik} \end{pmatrix}$ . Поэтому

$(e_i, \varepsilon_k) = 0$  при  $i = 1, \dots, k-1$  и  $(e_k, \varepsilon_k) = G_k$ .

Непосредственно из формулы для  $\varepsilon_k$  видно, что  $\varepsilon_k = G_{k-1}e_k + e$ , где  $e \in \langle e_1, \dots, e_{k-1} \rangle$ . Поэтому  $(\varepsilon_k - G_{k-1}e_k, \varepsilon_k) = 0$ . Следовательно,  $(\varepsilon_k, \varepsilon_k) = G_{k-1}(e_i, \varepsilon_k) = G_{k-1}G_k$ .

Чтобы завершить доказательство, нужно убедиться, что все векторы  $\varepsilon_1, \dots, \varepsilon_n$  ненулевые, т.е.  $G_k \neq 0$  для  $k = 1, \dots, n$ . Выберем произвольный ортонормированный базис и рассмотрим матрицу  $X$ , образованную столбцами координат векторов  $e_1, \dots, e_k$  относительно этого базиса. Яс-

но, что  $X^T X = \begin{pmatrix} g_{11} & \dots & g_{1k} \\ \dots & \dots & \dots \\ g_{k1} & \dots & g_{kk} \end{pmatrix}$ , поэтому  $G_k = (\det X)^2$ . Остаётся заметить, что  $\det X \neq 0$ , поскольку векторы  $e_1, \dots, e_k$  линейно независимы.  $\square$

Матрицу  $\begin{pmatrix} (e_1, e_1) & \dots & (e_1, e_k) \\ \dots & \dots & \dots \\ (e_k, e_1) & \dots & (e_k, e_k) \end{pmatrix}$ , возникшую при доказательстве теоремы 2.5.2.2, называют *матрицей Грама* системы векторов  $e_1, \dots, e_k$ , а её определитель — *определителем Грама*. Как только что было доказано, определитель Грама системы линейно независимых векторов положителен, а определитель Грама системы линейно зависимых векторов равен 0.

### 2.5.3. Ортогональные проекции

Пусть в пространстве  $V$  задано скалярное произведение и  $W$  — подпространство в  $V$ . Вектор  $w \in W$  называют *ортogonalной проекцией* вектора  $v \in V$  на подпространство  $W$ , если  $v - w \perp W$ .



**Теорема 2.5.3.1.** Для любого вектора  $v \in V$  существует единственная ортогональная проекция на подпространство  $W$ .

**Доказательство.** Выберем в пространстве  $W$  ортонормированный базис  $e_1, \dots, e_k$ . Рассмотрим вектор  $w = \lambda_1 e_1 + \dots + \lambda_k e_k$ . Условие  $v - w \perp e_i$  означает, что  $0 = (\lambda_1 e_1 + \dots + \lambda_k e_k - v, e_i) = \lambda_i - (v, e_i)$ , т.е.  $\lambda_i = (v, e_i)$ . Выбрав такие числа  $\lambda_i$ , получим требуемый вектор  $w = \sum_{i=1}^k (v, e_i) e_i$ .  $\square$

**Следствие 1.** Если  $e_1, \dots, e_n$  — ортонормированный базис пространства  $V$  и  $v \in V$ , то  $v = \sum_{i=1}^n (v, e_i) e_i$ .

Достаточно заметить, что вектор  $v - \sum_{i=1}^k (v, e_i) e_i$  ортогонален всему пространству  $V$ .

**Следствие 2.** Если  $w$  и  $w^\perp$  — ортогональные проекции вектора  $v$  на подпространства  $W$  и  $W^\perp$ , то  $v = w + w^\perp$ .

Для доказательства достаточно дополнить ортонормированный базис пространства  $W$  до ортонормированного базиса всего пространства и воспользоваться следствием 1.

**Теорема 2.5.3.2.** Если  $w$  — ортогональная проекция вектора  $v$  на подпространство  $W$  и  $w_1 \in W$ , то

$$\|v - w_1\|^2 = \|v - w\|^2 + \|w - w_1\|^2.$$

**Доказательство.** Пусть  $a = v - w$  и  $b = w - w_1 \in W$ . По определению  $a \perp b$ , поэтому  $\|a + b\|^2 = (a + b, a + b) = \|a\|^2 + \|b\|^2$ .  $\square$

**Следствие 1.**  $\|v\|^2 = \|w\|^2 + \|v - w\|^2$ .

Для доказательства достаточно положить  $w_1 = 0$ .

**Следствие 2.**  $\|u - w_1\| \geq \|v - w\|$ , причём равенство достигается только при  $w = w_1$ .

## 2.5.4. Симметрия относительно подпространства

Симметрией относительно подпространства  $W \subset V$  называют преобразование  $S_W$  пространства  $V$ , переводящее вектор  $v$  в вектор  $v'$ , для которого выполняется равенство  $v + v' = 2w$ , где  $w$  — ортогональная проекция вектора  $v$  на подпространство  $W$ .

В аффинном пространстве симметрия относительно подпространства допускает следующее геометрическое описание. Из точки  $A$  мы опускаем перпендикуляр  $AB$  на подпространство  $W$  и берём точку  $A'$ , для которой  $AA' = 2AB$ .

Если  $e_1, \dots, e_r$  — ортонормированный базис пространства  $W$ , то  $w = (v, e_1)e_1 + \dots + (v, e_r)e_r$ . Поэтому

$$S_W(v) = 2w - v = 2(v, e_1)e_1 + \dots + 2(v, e_r)e_r - v.$$

В частности, если  $W$  — гиперплоскость и  $a$  — единичный вектор, ортогональный  $W$ , то  $S_W(v) = v - 2(v, a)a$ , поскольку  $v = (v, e_1)e_1 + \dots + (v, e_r)e_r + (v, a)a = w + (v, a)a$ . Отметим, что формулу  $S_W(v) = v - 2(v, a)a$  легко проверить и непосредственно. Действительно,  $S_W(v) = v + \lambda a$ , где число  $\lambda$  характеризуется тем свойством, что вектор  $S_W(v) + v$  лежит в гиперплоскости  $W$ , т.е.  $(S_W(v) + v, a) = 0$ . Поэтому  $(v + \lambda a + v, a) = 0$ , а значит,  $0 = 2(v, a) + \lambda(a, a) = 2(v, a) + \lambda$ .

Если  $W_1, \dots, W_r$  — попарно ортогональные гиперплоскости и  $W = \bigcap_{i=1}^r W_i$ , то  $S_W = S_{W_1} \circ \dots \circ S_{W_r}$ , поскольку оператор ортогонального проецирования на  $W$  можно представить в виде композиции ортогональных проецирований на  $W_1, \dots, W_r$ .

**Теорема 2.5.4.1.** *а) Симметрия  $S_W$  относительно подпространства  $W$  является ортогональным преобразованием.*

*б) Определитель преобразования  $S_W$  равен 1 тогда и только тогда, когда коразмерность подпространства  $W$  чётна.*

**Доказательство.** а) Пусть  $w$  — ортогональная проекция вектора  $v$  на подпространство  $W$ . Требуется доказать, что  $(v, v) = (2w - v, 2w - v)$ , т.е.  $(v, w) = (w, w)$ . Последнее равенство следует из того, что  $v - w \perp w$ .

б) Симметрию относительно любого подпространства можно представить в виде композиции симметрий относительно гиперплоскостей, поэтому достаточно доказать, что определитель симметрии относительно гиперплоскости равен  $-1$ . Пусть  $e_1$  — вектор единичной длины, ортогональный данной гиперплоскости. Дополним этот вектор до ортонормированного базиса  $e_1, \dots, e_n$ . В таком базисе рассматриваемая симметрия задаётся диагональной матрицей  $\text{diag}(-1, 1, \dots, 1)$ .  $\square$

**Теорема 2.5.4.2.** *а) Любое ортогональное преобразование  $A$  в  $n$ -мерном пространстве можно представить в виде композиции не более  $n$  симметрий относительно гиперплоскостей.*

б) Пусть  $A$  — нетождественное ортогональное преобразование в  $n$ -мерном пространстве, причём  $\dim \operatorname{Ker}(A - I) = k$ . Тогда  $A$  можно представить в виде композиции не более  $n - k$  симметрий относительно гиперплоскостей и нельзя представить в виде композиции менее  $n - k$  симметрий относительно гиперплоскостей.

**Доказательство.** а) Выберем в рассматриваемом пространстве ортонормированный базис  $e_1, \dots, e_n$ . Предположим, что  $Ae_1 \neq e_1$ . Рассмотрим гиперплоскость, состоящую из векторов  $v$ , для которых  $\|v - e_1\| = \|v - Ae_1\|$ . Пусть  $S_1$  — симметрия относительно этой гиперплоскости. Тогда  $S_1 Ae_1 = e_1$  и, кроме того, если  $Av = v$ , то  $S_1 Av = v$ . Действительно,  $\|v - e_1\| = \|Av - Ae_1\| = \|v - Ae_1\|$ , поэтому вектор  $v$  лежит в рассматриваемой гиперплоскости. Если же  $Ae_1 = e_1$ , то вместо оператора  $S_1 A$  мы берём сам оператор  $A$ , который в таком случае обладает обоими указанными свойствами. Затем берём вектор  $e_2$  и строим оператор  $S_2 S_1 A$  (вместо него может получиться оператор  $S_1 A$  или даже  $A$ , если хотя бы один из векторов  $e_1, e_2$  неподвижен при действии оператора  $A$ ). Прделав такую операцию для всех векторов базиса, получим оператор  $S_k \dots S_1 A$ . Этот оператор оставляет неподвижными все векторы базиса, поэтому  $S_k \dots S_1 A = I$ , а значит,  $A = S_1^{-1} \dots S_k^{-1} = S_1 \dots S_k$ .

б) Если  $\dim \operatorname{Ker}(A - I) = k$ , то мы можем выбрать ортонормированный базис  $e_1, \dots, e_{n-k}$  в ортогональном дополнении к  $\operatorname{Ker}(A - I) = k$  и повторить предыдущие рассуждения. В результате получим представление оператора  $A$  в виде композиции не более  $n - k$  симметрий относительно гиперплоскостей.

Заметим, что если  $X$  — произвольный оператор, а  $S$  — симметрия относительно некоторой гиперплоскости  $W$ , то  $\dim \operatorname{Ker}(SX - I) \geq \dim \operatorname{Ker}(X - I) - 1$ , поскольку  $\operatorname{Ker}(SX - I) \supset \operatorname{Ker}(X - I) \cap W$ . Поэтому если  $A$  — композиция симметрий  $S_1, \dots, S_r$  относительно гиперплоскостей, то  $k = \dim \operatorname{Ker}(A - I) = \dim \operatorname{Ker}(S_1 \dots S_r - I) \geq n - r$ , а значит,  $r \geq n - k$ .  $\square$

### 2.5.5. Угол между вектором и подпространством

Углом между прямой  $l$  и подпространством  $W$  называют угол между вектором  $v$ , задающим  $l$ , и вектором  $w$  — ортогональной проекцией  $v$  на  $W$  (если  $w = 0$ , то  $v \perp W$ ). Так как  $v - w \perp w$ , то  $(v, w) = (w, w) \geq 0$ , т.е. угол между векторами  $v$  и  $w$  не тупой.

Если  $w$  и  $w^\perp$  — ортогональные проекции единичного вектора  $v$  на  $W$  и  $W^\perp$ , то  $\cos \angle(v, w) = \|w\|$  и  $\cos \angle(v, w^\perp) = \|w^\perp\|$  (см. рис. 2.2), поэтому  $\cos \angle(v, W) = \sin \angle(v, W^\perp)$ .

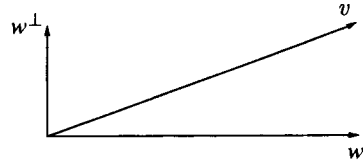


Рис. 2.2.

Пусть  $e_1, \dots, e_n$  — ортонормированный базис и  $v = x_1 e_1 + \dots + x_n e_n$  — единичный вектор. Тогда  $x_i e_i$  — проекция вектора  $v$  на  $i$ -ю ось координат, т.е.  $x_i = \cos \alpha_i$ , где  $\alpha_i$  — угол между векторами  $v$  и  $e_i$  (косинус угла между векторами  $v$  и  $e_i$  может отличаться знаком от косинуса угла между вектором  $v$  и  $i$ -й осью координат). Следовательно,

$$\sum_{i=1}^n \cos^2 \alpha_i = 1 \quad \text{и} \quad \sum_{i=1}^n \sin^2 \alpha_i = \sum_{i=1}^n (1 - \cos^2 \alpha_i) = n - 1.$$

**Теорема 2.5.5.1.** Пусть  $e_1, \dots, e_k$  — ортонормированный базис подпространства  $W \subset V$ ;  $\alpha_i$  — угол между векторами  $v$  и  $e_i$ ,  $\alpha$  — угол между  $v$  и  $W$ . Тогда  $\cos^2 \alpha = \sum_{i=1}^k \cos^2 \alpha_i$ .

**Доказательство.** Дополним базис  $e_1, \dots, e_k$  до базиса  $e_1, \dots, e_n$  пространства  $V$ . Тогда  $v = x_1 e_1 + \dots + x_n e_n$ , причём  $x_i = \cos \alpha_i$ , при  $i = 1, \dots, k$  и проекция вектора  $v$  на подпространство  $W$  равна  $x_1 e_1 + \dots + x_k e_k = w$ . Следовательно,  $\cos^2 \alpha = |w|^2 = x_1^2 + \dots + x_k^2 = \cos^2 \alpha_1 + \dots + \cos^2 \alpha_k$ .  $\square$

### 2.5.6. Проекции ортонормированных базисов

При рассмотрении ортогональных проекций ортонормированных базисов часто бывает полезен тот факт, что квадратная матрица ортогональна по строкам тогда и только тогда, когда она ортогональна по столбцам.

Начнём с простейшего случая, когда базис проецируется на прямую.

**Теорема 2.5.6.1.** Пусть  $e_1, \dots, e_n$  — векторы в  $\mathbb{R}^n$ . Сумма квадратов длин проекций этих векторов на любую прямую равна 1 тогда и только тогда, когда  $e_1, \dots, e_n$  — ортонормированный базис.

**Доказательство.** Пусть  $e_1 = (a_{11}, \dots, a_{1n}), \dots, e_n = (a_{n1}, \dots, a_{nn})$ . Рассмотрим матрицу  $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$ . Если прямая задаётся вектором-столбцом  $x$  единичной длины, то вектор-столбец  $Ax$  состоит из координат проекций векторов  $e_1, \dots, e_n$  на эту прямую. Сумма квадратов длин проекций на любую прямую равна 1 тогда и только тогда, когда

$$(a_{11}x_1 + \dots + a_{1n}x_n)^2 + \dots + (a_{n1}x_1 + \dots + a_{nn}x_n)^2 = x_1^2 + \dots + x_n^2$$

для любых  $x_1, \dots, x_n$ , т.е.  $\sum_{k=1}^n a_{ki}a_{kj} = \delta_{ij}$ . Это эквивалентно тому, что матрица  $A$  ортогональна по строкам, т.е.  $A$  ортогональна по столбцам. А ортогональность матрицы  $A$  по столбцам эквивалентна ортонормированности векторов  $e_1, \dots, e_n$   $\square$

Перейдём теперь к рассмотрению более сложной ситуации, когда векторы проецируются не на прямую, а на подпространство. Наше изложение во многом использует препринт [Еа].

**Теорема 2.5.6.2.** Векторы  $(x_1, y_1)$ ,  $(x_2, y_2)$  и  $(x_3, y_3)$  являются ортогональными проекциями трёх векторов  $\mathbb{R}^3$  равной длины тогда и только тогда, когда

$$(x_1 + iy_1)^2 + (x_2 + iy_2)^2 + (x_3 + iy_3)^2 = 0.$$

**Доказательство.** Векторы  $(x_1, y_1, z_1)$ ,  $(x_2, y_2, z_2)$  и  $(x_3, y_3, z_3)$  ортогональны и имеют одну и ту же длину  $\rho$  тогда и только тогда, когда матрица

$\frac{1}{\rho} \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$  ортогональна по столбцам, а это эквивалентно тому,

что она ортогональна по строкам, т.е. векторы  $(x_1, x_2, x_3)$ ,  $(y_1, y_2, y_3)$  и  $(z_1, z_2, z_3)$  ортогональны и имеют одну и ту же длину  $\rho$ . Таким образом, векторы  $(x_1, x_2, x_3)$  и  $(y_1, y_2, y_3)$  должны быть ортогональны и иметь равную длину. Это свойство не только необходимо, но и достаточно, поскольку пару ортонормированных векторов в  $\mathbb{R}^3$  всегда можно дополнить до тройки ортонормированных векторов.

Мы получили два уравнения  $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$  и  $x_1y_1 + x_2y_2 + x_3y_3 = 0$  для вещественных чисел. Эти два уравнения эквивалентны указанному комплексному уравнению.  $\square$

Непосредственно из теоремы 2.5.6.2 легко выводится аналогичное утверждение для проекций вершин правильного тетраэдра.

**Теорема 2.5.6.3.** Точки  $a = x_1 + iy_1$ ,  $b = x_2 + iy_2$ ,  $c = x_3 + iy_3$  и  $d = x_4 + iy_4$  являются ортогональными проекциями вершин правильного тетраэдра тогда и только тогда, когда

$$(a + b + c + d)^2 = 4(a^2 + b^2 + c^2 + d^2).$$

**Доказательство.** Правильный тетраэдр можно вписать в куб. Другая формулировка этого утверждения такова: четыре точки являются вершинами правильного тетраэдра тогда и только тогда, когда они являются концами векторов  $u$ ,  $v$ ,  $w$ ,  $u + v + w$ , где  $u$ ,  $v$ ,  $w$  — ортогональные векторы равной длины. Поэтому точки  $a$ ,  $b$ ,  $c$ ,  $d$  являются ортогональными проекциями вершин правильного тетраэдра тогда и только тогда, когда векторы  $a$ ,  $b$ ,  $c$  являются ортогональными проекциями трёх ортогональных векторов равной длины и  $d = a + b + c$ . Легко также проверить, что если  $d = a + b + c$ , то равенство  $(a + b + c + d)^2 = 4(a^2 + b^2 + c^2 + d^2)$  эквивалентно равенству  $a^2 + b^2 + c^2 = 0$ .  $\square$

### 2.5.7. Равные проекции ортогональных векторов

**Теорема 2.5.7.1.** Пусть  $e_1, \dots, e_n$  — ортогональный базис пространства  $V$ ,  $d_1, \dots, d_n$  — длины векторов  $e_1, \dots, e_n$ . Эти векторы имеют равные проекции на некоторое  $m$ -мерное подпространство  $W \subset V$  тогда и только тогда, когда  $d_i^2 \left( \frac{1}{d_1^2} + \dots + \frac{1}{d_n^2} \right) \geq m$  при  $i = 1, \dots, n$ .

**Доказательство.** [Ни] Выберем в пространстве  $W$  ортонормированный базис и дополним его до ортонормированного базиса  $\varepsilon_1, \dots, \varepsilon_n$  пространства  $V$ . Пусть  $(x_{1i}, \dots, x_{ni})$  — координаты вектора  $e_i$ , относительно базиса  $\varepsilon_1, \dots, \varepsilon_n$ ;  $y_{ki} = x_{ki}/d_i$ . Тогда  $\|y_{ki}\|$  — ортогональная матрица, причём длина проекции вектора  $e_i$  на подпространство  $W$  равна  $d$  тогда и только тогда, когда

$$y_{1i}^2 + \dots + y_{mi}^2 = (x_{1i}^2 + \dots + x_{mi}^2)/d_i^2 = d^2/d_i^2. \quad (1)$$

Если требуемое подпространство  $W$  существует, то  $d \leq d_i$  и  $m = \sum_{k=1}^m \sum_{i=1}^n y_{ki}^2 = \sum_{i=1}^n \sum_{k=1}^m y_{ki}^2 = d^2 \left( \sum_{i=1}^n 1/d_i^2 \right) \leq d_i^2 \left( \frac{1}{d_1^2} + \dots + \frac{1}{d_n^2} \right)$ .

Предположим теперь, что  $m \leq d_i^2 \left( \frac{1}{d_1^2} + \dots + \frac{1}{d_n^2} \right)$  при  $i = 1, \dots, n$  и построим ортогональную матрицу  $\|y_{ki}\|_1^n$ , обладающую свойством (1), где

$$d^2 = \frac{m}{\frac{1}{d_1^2} + \dots + \frac{1}{d_n^2}};$$

подпространство  $W$  после этого строится очевидным образом. Индукцией по  $n$  докажем, что если  $0 \leq \beta_i$  при  $i = 1, \dots, n$  и  $\beta_1 + \dots + \beta_n = m$ , то существует такая ортогональная матрица  $\|y_{ki}\|_1^n$ , что  $y_{1i}^2 + \dots + y_{mi}^2 = \beta_i$ . При  $n = 1$  утверждение очевидно. Предположим, что утверждение верно для  $n - 1$  и докажем его для  $n$ . Рассмотрим два случая.

а)  $m \leq n/2$ . Можно считать, что  $\beta_1 \geq \dots \geq \beta_n$ . Тогда  $\beta_{n-1} + \beta_n \leq 2m/n \leq 1$ , а значит, существует такая ортогональная матрица  $A = \|a_{ki}\|_1^{n-1}$ , что  $a_{1i}^2 + \dots + a_{mi}^2 = \beta_i$  при  $i = 1, \dots, n-2$  и  $a_{1,n-1}^2 + \dots + a_{m,n-1}^2 = \beta_{n-1} + \beta_n$ . Тогда матрица

$$\|y_{ki}\|_1^n = \begin{pmatrix} a_{11} & \dots & a_{1,n-2} & \alpha_1 a_{1,n-1} & -\alpha_2 a_{1,n-1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n-1,1} & \dots & a_{n-1,n-2} & \alpha_1 a_{n-1,n-1} & -\alpha_2 a_{n-1,n-1} \\ 0 & \dots & 0 & \alpha_2 & \alpha_1 \end{pmatrix},$$

где  $\alpha_1 = \sqrt{\frac{\beta_{n-1}}{\beta_{n-1} + \beta_n}}$  и  $\alpha_2 = \sqrt{\frac{\beta_n}{\beta_{n-1} + \beta_n}}$ , ортогональна по столбцам,  $\sum_{k=1}^m y_{ki}^2 = \beta_i$  при  $i = 1, \dots, n-2$ ,  $y_{1,n-1}^2 + \dots + y_{m,n-1}^2 = \alpha_1^2(\beta_{n-1} + \beta_n) = \beta_{n-1}$  и  $y_{1n}^2 + \dots + y_{mn}^2 = \alpha_2^2(\beta_{n-1} + \beta_n) = \beta_n$ . Кроме того, столбцы ортонормированы.

б) Пусть  $m > n/2$ . Тогда  $n - m \leq n/2$ , поэтому существует такая ортогональная матрица  $\|y_{ki}\|_1^n$ , что  $y_{m+1,i}^2 + \dots + y_{ni}^2 = 1 - \beta_i$  при  $i = 1, \dots, n$ , а значит,  $y_{1i}^2 + \dots + y_{mi}^2 = \beta_i$ .  $\square$

**Следствие.** Для ортогональных векторов  $e_1, \dots, e_n$  в  $\mathbb{R}^n$   $m$ -мерное подпространство  $W$ , образующее с ними данные углы  $\alpha_1, \dots, \alpha_n$ , существует тогда и только тогда, когда

$$\frac{\cos^2 \alpha_1 + \dots + \cos^2 \alpha_n}{\cos^2 \alpha_i} \geq m$$

при  $i = 1, \dots, n$ .

**Доказательство.** Если  $\alpha_i$  — угол между вектором  $e_i$  и подпространством  $W$ , то  $d_i = d / \cos \alpha_i$ .  $\square$

### 2.5.8. Определители, составленные из скалярных произведений

Пусть  $U$  и  $W$  —  $m$ -мерные подпространства в пространстве  $V$ ;  $e_1, \dots, e_m$  и  $\varepsilon_1, \dots, \varepsilon_m$  — их ортонормированные базисы.

**Теорема 2.5.8.1.** Пусть  $x_1, \dots, x_m$  и  $y_1, \dots, y_m$  — системы векторов в  $U$  и  $W$ . Тогда

$$|(x_i, y_j)|_1^m = |(x_i, e_j)|_1^m |(y_i, \varepsilon_j)|_1^m |(e_i, \varepsilon_j)|_1^m.$$

**Доказательство.** [Ко] Так как  $x_i = \sum_{k=1}^m (x_i, \varepsilon_k) \varepsilon_k + x'_i$ , где  $x'_i \perp W$  и  $y_j = \sum_{k=1}^m (y_j, \varepsilon_k) \varepsilon_k$ , то

$$|(x_i, y_j)|_1^m = \left| \sum_{k=1}^m (x_i, \varepsilon_k) (\varepsilon_k, y_j) \right|_1^m = |(x_i, \varepsilon_k)|_1^m |(\varepsilon_k, y_j)|_1^m.$$

А так как  $\varepsilon_k = \sum_{j=1}^k (\varepsilon_k, e_j) e_j + \varepsilon'_k$ , где  $\varepsilon_k \perp U$ , то

$$|(x_i, \varepsilon_k)|_1^m = \left| \sum_{j=1}^m (\varepsilon_k, e_j) (e_j, x_i) \right|_1^m = |(x_i, \varepsilon_j)|_1^m |(e_j, \varepsilon_k)|_1^m.$$

□

**Следствие 1.**  $(|(x_i, y_j)|_1^m)^2 = |(x_i, x_j)|_1^m |(y_i, y_j)|_1^m (|(e_i, \varepsilon_j)|_1^m)^2$ .

**Доказательство.** Из равенства  $x_i = \sum_{k=1}^m (x_i, e_k) e_k$  следует, что

$$|(x_i, x_j)|_1^m = \left| \sum_{k=1}^m (x_i, e_k) (e_k, x_j) \right|_1^m = |(x_i, e_k)|_1^m |(e_k, x_j)|_1^m.$$

Аналогично

$$|(y_i, y_j)|_1^m = |(y_i, \varepsilon_k)|_1^m |(\varepsilon_k, y_j)|_1^m.$$

Перемножим эти равенства и домножим полученное равенство на  $(|(e_i, \varepsilon_j)|_1^m)^2$ . Остаётся воспользоваться результатом теоремы 2.5.8.1. □

**Следствие 2.** Если  $x_1, \dots, x_m$  и  $y_1, \dots, y_m$  — ортонормированные базисы  $U$  и  $W$ , то  $(|(x_i, y_j)|_1^m)^2 = (|(e_i, \varepsilon_j)|_1^m)^2$ .

**Теорема 2.5.8.2.**  $|\det \|(e_i, \varepsilon_j)\|_1^m| \leq 1$ , причём равенство достигается только при  $U = V$ .

**Доказательство.** [Ко] По неравенству Адамара (см. §6.1.2)

$$(|(e_i, \varepsilon_j)|_1^m)^2 \leq \sum_{k=1}^m (e_1, \varepsilon_k)^2 \cdot \dots \cdot \sum_{k=1}^m (e_m, \varepsilon_k)^2 = l_1^2 \cdot \dots \cdot l_m^2,$$

где  $l_i$  — длина проекции вектора  $e_i$  на  $W$ . Ясно, что  $l_i \leq 1$ , причём  $l_1 = l_2 = \dots = l_m = 1$  тогда и только тогда, когда  $U = \langle e_1, \dots, e_m \rangle \subset W$ . □



### 2.5.9. Системы подпространств

**Теорема 2.5.9.1.** Пусть в пространстве  $V$  дана система  $k$ -мерных подпространств, причём любые два из них пересекаются по  $(k - 1)$ -мерному подпространству. Тогда все они либо имеют общее  $(k - 1)$ -мерное подпространство, либо лежат в одном  $(k + 1)$ -мерном подпространстве.

**Доказательство.** Предположим, что  $k$ -мерные пространства  $V_1$ ,  $V_2$  и  $V_3$  не имеют общего  $(k - 1)$ -мерного подпространства. Тогда  $(k - 1)$ -мерные пространства  $V_1 \cap V_2$  и  $V_1 \cap V_3$  различны, поэтому они порождают всё пространство  $V_1$ , а значит,  $\dim(V_1 \cap V_2 \cap V_3) = \dim(V_1 \cap V_2) + \dim(V_1 \cap V_3) - \dim V_1 = k - 2$ . Рассмотрим ортогональное дополнение  $W$  пространства  $V_1 \cap V_2 \cap V_3$ . Пересечения  $W$  с пространствами  $V_1$ ,  $V_2$  и  $V_3$  двумерны, и эти двумерные плоскости попарно пересекаются по прямым. Следовательно, они лежат в одном трёхмерном пространстве, а значит, пространства  $V_1$ ,  $V_2$  и  $V_3$  лежат в одном  $(k - 2) + 3 = (k + 1)$ -мерном пространстве. В частности,  $V_3 \subset \langle V_1, V_2 \rangle$ .

Докажем теперь, что любое другое данное подпространство  $V_4$  содержится в  $\langle V_1, V_2 \rangle$ . Если  $V_1$ ,  $V_2$  и  $V_4$  не имеют общего  $(k - 1)$ -мерного подпространства, то  $V_4 \subset \langle V_1, V_2 \rangle$ , поэтому пусть  $V_1 \cap V_2 \cap V_4 = V_1 \cap V_2$ . Так как  $V_4 \cap V_3 \neq V_1 \cap V_2$ , то  $V_4 \cap V_3$  содержит вектор  $v$ , не принадлежащий  $V_1 \cap V_2$ . Тогда  $\dim \langle V_1 \cap V_2, v \rangle = k$  и  $\langle V_1 \cap V_2, v \rangle \subset V_4$ , поэтому  $V_4 = \langle V_1 \cap V_2, v \rangle \subset \langle V_1, V_2, V_3 \rangle = \langle V_1, V_2 \rangle$ .  $\square$

Теорема 2.5.9.1 справедлива также и в пространстве  $V^*$ . Переходя к пространству  $V$ , для  $m = \dim V - k$  получим следующее утверждение.

**Теорема 2.5.9.2.** Пусть в пространстве  $V$  дана система  $m$ -мерных подпространств, причём любые два из них лежат в одном  $(m + 1)$ -мерном пространстве. Тогда все они либо имеют общее  $(m - 1)$ -мерное подпространство, либо лежат в одном  $(m + 1)$ -мерном подпространстве.

### 2.5.10. Теорема Шерка

Пусть в пространстве  $V$  заданы непустые множества векторов  $A_1, \dots, A_n$  (эти множества могут включать и нулевой вектор). Для любого целого  $m$ , где  $0 \leq m \leq n$ , рассмотрим следующие свойства:

$(P_{n,m})$  любые векторы  $a_1, \dots, a_m$ , где  $a_i \in A_i$ , порождают пространство размерности не более  $m$ ;

$(Q_{n,m})$  существует подпространство  $W$  размерности  $h$  ( $0 \leq h \leq m$ ), содержащее  $h + (n - m)$  из множеств  $A_i$ .

Свойства  $P_{n,n}$  и  $Q_{n,n}$  (с  $h = 0$ ) выполнены всегда.

**Теорема 2.5.10.1.** Свойства  $P_{n,m}$  и  $Q_{n,m}$  эквивалентны.

**Доказательство.** [Sc3] Если выполнено свойство  $Q_{n,m}$  и, например, множества  $A_1, \dots, A_{h+n-m}$  содержатся в  $W$ , то первые  $h+n-m$  из векторов  $a_1, \dots, a_n$  содержатся в  $h$ -мерном пространстве  $W$ , а остальные  $n-(h+n-m)$  векторов порождают пространство размерности не более  $m-h$ . Поэтому векторы  $a_1, \dots, a_n$  порождают пространство размерности не более  $h+(m-h)=m$ .

Предположим теперь, что выполнено свойство  $P_{n,m}$ . Доказательство свойства  $Q_{n,m}$  проведём двойной индукцией: по  $n$  и по  $f = \sum_{i=1}^n \dim \langle A_i \rangle$ . При  $n=1$  утверждение очевидно. В дальнейшем считаем, что утверждение доказано для  $n-1$ , где  $n \geq 2$ . Если выполнено свойство  $P_{n,0}$ , то каждое множество  $A_i$  состоит лишь из нулевого вектора, поэтому выполнено свойство  $Q_{n,0}$ . В дальнейшем считаем, что  $m \geq 1$ .

Если для множеств  $A_1, \dots, A_{n-1}$  выполнено свойство  $P_{n-1,m-1}$ , то по предположению индукции для них выполнено свойство  $Q_{n-1,m-1}$ , а значит, для множеств  $A_1, \dots, A_n$  выполнено свойство  $Q_{n,m}$ . Поэтому можно считать, что никакие  $n-1$  множеств из  $A_1, \dots, A_n$  не обладают свойством  $P_{n-1,m-1}$ . Тогда  $m < n$  и существуют векторы  $a_1, \dots, a_{n-1}$ , порождающие пространство  $U$ , размерность которого не меньше  $m$ . Из свойства  $P_{n,m}$  следует, что  $\dim U = m$  и  $A_n \subset U$ . В частности,  $\dim A_n \leq m$ . Аналогично  $\dim A_i \leq m$  при  $i = 1, \dots, n$ .

Докажем теперь, что свойство  $Q_{n,m}$  выполнено, если  $A_n = \{0\}$ . Для множеств  $A_1, \dots, A_{n-1}$  выполнено свойство  $P_{n-1,m}$ . Если  $m = n-1$ , то свойство  $Q_{n,m}$  выполнено с  $h = 0$ . Если  $m < n-1$ , то согласно предположению индукции для множеств  $A_1, \dots, A_{n-1}$  выполнено свойство  $Q_{n-1,m}$ , а значит, для множеств  $A_1, \dots, A_n$  выполнено свойство  $Q_{n,m}$ . Поэтому в дальнейшем считаем, что

$$A_i \neq \{0\} \quad \text{при} \quad i = 1, \dots, n. \quad (1)$$

**Лемма.** Пусть  $k$ -мерное подпространство  $W_1$ , где  $1 \leq k \leq m$ , содержит  $k$  из множеств  $A_1, \dots, A_n$ . Тогда для этих множеств выполнено свойство  $Q_{n,m}$ .

**Доказательство.** Можно считать, что  $A_1, \dots, A_k$  содержатся в  $W_1$ . Кроме того, можно считать, что число  $k$  в условии леммы выбрано минимальным. Тогда либо  $k=1$ , либо любое  $h$ -мерное подпространство, где  $0 < h < k$ , содержит менее  $h$  множеств из  $A_1, \dots, A_n$ . Во втором случае свойство  $Q_{k,k-1}$  не выполнено для множеств  $A_1, \dots, A_k$ , поэтому для них не выполнено свойство  $P_{k,k-1}$ , а значит, векторы  $b_1, \dots, b_k$ , где  $b_i \in A_i$ ,

можно выбрать так, что они образуют базис  $W_1$ . При  $k = 1$  аналогичное утверждение очевидно следует из (1).

Пусть  $p$  — проекция  $V$  параллельно  $W_1$  на подпространство, дополнительное к  $W_1$ . Векторы  $b_1, \dots, b_k, a_{k+1}, \dots, a_n$  порождают подпространство размерности не более  $m$ , поэтому векторы  $p(a_{k+1}), \dots, p(a_n)$  порождают пространство размерности не более  $m - k$ . Следовательно, множества  $p(A_{k+1}), \dots, p(A_n)$  обладают свойством  $P_{n-k, m-k}$ . Согласно предположению индукции они обладают свойством  $Q_{n-k, m-k}$ , т.е. для некоторого  $g$  ( $0 \leq g \leq m - k$ ) множества  $p(A_{i_1}), \dots, p(A_{i_{g+n-m}})$  лежат в  $g$ -мерном подпространстве. Следовательно, множества  $A_1, \dots, A_k, A_{i_1}, \dots, A_{i_{g+n-m}}$  лежат в  $(g + k)$ -мерном подпространстве. Это завершает доказательство леммы с  $h = g + k$ .  $\square$

Завершим теперь доказательство свойства  $Q_{n, m}$ . Если  $f = \sum_{i=1}^n \dim \langle A_i \rangle < n$ , то утверждение очевидно следует из (1). Предположим теперь, что утверждение верно для  $f - 1$ . Согласно лемме можно считать, что  $\dim A_n > 1$ . Пусть  $B_n$  состоит из единственного элемента  $b_n \in A_n, b_n \neq 0$ . Множества  $A_1, \dots, A_{n-1}, B_n$  обладают свойством  $P_{n, m}$  и  $\sum_{i=1}^{n-1} \dim \langle A_i \rangle + \dim \langle B_n \rangle < f$ , так как  $\dim B_n = 1 < \dim A_n$ . Следовательно, существует  $k$ -мерное подпространство  $Z$ , содержащее  $k + (n - m)$  из множеств  $A_1, \dots, A_{n-1}$ , причём  $0 \leq k \leq m$ . Согласно (1)  $k \geq 1$ . Пространство  $Z$  содержит  $k + (n - m) - 1 \geq k$  множеств из  $A_1, \dots, A_n$ . Остаётся воспользоваться леммой.  $\square$

## Задачи

**5.1.** Пусть  $e_1, \dots, e_n$  — ортонормированный базис,  $\varepsilon_1, \dots, \varepsilon_k$  — ортонормированная система векторов,  $\alpha_{ij}$  — угол между векторами  $e_i$  и  $\varepsilon_j$ . Докажите, что  $\sum_{i=1}^n \sum_{j=1}^k \cos^2 \alpha_{ij} = k$ .

**5.2.** Пусть  $v$  и  $w$  — два вектора,  $v'$  и  $w'$  — их ортогональные проекции на одно и то же подпространство  $V$ . Докажите, что  $(v', w) = (v, w')$ .

**5.3.** В  $n$ -мерном пространстве  $V$  даны  $m$ -мерные подпространства  $U$  и  $W$ , причём  $u \perp W$  для некоторого  $u \in U \setminus \{0\}$ . Докажите, что  $w \perp U$  для некоторого  $w \in W \setminus \{0\}$ .

**5.4.** В  $n$ -мерном евклидовом пространстве даны две системы векторов  $x_1, \dots, x_k$  и  $y_1, \dots, y_k$ , причём  $(x_i, x_j) = (y_i, y_j)$  при всех  $i, j$ . Докажите, что существует ортогональный оператор  $U$ , отображающий  $x_i$  в  $y_i$ .

**5.5.** Предположим, что попарные углы между векторами  $e_1, \dots, e_n$  больше  $90^\circ + \alpha$ , где  $\alpha > 0$ . Докажите, что  $n < \frac{1}{\sin \alpha} + 1$ .

**5.6.** Докажите, что многочлен  $x_1^2 + \dots + x_n^2$  нельзя представить в виде суммы менее чем  $n$  квадратов многочленов с вещественными коэффициентами от переменных  $x_1, \dots, x_n$ .

**5.7.** В  $\mathbb{R}^n$ , где  $n \geq 2$ , дано замкнутое выпуклое множество  $C$ , не содержащее точку  $O$ . Докажите, что через точку  $O$  можно провести: а) прямую; б) гиперплоскость, не пересекающую  $C$ .

**5.8.** Докажите, что любое ортогональное преобразование с положительным определителем в пространстве размерности  $n \geq 3$  можно представить в виде композиции отражений относительно подпространств размерности  $n - 2$ .

**5.9.** Пусть  $V$  — линейное подпространство в  $\mathbb{R}^n$ ,  $V^\perp$  — его ортогональное дополнение. Докажите, что хотя бы в одном из этих двух подпространств есть ненулевой вектор с неотрицательными координатами.

**5.10.** а) Пусть  $V$  — линейное подпространство в  $\mathbb{R}^n$ , которое не содержит ненулевых векторов с неотрицательными координатами. Докажите, что если  $\dim V \leq n - 2$ , то  $V$  содержится в гиперплоскости, которая тоже не содержит ненулевых векторов с неотрицательными координатами.

б) С помощью задачи а) получите ещё одно решение задачи 5.9.

**5.11.** Пусть  $A$  — произвольная вещественная матрица размером  $m \times n$ . Докажите, что либо найдётся ненулевой вектор  $x = (x_1, \dots, x_n)$  с неотрицательными координатами, для которого вектор  $Ax^T$  имеет неотрицательные координаты, либо найдётся ненулевой вектор  $y = (y_1, \dots, y_m)$ , для которого вектор  $-A^T y^T$  имеет неотрицательные координаты (*теорема об альтернативе*).

**5.12.** Пусть  $A$  — произвольная вещественная матрица размером  $m \times n$ ,  $J$  — матрица такого же размера, все элементы которой равны 1. Докажите, что существует вещественное число  $\lambda$ , для которого все координаты вектора  $(A - \lambda J)x^T$  неотрицательны для некоторого ненулевого вектора  $x = (x_1, \dots, x_n)$  с неотрицательными координатами и все координаты вектора  $(-A^T + \lambda J^T)y^T$  неотрицательны для некоторого ненулевого вектора  $y = (y_1, \dots, y_m)$  с неотрицательными координатами (*теорема о минимаксе*).

**5.13.** а) Докажите, что сумма квадратов длин проекций сторон плоского четырёхугольника на любую прямую, лежащую в его плоскости, постоянна тогда и только тогда, когда векторы сторон этого четырёхугольника являются проекциями векторов, идущих из центра правильного тетраэдра в его вершины.

б) Докажите, что сумма квадратов длин проекций сторон плоского  $n$ -угольника на любую прямую, лежащую в его плоскости, постоянна тогда

и только тогда, когда векторы сторон этого  $n$ -угольника являются проекциями векторов, идущих из центра правильного  $(n - 1)$ -мерного симплекса в его вершины.

## 2.6. Ортогональные многочлены

В пространстве многочленов степени не выше  $n - 1$  можно рассмотреть скалярное произведение  $(f, g) = \int_a^b w(x)f(x)g(x)dx$ , где  $w(x)$  — неотрицательная *весовая функция*. Если применить процесс ортогонализации относительно этого скалярного произведения к многочленам  $1, x, x^2, \dots, x^{n-1}$ , то мы получим последовательность многочленов  $\varphi_0(x), \varphi_1(x), \dots, \varphi_{n-1}(x)$ . При возрастании  $n$  начальный участок этой последовательности не изменяется, поэтому мы получаем бесконечную последовательность *ортогональных многочленов*. Ортогональность системы векторов не нарушается при замене их на пропорциональные векторы, поэтому общее определение таково:  $\{\varphi_k(x)\}, k = 0, 1, \dots$ , — последовательность ортогональных многочленов, если  $\varphi_k(x)$  — многочлен степени  $k$  и  $(\varphi_i, \varphi_j) = 0$  при  $i \neq j$ .

Мы сначала обсудим некоторые общие свойства последовательностей ортогональных многочленов, а затем — свойства наиболее известных последовательностей ортогональных многочленов для конкретных  $a$  и  $b$  и весовых функций  $w(x)$ .

### 2.6.1. Общие свойства

Пусть  $\{\varphi_k(x)\}$  — последовательность ортогональных многочленов.

**Теорема 2.6.1.1.** *Внутри отрезка  $[a, b]$  многочлен  $\varphi_n$  имеет ровно  $n$  некратных корней.*

**Доказательство.** Пусть  $x_1, \dots, x_m$  — корни многочлена  $\varphi_n(x)$  на отрезке  $[a, b]$ , причём каждый корень берётся столько раз, какова его кратность. Предположим, что либо  $m < n$ , либо  $m = n$ , но среди данных корней есть совпадающие. В первом случае рассмотрим многочлен  $\psi(x)$ , равный  $(x - x_1) \dots (x - x_m)$ , а во втором случае поделим этот многочлен на  $(x - x_i)^2$ , где  $x_i$  — кратный корень. Многочлен  $\psi(x)$  обладает двумя свойствами: 1) на всём отрезке  $[a, b]$  многочлен  $\varphi_n(x)\psi(x)$  либо неотрицательный, либо неположительный; 2) степень многочлена  $\psi(x)$  строго меньше  $n$ . Из свойства 1) следует, что  $(\varphi_n, \psi) \neq 0$  (ясно, что многочлен  $\varphi_n(x)\psi(x)$  ненулевой). Из свойства 2) следует, что  $\psi = \lambda_0\varphi_0 + \dots + \lambda_{n-1}\varphi_{n-1}$ , где  $\lambda_0,$

$\dots, \lambda_{n-1}$  — константы. Поэтому  $(\varphi_n, \psi) = 0$ . Приходим к противоречию.  $\square$

**Теорема 2.6.1.2.** Любые три последовательных ортогональных многочлена связаны соотношением

$$\varphi_{n+1}(x) = (a_n x + b_n)\varphi_n(x) + c_n \varphi_{n-1}(x), \quad n = 1, 2, \dots,$$

где  $a_n, b_n$  и  $c_n$  — некоторые константы.

**Доказательство.** Пусть  $\varphi_n(x) = k_n x^n + k'_n x^{n-1} + \dots$ . Коэффициент многочлена  $\varphi_{n+1}(x) - a_n x \varphi_n(x)$  при  $x^{n+1}$  равен  $k_{n+1} - a_n k_n$ , поэтому если  $a_n = k_{n+1}/k_n$ , то этот многочлен имеет степень не выше  $n$ , а значит,

$$\varphi_{n+1}(x) - a_n x \varphi_n(x) = \lambda_0 \varphi_0(x) + \dots + \lambda_n \varphi_n(x)$$

для некоторых констант  $\lambda_0, \dots, \lambda_n$ .

Из условия ортогональности следует, что  $(\varphi_m(x), x^k) = 0$  при  $k \leq m-1$ . Поэтому  $(x \varphi_n(x), x^k) = (\varphi_n(x), x^{k+1}) = 0$  при  $k \leq n-2$ . Таким образом,

$$(\lambda_0 \varphi_0(x) + \dots + \lambda_n \varphi_n(x), x^k) = 0$$

при  $k \leq n-2$ . Любой многочлен  $\varphi_0, \dots, \varphi_{n-2}$  линейно выражается через  $1, x, \dots, x^{n-2}$ , поэтому

$$(\lambda_0 \varphi_0(x) + \dots + \lambda_n \varphi_n(x), \varphi_k(x)) = 0$$

при  $k \leq n-2$ . Следовательно,  $\lambda_0 = \dots = \lambda_{n-2} = 0$ . В итоге получаем

$$\varphi_{n+1}(x) - a_n x \varphi_n(x) = \lambda_{n-1} \varphi_{n-1}(x) + \lambda_n \varphi_n(x).$$

Остаётся положить  $b_n = \lambda_n$  и  $c_n = \lambda_{n-1}$ .  $\square$

**Замечание.** Удобно считать, что  $\varphi_{-1}(x) = 0$ . Тогда рекуррентное соотношение из теоремы 2.6.1.2 будет выполняться и при  $n = 0$ .

## 2.6.2. Многочлены Лежандра

Многочлены  $P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$  называют *многочленами Лежандра*. Следующее утверждение показывает, что многочлены Лежандра образуют ортогональную систему относительно скалярного произведения  $(f, g) = \int_{-1}^1 f(x)g(x) dx$ .

**Теорема 2.6.2.1.**  $\int_{-1}^1 P_m(x)P_n(x) dx = \frac{2}{2n+1} \delta_{mn}$ .

**Доказательство.** Интегрирование по частям показывает, что

$$\int_{-1}^1 \frac{d^m}{dx^m} (x^2 - 1)^m \frac{d^n}{dx^n} (x^2 - 1)^n dx = (-1)^m \int_{-1}^1 (x^2 - 1)^m \frac{d^{n+m}}{dx^{n+m}} (x^2 - 1)^n dx,$$

поскольку  $\frac{d^k}{dx^k} (x^2 - 1)^m$  обращается в нуль, когда  $x = \pm 1$  и  $k \leq m - 1$ . Можно считать, что  $m \geq n$ . Если  $m > n$ , то  $\frac{d^{n+m}}{dx^{n+m}} (x^2 - 1)^n = 0$ . Остаётся разобрать случай, когда  $m = n$ . В этом случае

$$\begin{aligned} \int_{-1}^1 \left( \frac{d^n}{dx^n} (x^2 - 1) \right)^2 dx &= (-1)^n (2n)! \int_{-1}^1 (x^2 - 1)^n dx = \\ &= 2(2n)! \int_0^1 (x^2 - 1)^n dx = \\ &= 2(2n)! \int_0^{\pi/2} \sin^{2n+1} \theta d\theta; \end{aligned}$$

мы сделали замену  $x = \cos \theta$  и поменяли местами верхний и нижний пределы интегрирования. Пусть  $U_n = \int_0^{\pi/2} \sin^{2n+1} \theta d\theta$ . Интегрируя по частям, получаем

$$\begin{aligned} U_n &= - \int_0^{\pi/2} \sin^{2n} \theta d(\cos \theta) = \\ &= - \sin^{2n} \theta \cos \theta \Big|_0^{\pi/2} + \int_0^{\pi/2} \cos \theta d(\sin^{2n} \theta) = \\ &= 2n \int_0^{\pi/2} \sin^{2n-1} \theta \cos^2 \theta d\theta = \\ &= 2n U_{n-1} - 2n U_n, \end{aligned}$$

т.е.  $U_n = \frac{2n}{2n+1} U_{n-1}$ . Учитывая, что  $U_0 = 1$ , получаем  $U_n = \frac{2 \cdot 4 \cdot \dots \cdot (2n)}{3 \cdot 5 \cdot \dots \cdot (2n+1)}$ . Таким образом,

$$\int_{-1}^1 (P_n(x))^2 dx = \frac{1}{2^{2n}(n!)^2} \cdot 2(2n)! \cdot \frac{2 \cdot 4 \cdot \dots \cdot (2n)}{3 \cdot 5 \cdot \dots \cdot (2n+1)} = \frac{2}{2n+1},$$

поскольку  $2 \cdot 4 \cdot \dots \cdot (2n) = 2^n n!$  и  $2^n n! (3 \cdot 5 \cdot \dots \cdot (2n+1)) = (2n+1)!$ .  $\square$

**Теорема 2.6.2.2.** Любые три последовательных многочлена Лежандра связаны соотношением

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x), \quad n \geq 1.$$

**Доказательство.** Легко видеть, что  $P_n(x) = k_n x^n + \dots$ , где  $k_n = \frac{(2n)!}{2^n(n!)^2}$ . Поэтому согласно теореме 2.6.1.2  $P_{n+1}(x) - a_n x P_n(x) = b_n P_n(x) + c_n P_{n-1}(x)$ , где  $a_n = \frac{k_{n+1}}{k_n} = \frac{2n+1}{n+1}$ . Чтобы найти константы  $b_n$  и  $c_n$ , мы воспользуемся тем, что  $P_n(1) = 1$  и  $P_n(-1) = (-1)^n$ . Эти равенства следуют из того, что

$$\begin{aligned} P_n(x) &= \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x-1)^n (x+1)^n) = \\ &= \frac{1}{2^n n!} \sum_{k=0}^n \binom{n}{k} \frac{n!}{(n-k)!} (x-1)^{n-k} \frac{n!}{k!} (x+1)^k. \end{aligned}$$

Для  $b_n$  и  $c_n$  при  $x = \pm 1$  получаем уравнения  $1 - \frac{2n+1}{n+1} = b_n + c_n$ ,  $1 - \frac{2n+1}{n+1} = -b_n + c_n$ , откуда  $b_n = 0$  и  $c_n = -\frac{n}{n+1}$ .  $\square$

### 2.6.3. Многочлены Эрмита

Многочлены  $H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$  называют *многочленами Эрмита*. Прежде всего заметим, что  $H_n(x)$  — многочлен степени  $n$ . Действительно,  $(f(x)e^{-x^2})' = f'(x)e^{-x^2} - 2xf(x)e^{-x^2}$ , поэтому  $H_{n+1}(x) = -H_n'(x) + 2xH_n(x)$ . Остаётся заметить, что  $H_0(x) = 1$ . Следующее утверждение показывает, что многочлены Эрмита образуют ортогональную систему относительно скалярного произведения  $(f, g) = \int_{-\infty}^{\infty} e^{-x^2} f(x)g(x) dx$ .

**Теорема 2.6.3.1.**  $\int_{-\infty}^{\infty} e^{-x^2} H_n(x)H_m(x) dx = 2^n n! \sqrt{\pi} \delta_{mn}$ .

**Доказательство.** Можно считать, что  $m \leq n$ . По определению

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x)H_m(x) dx = (-1)^n \int_{-\infty}^{\infty} H_m(x)\varphi^{(n)}(x) dx,$$

где  $\varphi(x) = e^{-x^2}$ . Интегрируя по частям, получаем

$$\begin{aligned} (-1)^n \int_{-\infty}^{\infty} H_m(x)\varphi^{(n)}(x) dx &= \\ &= (-1)^n H_m(x)\varphi^{(n-1)}(x) \Big|_{-\infty}^{\infty} + (-1)^{n+1} \int_{-\infty}^{\infty} H_m'(x)\varphi^{(n-1)}(x) dx. \end{aligned}$$

Ясно, что  $H_m(x)\varphi^{(n-1)}(x)$  представляет собой произведение многочлена на  $e^{-x^2}$ , поэтому при  $x = \pm\infty$  это выражение обращается в нуль. Повторяя интегрирование по частям, приходим к



$(-1)^{n+m} \int_{-\infty}^{\infty} H_m^{(m)}(x) \varphi^{(n-m)}(x) dx$ . Если  $m < n$ , то можно ещё раз проинтегрировать по частям, и мы получим 0, поскольку  $H_m^{(m+1)}(x) = 0$ . Если  $m = n$ , то мы получаем  $n! 2^n \int_{-\infty}^{\infty} e^{-x^2} dx = n! 2^n \sqrt{\pi}$ , поскольку  $H_m^{(m)}(x) = (2^n x^n + \dots)^{(n)} = n! 2^n$ .  $\square$

**Теорема 2.6.3.2.** Любые три последовательных многочлена Эрмита связаны соотношением

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \quad n \geq 1.$$

**Доказательство.** Пусть  $\varphi(x) = e^{-x^2}$ . Тогда  $\varphi'(x) = -2xe^{-x^2}$ , поэтому  $\varphi'(x) + 2x\varphi(x) = 0$ . Дифференцируя это равенство, получаем  $\varphi''(x) + 2x\varphi'(x) + 2\varphi(x) = 0$ , затем  $\varphi'''(x) + 2x\varphi''(x) + 4\varphi'(x) = 0$  и т.д. Индукцией по  $n$  легко доказать, что соответствующее равенство имеет вид  $\varphi^{(n+1)}(x) + 2x\varphi^{(n)}(x) + 2n\varphi^{(n-1)}(x) = 0$ , что эквивалентно требуемому равенству.  $\square$

**Следствие.** Для любого натурального  $n$  выполняется равенство  $H_n'(x) = 2nH_{n-1}(x)$ .

**Доказательство.** Достаточно сравнить полученное вначале рекуррентное соотношение  $H_{n+1}(x) = -H_n'(x) + 2xH_n(x)$  с соотношением из теоремы.  $\square$

#### 2.6.4. Многочлены Чебышева

Многочленом Чебышева называют многочлен  $T_n(x)$ , для которого  $T_n(\cos \varphi) = \cos n\varphi$  для всех  $\varphi$ . Формула

$$\cos(n+1)\varphi + \cos(n-1)\varphi = 2 \cos n\varphi \cos \varphi$$

показывает, что  $T_n(x)$  действительно является многочленом (степени  $n$ ), причём семейство этих многочленов определяется рекуррентным соотношением

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

и начальными условиями  $T_0(x) = 1$  и  $T_1(x) = x$ .

Следующее утверждение показывает, что многочлены Чебышева образуют ортогональную систему относительно скалярного произведения  $(f, g) = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x)g(x) dx$ .

**Теорема 2.6.4.1.**  $\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) dx = \frac{\pi}{2} \delta_{mn}$ , если одно из чисел  $n$  и  $m$  больше нуля;  $\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_0(x) T_0(x) dx = \pi$ .

**Доказательство.** Сделав замену  $x = \cos \varphi$ , получим

$$\begin{aligned} \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) dx &= \int_0^\pi \cos n\varphi \sin m\varphi d\varphi = \\ &= \int_0^\pi \frac{\cos(n+m)\varphi + \cos(n-m)\varphi}{2} d\varphi. \end{aligned}$$

Если  $n \neq m$ , то мы получаем 0, поскольку  $\int_0^\pi \cos k\varphi = 0$  при  $k \neq 0$ . Если же  $n = m \neq 0$ , то мы получаем  $\int_0^\pi \frac{d\varphi}{2} = \frac{\pi}{2}$ .  $\square$

### Задачи

**6.1.** а) Докажите, что многочлен Лежандра  $P_n(x)$  равен определителю матрицы

$$\begin{pmatrix} x & 1 & 0 & \dots & 0 & 0 \\ \frac{1}{2} & \frac{3}{2}x & 1 & \dots & 0 & 0 \\ 0 & \frac{5}{2} & \frac{5}{2}x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{2n-3}{n-1}x & 1 \\ 0 & 0 & 0 & \dots & \frac{n-1}{n} & \frac{2n-1}{n}x \end{pmatrix}.$$

б) Докажите, что многочлен Эрмита  $H_n(x)$  равен определителю матрицы

$$\begin{pmatrix} 2x & 1 & 0 & \dots & 0 & 0 \\ 2 & 2x & 2 & \dots & 0 & 0 \\ 0 & 2 & 2x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2x & n-1 \\ 0 & 0 & 0 & \dots & 2 & 2x \end{pmatrix}.$$

в) Докажите, что многочлен Чебышева  $T_n(x)$  равен определителю матрицы порядка  $n$

$$\begin{pmatrix} x & 1 & 0 & \dots & 0 & 0 \\ 1 & 2x & 1 & \dots & 0 & 0 \\ 0 & 1 & 2x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2x & 1 \\ 0 & 0 & 0 & \dots & 1 & 2x \end{pmatrix}.$$

## 2.7. Комплексификация и овеществление. Эрмитовы пространства

### 2.7.1. Комплексификация

Комплексификацией линейного пространства  $V$  над  $\mathbb{R}$  называют множество пар  $(a, b)$ , где  $a, b \in V$ , на котором введена структура линейного пространства над  $\mathbb{C}$  следующим образом:

$$\begin{aligned}(a, b) + (a_1, b_1) &= (a + a_1, b + b_1), \\ (x + iy)(a, b) &= (xa - yb, xb + ya).\end{aligned}$$

Такие пары векторов можно записывать в виде  $a + ib$ . Комплексификацию пространства  $V$  будем обозначать  $V^{\mathbb{C}}$ ; подпространство, состоящее из пар  $(a, 0)$ , называют вещественным подпространством  $V^{\mathbb{C}}$  или вещественной формой  $V^{\mathbb{C}}$ .

Оператору  $A : V \rightarrow V$  можно сопоставить оператор  $A^{\mathbb{C}} : V^{\mathbb{C}} \rightarrow V^{\mathbb{C}}$ , заданный формулой  $A^{\mathbb{C}}(a + ib) = Aa + iAb$ . Оператор  $A^{\mathbb{C}}$  называют комплексификацией оператора  $A$ . Если  $e_1, \dots, e_n$  — базис пространства  $V$ , то  $e_1, \dots, e_n$  — базис (над  $\mathbb{C}$ ) пространства  $V^{\mathbb{C}}$ ; в этих базисах матрицы операторов  $A$  и  $A^{\mathbb{C}}$  совпадают.

Пространство  $V^{\mathbb{C}}$  с фиксированным вещественным подпространством  $V$  называют вещественно-комплексным пространством. В этом пространстве определена операция сопряжения, переводящая вектор  $v + iw$  в вектор  $\overline{v + iw} = v - iw$  ( $v, w \in V$ ). Пространство  $V^{\mathbb{C}}$  является линейным пространством размерности  $2 \dim V$  над  $\mathbb{R}$ . Операция сопряжения является линейным преобразованием этого пространства над  $\mathbb{R}$ , но не над  $\mathbb{C}$ . Точки пространства  $V$  (и только они) остаются при сопряжении неподвижными.

### 2.7.2. Овеществление

Линейное пространство  $V$  над  $\mathbb{C}$  является также и линейным пространством над  $\mathbb{R}$  (элементы  $V$  можно умножать на все комплексные числа, поэтому их заведомо можно умножать на вещественные числа). Полученное пространство над  $\mathbb{R}$  называют овеществлением  $V$ ; будем обозначать его  $V_{\mathbb{R}}$ .

Линейное отображение  $A : V \rightarrow W$  можно рассматривать как линейное отображение  $A_{\mathbb{R}} : V_{\mathbb{R}} \rightarrow W_{\mathbb{R}}$  над  $\mathbb{R}$ . Отображение  $A_{\mathbb{R}}$  называют овеществлением отображения  $A$ .

Если  $e_1, \dots, e_n$  — базис  $V$  над  $\mathbb{C}$ , то  $e_1, \dots, e_n, ie_1, \dots, ie_n$  — базис  $V_{\mathbb{R}}$ . Легко проверить, что если  $A = B + iC$  — матрица линейного отображения  $A : V \rightarrow W$  относительно базисов  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_m$ , причём матрицы  $B$  и  $C$  вещественны, то матрица линейного отображения  $A_{\mathbb{R}}$  относительно базисов  $e_1, \dots, e_n, ie_1, \dots, ie_n$  и  $\varepsilon_1, \dots, \varepsilon_m, i\varepsilon_1, \dots, i\varepsilon_m$  имеет вид  $\begin{pmatrix} B & -C \\ C & B \end{pmatrix}$ .

**Теорема 2.7.2.1.** Если  $A : V \rightarrow V$  — линейное отображение над  $\mathbb{C}$ , то  $\det A_{\mathbb{R}} = |\det A|^2$ .

**Доказательство.**  $\begin{vmatrix} B & -C \\ C & B \end{vmatrix} = \begin{vmatrix} B+iC & -C+iB \\ C & B \end{vmatrix} = \begin{vmatrix} B+iC & 0 \\ C & B-iC \end{vmatrix} = |B+iC| \cdot |B-iC| = |\det A|^2$ .  $\square$

Овеществление  $V_{\mathbb{R}}$  линейного пространства  $V$  над  $\mathbb{C}$  имеет каноническую ориентацию, заданную базисом  $e_1, \dots, e_n, ie_1, \dots, ie_n$ , где  $e_1, \dots, e_n$  — некоторый базис  $V$  над  $\mathbb{C}$ . Эта ориентация не зависит от выбора базиса  $e_1, \dots, e_n$ , поскольку согласно теореме 2.7.2.1 для любого линейного над  $\mathbb{C}$  отображения  $A$  определитель отображения  $A : V \rightarrow V$  положителен.

### 2.7.3. Эрмитовы пространства

Пусть  $V$  — линейное пространство над  $\mathbb{C}$ . Эрмитовым произведением в  $V$  называют отображение  $V \times V \rightarrow \mathbb{C}$ , которое сопоставляет каждой паре векторов  $x, y \in V$  комплексное число  $(x, y)$  и обладает следующими свойствами:

- 1)  $(x, y) = \overline{(y, x)}$ ;
- 2)  $(\alpha x + \beta y, z) = (\alpha x, z) + (\beta y, z)$ ;
- 3)  $(x, x)$  — действительное положительное число для любого  $x \neq 0$ .

Пространство  $V$ , в котором задано эрмитово произведение, называют эрмитовым. Стандартное эрмитово произведение в  $\mathbb{C}^n$  имеет вид  $x_1 \bar{y}_1 + \dots + x_n \bar{y}_n$ .

Базис  $e_1, \dots, e_n$  эрмитова пространства  $V$  называют ортонормированным, если  $(e_i, e_j) = \delta_{ij}$ . Пусть  $\|a_{ij}\|_1^n$  — матрица оператора  $A$  относительно ортонормированного базиса  $e_1, \dots, e_n$ , т.е.  $Ae_j = \sum_i a_{ij} e_i$ . Тогда  $a_{ij} = (Ae_j, e_i)$ . В частности, значения  $(Ax, y)$  однозначно определяют оператор  $A$ .

Если  $V$  — эрмитово пространство, то его овеществление  $V_{\mathbb{R}}$  можно снабдить скалярным произведением

$$(x, y)_{\mathbb{R}} = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2),$$

где  $\|x\|^2 = (x, x)_{\mathbb{C}}$  — квадрат длины вектора относительно эрмитова произведения.

Например, стандартному эрмитову произведению в  $\mathbb{C}^n$  соответствует скалярное произведение

$$(x, y)_{\mathbb{R}} = \frac{1}{2} \sum_k (x_k \bar{y}_k + \bar{x}_k y_k) = \operatorname{Re} \sum_k x_k \bar{y}_k.$$

Если  $x_k = a_k + ib_k$  и  $y_k = c_k + id_k$ , где  $a_k, b_k, c_k, d_k \in \mathbb{R}$ , то  $(x, y)_{\mathbb{R}} = \sum_k (a_k c_k + b_k d_k)$ .

**Теорема 2.7.3.1.** *Скалярное произведение  $(x, y)_{\mathbb{R}}$  обладает следующими свойствами:*

- а)  $(x, ix)_{\mathbb{R}} = 0$ ;
- б)  $(x, iy)_{\mathbb{R}} + (ix, y)_{\mathbb{R}} = 0$ ;
- в)  $(x, y)_{\mathbb{R}}^2 + (ix, y)_{\mathbb{R}}^2 \leq \|x\|^2 \|y\|^2$ .

**Доказательство.** а) Ясно, что  $(ix, x)_{\mathbb{R}} + (x, ix)_{\mathbb{R}} = (ix, x)_{\mathbb{R}} + \overline{(ix, x)_{\mathbb{R}}} = 0$ . Поэтому  $\|x + ix\|^2 = \|x\|^2 + \|ix\|^2 = 2\|x\|^2$ .

б) Достаточно воспользоваться равенством  $\|x + iy\|^2 = \|ix - y\|^2$ .

в) Можно считать, что длина вектора  $x$  равна 1. Тогда длина вектора  $ix$  тоже равна 1. Согласно а) эти векторы ортогональны. Дополним векторы  $x$  и  $ix$  до ортонормированного базиса. Пусть  $(y_1, y_2, \dots)$  — координаты вектора  $y$  относительно этого базиса. Тогда  $y_1^2 + y_2^2 \leq \|y\|^2$ .  $\square$

#### 2.7.4. Унитарные, эрмитовы и косоэрмитовы операторы

Линейный оператор  $A^*$  называют *сопряжённым* с  $A$ , если  $(Ax, y) = (x, A^*y) = \overline{(A^*y, x)}$ . Пусть  $\|a_{ij}\|_1^n$  и  $\|b_{ij}\|_1^n$  — матрицы операторов  $A$  и  $A^*$  относительно ортонормированного базиса. Тогда  $a_{ij} = (Ae_j, e_i) = \overline{(A^*e_i, e_j)} = \bar{b}_{ij}$ . Матрицу  $X^* = \bar{X}^T$  называют *сопряжённой* с  $X$  (она получается из  $X$  комплексным сопряжением всех элементов и транспонированием).

Линейный оператор  $A$  называют *унитарным*, если  $(Ax, Ay) = (x, y)$ , т.е. унитарный оператор сохраняет эрмитово произведение. Если оператор  $A$  унитарен, то  $(x, y) = (Ax, Ay) = (x, A^*Ay)$ . Поэтому  $A^*A = I =$

$AA^*$ , т.е. строки и столбцы матрицы  $A$  образуют ортонормированную систему векторов.

Линейный оператор  $A$  называют *эрмитовым* (соответственно *косоэрмитовым*), если  $A^* = A$  (соответственно  $A^* = -A$ ). Линейный оператор эрмитов или косоэрмитов тогда и только тогда, когда его матрица  $A$  относительно ортонормированного базиса эрмитова, т.е.  $A^* = A$ , или косоэрмитова, т.е.  $A^* = -A$ ; в этом случае его матрица относительно любого ортонормированного базиса эрмитова или косоэрмитова.

Аналогом вещественных симметрических матриц в комплексном случае являются, как правило, эрмитовы матрицы. Иногда рассматривают также *комплексные симметрические* или *кососимметрические* матрицы, для которых  $A^T = A$  или  $A^T = -A$ .

**Теорема 2.7.4.1.** Пусть  $A$  — комплексный оператор, причём  $(Ax, x) \equiv 0$ . Тогда  $A = 0$ .

**Доказательство.** Запишем равенство  $(Ax, x) = 0$  для  $x = u + v$  и  $x = u + iv$ . Учитывая, что  $(Av, v) = (Au, u) = 0$ , получим  $(Av, u) + (Au, v) = 0$  и  $i(Av, u) - i(Au, v) = (Aiv, u) + (Au, iv) = 0$ . Следовательно,  $(Au, v) = 0$  для любых  $u, v \in V$ .  $\square$

**Замечание.** В вещественном случае тождество  $(Ax, x) \equiv 0$  означает, что  $A$  — кососимметрический оператор (теорема 4.3.1.2).

**Теорема 2.7.4.2.** Пусть  $A$  — комплексный оператор, причём  $(Ax, x) \in \mathbb{R}$  для любого  $x$ . Тогда  $A$  — эрмитов оператор.

**Доказательство.** Так как  $(Ax, x) = \overline{(Ax, x)} = (x, Ax)$ , то  $((A - A^*)x, x) = (Ax, x) - (A^*x, x) = (Ax, x) - (x, Ax) = 0$ . Согласно теореме 2.7.4.1  $A - A^* = 0$ .  $\square$

**Теорема 2.7.4.3.** Любой комплексный оператор однозначно представим в виде  $A = B + iC$ , где  $B$  и  $C$  — эрмитовы операторы.

**Доказательство.** Если  $A = B + iC$ , где  $B$  и  $C$  — эрмитовы операторы, то  $A^* = B^* - iC^* = B - iC$ , поэтому  $2B = A + A^*$  и  $2iC = A - A^*$ . Легко проверить, что операторы  $(A + A^*)/2$  и  $(A - A^*)/2i$  эрмитовы.  $\square$

**Замечание.** Оператор  $iC$  косоэрмитов тогда и только тогда, когда оператор  $C$  эрмитов, поэтому любой оператор  $A$  однозначно представим в виде суммы эрмитова и косоэрмитова оператора.

### 2.7.5. Нормальные операторы

Оператор  $A$  называют *нормальным*, если  $A^*A = AA^*$ . Легко проверить, что унитарные, эрмитовы и косоэрмитовы операторы нормальны.

**Теорема 2.7.5.1.** Оператор  $A = B + iC$ , где  $B$  и  $C$  — эрмитовы операторы, нормален тогда и только тогда, когда  $BC = CB$ .

**Доказательство.** Так как  $A^* = B^* - iC^* = B - iC$ , то  $A^*A = B^2 + C^2 + i(BC - CB)$  и  $AA^* = B^2 + C^2 - i(BC - CB)$ . Поэтому равенство  $A^*A = AA^*$  эквивалентно равенству  $BC - CB = 0$ .  $\square$

### 2.7.6. Комплексные структуры

Если  $V$  — линейное пространство над  $\mathbb{R}$ , то для задания на  $V$  структуры линейного пространства над  $\mathbb{C}$  необходимо задать операцию  $J$  умножения на  $i$ , т.е.  $Jv = iv$ . Это линейное отображение  $J : V \rightarrow V$  должно обладать следующим свойством:  $J^2v = i(iv) = -v$ , т.е.  $J^2 = -I$ . Ясно также, что если в пространстве  $V$  над  $\mathbb{R}$  задан такой линейный оператор  $J$ , то  $V$  можно превратить в пространство над  $\mathbb{C}$ , определив умножение на комплексные числа формулой  $(a + ib)v = av + bJv$ . В частности, пространство  $V$  в этом случае имеет чётную размерность.

**Определение.** Пусть  $V$  — линейное пространство над  $\mathbb{R}$ . Линейный оператор  $J : V \rightarrow V$  называют комплексной структурой, если  $J^2 = -I$ .

Выберем в пространстве  $\mathbb{R}^{2n}$  базис и обозначим базисные векторы  $e_1, \dots, e_n, ie_1, \dots, ie_n$ . Отображение  $J : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ , заданное формулами  $Je_k = ie_k$  и  $J(ie_k) = -e_k$ , является комплексной структурой. Относительно базиса  $e_1, \dots, e_n, ie_1, \dots, ie_n$  матрица комплексной структуры  $J$  имеет вид  $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ .

**Теорема 2.7.6.1.** Для любой комплексной структуры  $J$  в пространстве  $\mathbb{R}^{2n}$  можно выбрать базис  $e_1, \dots, e_{2n}$  так, чтобы матрица комплексной структуры  $J$  относительно этого базиса имела вид  $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ .

**Доказательство.** Прежде всего заметим, что у оператора комплексной структуры нет собственных векторов, поскольку если  $Jv = \lambda v$ , то  $-v = J^2v = \lambda^2v$ . Выберем в качестве  $e_1$  любой ненулевой вектор и положим  $e_{2n+1} = Je_1$ . Если  $k < n$  и мы уже выбрали векторы  $e_1, \dots, e_k$  и  $e_{n+1}$ ,

...,  $e_{n+k}$ , то в качестве  $e_{k+1}$  выбираем любой вектор, не принадлежащий пространству  $V_k$ , порождённому выбранными векторами, и полагаем  $e_{n+k+1} = Je_{k+1}$ .

Проверим, что если вектор  $v = e_{k+1}$  не принадлежит пространству  $V_k$  (инвариантному относительно оператора комплексной структуры  $J$ ), то пространство  $V_k$  пересекает пространство, порождённое векторами  $v$  и  $Jv$ , только по нулю. Если  $\lambda v + \mu Jv \in V_k$ , то  $J(\lambda v + \mu Jv) = \lambda Jv - \mu v \in V_k$ , поэтому

$$\lambda(\lambda v + \mu Jv) - \mu(\lambda Jv - \mu v) = (\lambda^2 + \mu^2)v \in V_k.$$

□

Оператор  $A$  в  $\mathbb{R}^{2n}$  является комплексно линейным оператором тогда и только тогда, когда  $A(iv) = iA(v)$  для любого вектора  $v$ , т.е.  $AJ = JA$ . Легко видеть, что это условие выполняется тогда и только тогда, когда  $A = \begin{pmatrix} B & -C \\ C & B \end{pmatrix}$ , где  $B$  и  $C$  — некоторые вещественные матрицы порядка  $n$ .

Если  $A$  — невырожденное преобразование  $\mathbb{R}^{2n}$ , то для базиса  $Ae_1, \dots, Ae_n, A(ie_1), \dots, A(ie_n)$  аналогично можно определить комплексную структуру  $J_A$ . Комплексные структуры  $J$  и  $J_A$  совпадают тогда и только тогда, когда оператор  $A$  является комплексно линейным оператором в пространстве  $\mathbb{R}^{2n}$ , на котором структура комплексного пространства определена оператором  $J$ .

Собственные значения оператора  $J : V \rightarrow V$  чисто мнимые, поэтому для более детального изучения этого оператора рассмотрим комплексификацию  $V^{\mathbb{C}}$  пространства  $V$ . Следует отметить, что операция умножения на  $i$  в  $V^{\mathbb{C}}$  никак не связана с комплексной структурой  $J$  и её комплексификацией  $J^{\mathbb{C}}$ .

**Теорема 2.7.6.2.**  $V^{\mathbb{C}} = V_+ \oplus V_-$ , где  $V_+ = \text{Ker}(J^{\mathbb{C}} - iI) = \text{Im}(J^{\mathbb{C}} + iI)$  и  $V_- = \text{Ker}(J^{\mathbb{C}} + iI) = \text{Im}(J^{\mathbb{C}} - iI)$ .

**Доказательство.** Так как  $(J^{\mathbb{C}} - iI)(J^{\mathbb{C}} + iI) = (J^{\mathbb{C}})^2 + I = 0$ , то  $\text{Im}(J^{\mathbb{C}} + iI) \subset \text{Ker}(J^{\mathbb{C}} - iI)$ . Аналогично  $\text{Im}(J^{\mathbb{C}} - iI) \subset \text{Ker}(J^{\mathbb{C}} + iI)$ . С другой стороны,  $-i(J^{\mathbb{C}} + iI) + i(J^{\mathbb{C}} - iI) = 2I$ , поэтому  $V^{\mathbb{C}} \subset \text{Im}(J^{\mathbb{C}} + iI) + \text{Im}(J^{\mathbb{C}} - iI)$ . Учитывая, что  $\text{Ker}(J^{\mathbb{C}} - iI) \cap \text{Ker}(J^{\mathbb{C}} + iI) = 0$ , получаем требуемое. □

**Замечание.** Ясно, что  $V_+ = \bar{V}_-$ .

**Теорема 2.7.6.3.** Если  $J$  — оператор комплексной структуры, то  $\det J = 1$ .



**Первое доказательство.** Если мы рассмотрим  $J$  как оператор в  $\mathbb{C}^{2n}$ , то согласно теореме 2.7.6.2 этот оператор в некотором базисе имеет диагональный вид с элементами  $\pm i$  на диагонали, причём элементов  $i$  и  $-i$  поровну. Определитель такого оператора равен 1. Ясно также, что определитель матрицы с вещественными элементами не изменится, если мы будем рассматривать её как матрицу с комплексными элементами.  $\square$

**Второе доказательство.** Согласно теореме 2.7.6.1 в некотором базисе матрица оператора  $J$  имеет вид  $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ . Несложно проверить, что определитель такой матрицы равен 1.  $\square$

Пусть  $J$  — оператор комплексной структуры в вещественном пространстве  $V$ . Тогда в пространстве  $V$  оператор  $J$  задаёт ориентацию, соответствующую канонической ориентации комплексного пространства. А именно, согласно теореме 2.7.6.1 в некотором базисе матрица оператора  $J$  имеет вид  $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ . Ориентация этого базиса — это и есть каноническая ориентация комплексного пространства. Она не зависит от выбора базиса.

**Замечание.** Если в пространстве  $V$  выбрана ориентация, то ориентация, которую задаёт оператор комплексной структуры  $J$ , может не совпадать с выбранной ориентацией. При этом оператор  $J$  всегда сохраняет ориентацию пространства  $V$ , поскольку согласно теореме 2.7.6.3 его определитель равен 1.

### 2.7.7. Комплексно сопряжённое пространство

Пусть  $V$  — линейное пространство над  $\mathbb{C}$ . Рассмотрим пространство  $\overline{V}$ , совпадающее как множество и как группа по сложению с  $V$ , а произведение  $\lambda$  на  $v$  в  $\overline{V}$  равно произведению  $\bar{\lambda}$  на  $v$  в  $V$ . Другими словами, если элемент пространства  $\overline{V}$ , соответствующий элементу  $v \in V$ , обозначить  $\bar{v}$ , то  $\lambda \bar{v} = \bar{\lambda} v$ . Пространство  $\overline{V}$  мы назовём *комплексно сопряжённым* с  $V$ .

Легко проверить, что  $\overline{V}$  — линейное пространство над  $\mathbb{C}$ . Оно, конечно, изоморфно  $V$ , но изоморфизм не канонический. Ясно также, что о вещественности пространств  $V$  и  $\overline{V}$  канонически изоморфны.

Отображение  $f : V \rightarrow \overline{V}$  линейно над  $\mathbb{C}$  тогда и только тогда, когда при естественном отождествлении  $V$  с  $\overline{V}$  получается *антилинейное* отображение, т.е.  $f(\lambda v) = \bar{\lambda} f(v)$  и  $f(v + w) = f(v) + f(w)$ . Действительно,

пусть  $f(v) = \bar{w}$  и  $f(\lambda v) = \bar{u}$ . Условие линейности заключается в том, что  $\bar{u} = \lambda f(v) = \lambda \bar{w} = \bar{\lambda w}$ , т.е.  $u = \bar{\lambda w}$ . При естественном отождествлении  $V$  с  $\bar{V}$  это равенство запишется в виде  $f(\lambda v) = \bar{\lambda} f(v)$ .

**Теорема 2.7.7.1.** *Пространство  $(V_{\mathbb{R}})^{\mathbb{C}}$  канонически изоморфно  $V \oplus \bar{V}$ .*

**Доказательство.** Овеществления пространств  $(V_{\mathbb{R}})^{\mathbb{C}}$  и  $V \oplus \bar{V}$  канонически изоморфны, так как они изоморфны пространству  $W$ , состоящему из пар  $(x, y)$ , где  $x, y \in V$ . В пространстве  $W$  определена комплексная структура  $J_1(x, y) = (-y, x)$ , соответствующая комплексной структуре пространства  $(V_{\mathbb{R}})^{\mathbb{C}}$ , и комплексная структура  $J_2(x, y) = (ix, iy)$ , соответствующая комплексной структуре пространства  $V$ . Легко проверить, что  $W = \text{Ker}(J_1 - J_2) \oplus \text{Ker}(J_1 + J_2)$ . В самом деле,  $\text{Ker}(J_1 - J_2)$  состоит из пар  $(a, -ia)$ ,  $\text{Ker}(J_1 + J_2)$  состоит из пар  $(b, ib)$ , поэтому  $\text{Ker}(J_1 - J_2) \cap \text{Ker}(J_1 + J_2) = 0$ ; кроме того, любой элемент  $(x, y) \in W$  можно представить в виде суммы элементов  $(a, -ia)$  и  $(b, ib)$ , где  $a = (x + iy)/2$  и  $b = (x - iy)/2$ . Так как  $J_1(a, -ia) = (ia, a)$  и  $J_1(b, ib) = (-ib, b)$ , пространства  $\text{Ker}(J_1 - J_2)$  и  $\text{Ker}(J_1 + J_2)$  канонически изоморфны  $V$  и  $\bar{V}$  соответственно.  $\square$

**Теорема 2.7.7.2.** *Канонические ориентации овеществлений пространств  $V$  и  $\bar{V}$  совпадают тогда и только тогда, когда размерности пространств  $V$  и  $\bar{V}$  над  $\mathbb{C}$  чётны.*

**Доказательство.** Заметим, что  $i\bar{v} = \overline{\bar{i}v} = \overline{-iv}$ , и применим естественное отождествление пространств  $V$  и  $\bar{V}$ . Тогда каноническая ориентация овеществления  $V$  задаётся базисом  $e_1, \dots, e_n, ie_1, \dots, ie_n$ , а каноническая ориентация овеществления  $\bar{V}$  — базисом  $e_1, \dots, e_n, -ie_1, \dots, -ie_n$ . Определитель матрицы перехода от одного базиса к другому равен  $(-1)^n$ .  $\square$

## 2.7.8. Метрика Фубини—Штуди

Пусть  $x$  и  $y$  — ненулевые векторы в эрмитовом пространстве  $V$ . Угол  $\theta$ ,  $0 \leq \theta \leq \pi/2$ , для которого  $\cos \theta = \frac{|(x, y)|}{\|x\| \cdot \|y\|}$ , называют *эрмитовым углом* между векторами  $x$  и  $y$ . Эрмитов угол между комплексными прямыми  $l_x = \{\lambda x | \lambda \in \mathbb{C}\}$  и  $l_y = \{\lambda y | \lambda \in \mathbb{C}\}$  определяется как эрмитов угол между любой парой ненулевых векторов, принадлежащих этим прямым.

В комплексном проективном пространстве, точками которого служат комплексные прямые, метрика Фубини—Штуди определяется следующим образом: расстояние между комплексными прямыми равно эрмитову углу между ними.

**Теорема 2.7.8.1.** Метрика Фубини–Штуди задаётся формулой

$$ds^2 = \frac{1}{\|x\|^4} (\|x\|^2 \|dx\|^2 - (x, d\bar{x})(\bar{x}, dx));$$

здесь  $ds$  — расстояние между  $x$  и  $x + dx$  для бесконечно малого  $dx$ .

**Доказательство.** Формула  $\cos ds = 1 - \frac{ds^2}{2} + \dots$  показывает, что

$$\frac{ds^2}{2} = 1 - \frac{|(x, x + dx)|}{\|x\| \cdot \|x + dx\|} = \frac{\|x\| \cdot \|x + dx\| - |(x, x + dx)|}{\|x\| \cdot \|x + dx\|}.$$

Умножим числитель и знаменатель на  $\|x\| \cdot \|x + dx\| + |(x, x + dx)|$ . Полученный в результате этого знаменатель с точностью до бесконечно малой величины равен  $2\|x\|^4$ . Таким образом,

$$ds^2 = \frac{1}{\|x\|^4} (\|x\|^2 \|x + dx\|^2 - |(x, x + dx)|^2).$$

Ясно, что  $\|x + dx\|^2 = \|x\|^2 + (x, d\bar{x}) + (\bar{x}, dx) + \|dx\|^2$  и

$$\begin{aligned} |(x, x + dx)|^2 &= (\bar{x}, x + dx)(x, \bar{x} + d\bar{x}) = \\ &= \|x\|^4 + \|x\|^2 ((x, d\bar{x}) + (\bar{x}, dx)) + \|x\|^2 \|dx\|^2. \end{aligned}$$

После очевидных сокращений получаем требуемое.  $\square$

### 2.7.9. Кэлеров угол

Пусть  $V$  — линейное подпространство над  $\mathbb{C}$ . Подпространство  $U \subset V_{\mathbb{R}}$  называют *голоморфным*, если  $iU = U$ ; здесь  $iU = \{iu | u \in U\}$ .

Пусть  $V$  — эрмитово пространство. Подпространство  $U \subset V_{\mathbb{R}}$  называют *антиголоморфным*, или *вполне вещественным*, если  $iU \perp U$ , где имеется в виду ортогональность в смысле скалярного произведения  $(, )_{\mathbb{R}}$  в  $V_{\mathbb{R}}$ , индуцированного эрмитовым произведением в  $V$ .

Для ненулевых векторов  $x, y$  в эрмитовом пространстве  $V$  можно определить *кэлеров угол*  $\theta_K(x, y)$ ,  $0 \leq \theta_K \leq \pi$ , следующим образом. Пусть  $\angle(x, y)$  — угол между векторами  $x$  и  $y$  в пространстве  $V_{\mathbb{R}}$  со скалярным произведением  $(, )_{\mathbb{R}}$  (точнее говоря, угол поворота от вектора  $x$  к вектору  $y$ ). Тогда

$$\cos \theta_K(x, y) \sin \angle(x, y) = \frac{(ix, y)_{\mathbb{R}}}{\|x\| \cdot \|y\|} = \cos \angle(ix, y).$$

Прежде всего нужно проверить, что  $|\cos \theta_K(x, y)| \leq 1$ , т.е.  $\cos^2 \angle(ix, y) \leq \sin^2 \angle(x, y)$ . Согласно теореме 2.7.3.1 в)  $\cos^2 \angle(x, y) + \cos^2 \angle(ix, y) \leq 1$ . Это неравенство эквивалентно требуемому неравенству.

Ясно, что  $\cos \theta_K(x, y) = \cos \theta_K(y, x)$ , поскольку  $(ix, y)_{\mathbb{R}} = -(x, iy)_{\mathbb{R}}$  и  $\angle(x, y) = -\angle(y, x)$ .

**Теорема 2.7.9.1.** а) Кэлеров угол зависит только от плоскости в  $V_{\mathbb{R}}$ , натянутой на векторы  $x$  и  $y$ .

б) Плоскость голоморфная тогда и только тогда, когда  $\theta_K = 0$  или  $\pi$ .

в) Плоскость антиголоморфная тогда и только тогда, когда  $\theta_K = \pi/2$ .

**Доказательство.** а) Достаточно проверить, что  $\cos \theta_K(x, y) = \cos \theta_K(x, y + tx)$  для любого  $t \in \mathbb{R}$ . Ясно, что

$$\cos^2 \theta_K(x, y) = \frac{\cos^2 \angle(ix, y)}{\sin^2 \angle(x, y)} = \frac{(ix, y)_{\mathbb{R}}^2}{\|x\|^2 \|y\|^2 - (x, y)_{\mathbb{R}}^2}.$$

Учитывая, что  $(ix, x)_{\mathbb{R}} = 0$ , получаем

$$\begin{aligned} (ix, y + tx)_{\mathbb{R}} &= (ix, y)_{\mathbb{R}}, \\ \|x\|^2 \|y + tx\|^2 - (x, y + tx)_{\mathbb{R}}^2 &= \|x\|^2 \|y\|^2 + 2t\|x\|^2 (y, x)_{\mathbb{R}} + t^2 \|x\|^4 - \\ &\quad - (x, y)_{\mathbb{R}}^2 - 2t\|x\|^2 (y, x)_{\mathbb{R}} - t^2 \|x\|^4 = \\ &= \|x\|^2 \|y\|^2 - (x, y)_{\mathbb{R}}^2. \end{aligned}$$

Остаётся заметить, что числа  $\cos \theta_K(x, y)$  и  $\cos \theta_K(x, y + tx)$  имеют одинаковые знаки, поскольку  $(ix, y + tx)_{\mathbb{R}} = (ix, y)_{\mathbb{R}}$  и направление поворота от вектора  $x$  к вектору  $y$  совпадает с направлением поворота от вектора  $x$  к вектору  $y + tx$ .

б) В качестве базиса голоморфной плоскости можно выбрать векторы  $x$  и  $ix$ . Ясно, что  $\sin \angle(x, ix) = 1$  и  $\frac{(ix, ix)_{\mathbb{R}}}{\|x\| \cdot \|ix\|} = 1$ . Поэтому  $\cos \theta_K(x, ix) = 1$ .

Наоборот, если  $\sin \angle(x, y) = \cos \angle(ix, y)$ , то вектор  $y$  лежит в плоскости, натянутой на векторы  $x$  и  $ix$ . Действительно, пусть  $(y_1, y_2, \dots)$  — координаты вектора  $y$  относительно ортонормированного базиса  $x, ix, \dots$ . Можно считать, что  $\|y\| = 1$ . Тогда  $\sin \angle(x, y) = 1 - y_1^2$  и  $\cos \angle(ix, y) = y_2^2$ , поэтому  $y_1^2 + y_2^2 = 1$ , т.е. все остальные координаты вектора  $y$  нулевые.

в) Если плоскость, натянутая на векторы  $x$  и  $y$ , антиголоморфная, то  $(x, iy)_{\mathbb{R}} = 0$ . Поэтому  $\cos \theta_K(x, y) = 0$ .

Наоборот, если  $\cos \theta_K(x, y) = 0$ , то  $(ix, y)_{\mathbb{R}} = 0$ . Поэтому  $(x, iy)_{\mathbb{R}} = -(ix, y)_{\mathbb{R}} = 0$ . Кроме того,  $(ix, x)_{\mathbb{R}} = 0$  и  $(iy, y)_{\mathbb{R}} = 0$ . Таким образом, если  $U$  — плоскость в  $V_{\mathbb{R}}$ , натянутая на векторы  $x$  и  $y$ , то  $iU \perp U$ .  $\square$

Таким образом, кэлеров угол между векторами  $x$  и  $y$ , является мерой отклонения от голоморфности плоскости, натянутой на  $x$  и  $y$ .

## Задачи

7.1. Докажите, что матрица овеществления эрмитова оператора симметрическая.

7.2. Выразите характеристический многочлен матрицы  $A_{\mathbb{R}}$  через характеристический многочлен матрицы  $A$ .

7.3. Докажите, что базисы  $e_1, \dots, e_n, ie_1, \dots, ie_n$  и  $e_1, ie_1, \dots, e_n, ie_n$  в  $V_{\mathbb{R}}$  одинаково ориентированы тогда и только тогда, когда число  $\frac{n(n-1)}{2}$  чётно.

7.4. Пусть  $W$  — комплексное пространство размерности  $n$ ,  $V \subset W$  — некоторое вещественное подпространство размерности  $n$ . Докажите, что  $V$  может быть вещественной формой  $W$  тогда и только тогда, когда  $V \cap iV = 0$ .

7.5. Пусть  $V^{\mathbb{C}} = W \oplus \bar{W}$ , где  $W$  — некоторое комплексное подпространство. Докажите, что в пространстве  $V$  можно выбрать комплексную структуру  $J$  так, что  $W = \text{Ker}(J^{\mathbb{C}} - I)$ .

7.6. Рассмотрим вещественно линейное отображение комплексной плоскости  $\mathbb{C}$  в себя:  $Az = az + b\bar{z}$ , где  $a, b \in \mathbb{C}$ . Докажите, что это отображение вырожденное тогда и только тогда, когда  $|a| = |b|$ .

7.7. Пусть  $V$  — одномерное пространство над  $\mathbb{C}$ . Докажите, что если пространства  $V$  и  $\bar{V}$  естественно отождествлены, то с геометрической точки зрения любой изоморфизм  $f : V \rightarrow \bar{V}$  над  $\mathbb{C}$  представляет собой композицию симметрии относительно прямой и гомотетии  $v \mapsto kv$ .

7.8. В пространстве  $\mathbb{C}^n$  укажите комплексное подпространство размерности  $[n/2]$ , на котором квадратичная форма  $B(x, y) = x_1y_1 + \dots + x_ny_n$  тождественно равна нулю.

7.9. Докажите, что если матрица оператора комплексной структуры ортогональная, то она кососимметрическая.

## 2.8. Нормированные пространства

Линейное пространство  $V$  над  $\mathbb{R}$  называют *нормированным*, если для любого вектора  $v \in V$  задана его *норма*  $\|v\| \in \mathbb{R}$ , обладающая следующими свойствами:

- 1)  $\|v\| \geq 0$  для любого  $v \in V$ ;
- 2)  $\|v\| = 0$  тогда и только тогда, когда  $v = 0$ ;
- 3)  $\|\lambda v\| = |\lambda| \cdot \|v\|$  для любых  $\lambda \in \mathbb{R}$ ,  $v \in V$ ;
- 4)  $\|v + w\| \leq \|v\| + \|w\|$  для любых  $v, w \in V$  (*неравенство треугольника*).

**Пример 2.8.0.1.** Если  $V$  — пространство со скалярным произведением, то в нём можно задать норму, положив  $\|v\| = \sqrt{(v, v)}$ .

**Пример 2.8.0.2.** В пространстве  $\mathbb{R}^n$  можно задать норму, положив  $\|(x_1, \dots, x_n)\| = \sum_{i=1}^n |x_i|$ .

**Пример 2.8.0.3.** В пространстве  $\mathbb{R}^n$  можно задать норму, положив  $\|(x_1, \dots, x_n)\| = \max_{1 \leq i \leq n} |x_i|$ .

**Пример 2.8.0.4.** Пусть  $p > 1$ . В пространстве  $\mathbb{R}^n$  можно задать норму  $l_p$ , положив

$$l_p(x_1, \dots, x_n) = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

**Доказательство.** Свойства 1)–3) для этой нормы очевидны, а свойство 4) — это известное *неравенство Минковского*

$$\left( \sum_{i=1}^n |x_i + y_i|^p \right)^{1/p} \leq \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} + \left( \sum_{i=1}^n |y_i|^p \right)^{1/p}.$$

□

Из свойств нормы следует, что функция  $f(\lambda) = \|\lambda v + w\|$  непрерывно зависит от  $\lambda$ . Действительно,  $\|\lambda v + w\| - |\delta| \cdot \|v\| \leq \|(\lambda + \delta)v + w\| \leq \|\lambda v + w\| + |\delta| \cdot \|v\|$ .

### 2.8.1. Тожество параллелограмма

Нормированное пространство  $V$  называют *пространством со скалярным произведением*, если на  $V$  существует скалярное произведение, для которого  $\|v\| = \sqrt{(v, v)}$  для всех  $v \in V$ .

Если  $V$  — пространство со скалярным произведением, то

$$\begin{aligned}\|u \pm v\|^2 &= (u \pm v, u \pm v) = \\ &= \|u\|^2 + \|v\|^2 \pm 2(u, v),\end{aligned}$$

поэтому

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2. \quad (2.1)$$

Тожество (2.1) называют *тождеством параллелограмма* (с геометрической точки зрения оно выражает то свойство, что сумма квадратов длин сторон параллелограмма равна сумме квадратов длин его сторон).

**Теорема 2.8.1.1 (Йордан—фон Нойман [J2]).** *Нормированное пространство  $V$  является пространством со скалярным произведением тогда и только тогда, когда в нём выполняется тождество параллелограмма.*

**Доказательство.** Если  $V$  — пространство со скалярным произведением  $(u, v)$ , то

$$4(u, v) = \|u + v\|^2 - \|u - v\|^2. \quad (2.2)$$

Равенство (2.2) можно взять в качестве определения скалярного произведения в нормированном пространстве. Нужно проверить, что если выполняется тождество параллелограмма, то в результате мы действительно получим некоторое скалярное произведение.

Запишем тождество параллелограмма для векторов  $u \pm w$  и  $v$ :

$$\|u \pm w + v\|^2 + \|u \pm w - v\|^2 = 2\|u \pm w\|^2 + 2\|v\|^2.$$

Следовательно,

$$\begin{aligned}\|u + w + v\|^2 + \|u + w - v\|^2 - \|u - w + v\|^2 - \|u - w - v\|^2 = \\ = 2\|u + w\|^2 - 2\|u - w\|^2,\end{aligned}$$

т.е.

$$4(u + v, w) + 4(u - v, w) = 8(u, w). \quad (2.3)$$

Положим в равенстве (2.3)  $u = v$ . После сокращений получим  $(2u, v) = 2(u, v)$ ; при этом нужно воспользоваться тем, что  $(0, w) = \|w\|^2 - \| -w\|^2 = 0$ .

Положим в равенстве (2.3)  $u + v = a$  и  $u - v = b$ . Воспользовавшись тождеством  $2(u, v) = 2(u, v)$ , получим

$$(a, w) + (b, w) = 2 \left( \frac{a+b}{2}, w \right) = (a+b, w). \quad (2.4)$$

Мы доказали аддитивность скалярного произведения по первому переменному. Аддитивность по второму переменному доказывается аналогично.

Из равенства (2.4) следует, что  $(na, w) = n(a, w)$  для любого натурального  $n$ . Поэтому  $n \left( \frac{m}{n}a, w \right) = \left( n \frac{m}{n}a, w \right) = m(a, w)$ , а значит,  $\left( \frac{m}{n}a, w \right) = \frac{m}{n}(a, w)$  для любых натуральных  $m$  и  $n$ . Формула (2.2) показывает, что  $(u, v) = -(u, -v)$  и  $(\lambda u, v)$  непрерывно зависит от  $\lambda$ . Поэтому  $(\lambda u, v) = \lambda(u, v)$  для всех  $\lambda \in \mathbb{R}$ .

Ясно также, что  $(u, u) = \|u\|^2 \geq 0$ .  $\square$

### 2.8.2. Эквивалентные нормы

Пусть  $V$  — нормированное  $n$ -мерное пространство. Единичным шаром в  $V$  называют множество  $D^n = \{v \in V \mid \|v\| \leq 1\}$ , а единичной сферой в  $V$  называют множество  $S^{n-1} = \{v \in V \mid \|v\| = 1\}$ .

**Теорема 2.8.2.1.** а) Единичный шар  $D^n$  — выпуклое множество, симметричное относительно начала координат.

б) Пусть  $(\cdot, \cdot)$  — некоторое скалярное произведение в  $V$ . Тогда можно выбрать положительные числа  $C_1$  и  $C_2$  так, что  $C_1 \leq (v, v) \leq C_2$  для любого  $v \in S^{n-1}$ .

**Доказательство.** а) Симметричность  $D^n$  относительно начала координат очевидна, поскольку  $\| -v \| = \|v\|$ . Докажем теперь, что множество  $D^n$  выпуклое. Пусть  $\|u\| \leq 1$  и  $\|v\| \leq 1$ . Пусть, далее,  $t \in [0, 1]$ . Тогда

$$\begin{aligned} \|tu + (1-t)v\| &\leq \|tu\| + \|(1-t)v\| = \\ &= t\|u\| + (1-t)\|v\| \leq \\ &\leq t + (1-t) = 1. \end{aligned}$$

б) Функция  $f(u) = \|u\|$  непрерывна на (обычной) единичной сфере, заданной уравнением  $(u, u) = 1$ . Поэтому она достигает на ней минимального значения  $c_1$  и максимального значения  $c_2$ . Это означает, что если  $(u, u) = 1$ , то  $c_1 \leq \|u\| \leq c_2$  (число  $c_1$  положительно, потому что  $\|u\| > 0$ , если  $u \neq 0$ ).



Любой вектор  $v$  можно представить в виде  $v = tu$ , где  $t \in \mathbb{R}$  и  $(u, u) =$

1. Ясно также, что  $\frac{\|v\|^2}{(v, v)} = \frac{\|tu\|^2}{(tu, tu)} = \frac{\|u\|^2}{(u, u)}$ , поэтому  $c_1 \leq \frac{\|v\|^2}{(v, v)} \leq c_2$ , а значит,

$$\frac{1}{c_1} \|v\|^2 \leq (v, v) \leq \frac{1}{c_2} \|v\|^2.$$

Таким образом, можно положить  $C_1 = 1/c_2^2$  и  $C_2 = 1/c_1^2$ .  $\square$

**Замечание 1.** Единичная сфера  $S^{n-1}$  (так же, как и единичный шар  $D^n$ ) полностью определяет норму. Действительно, для любого ненулевого вектора  $v$  существует единственное  $\lambda > 0$ , для которого  $\lambda v \in S^{n-1}$ . При этом  $\lambda \|v\| = 1$ , т.е.  $\|v\| = \lambda^{-1}$ .

**Замечание 2.** В утверждении б) существенно, что мы имеем дело с конечномерным пространством (в этом случае сфера компактна). Для бесконечномерных пространств это утверждение неверно; не верны и приведённые ниже следствия из него.

Пусть  $V$  и  $W$  — нормированные пространства. Отображение  $f : V \rightarrow W$  называют *непрерывным*, если для любого  $v \in V$  и любого  $\varepsilon > 0$  можно выбрать  $\delta > 0$  так, что из неравенства  $\|v_1\| < \delta$  следует неравенство  $\|f(v + v_1) - f(v)\| < \varepsilon$ .

**Теорема 2.8.2.2.** Любое линейное отображение нормированных пространств непрерывно.

**Доказательство.** Если  $f$  — линейное отображение, то  $f(v + v_1) - f(v) = f(v_1)$ , поэтому  $\|f(v + v_1) - f(v)\| = \|f(v_1)\|$ . Таким образом, достаточно доказать, что для любого  $\varepsilon > 0$  можно выбрать  $\delta > 0$  так, что если  $\|v_1\| < \delta$ , то  $\|f(v_1)\| < \varepsilon$ .

Пусть  $f(\sum_{j=1}^n x_j e_j) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j e_i$ . Обозначим наибольшее из чисел  $|a_{ij}|$  через  $A$ . Тогда относительно стандартного скалярного произведения квадрат длины вектора  $v_1 = \sum x_j e_j$  равен  $\sum x_j^2$ , а квадрат длины вектора  $f(v_1)$  равен

$$\sum_{i=1}^m \left( \sum_{j=1}^n a_{ij} x_j \right)^2 \leq \sum_{i=1}^m \left( \sum_{j=1}^n A |x_j| \right)^2 \leq n A^2 \sum x_j^2.$$

Согласно теореме 2.8.2.1 можно выбрать положительные константы  $C_1$  и  $C_2$  так, что  $C_1 \|v_1\|^2 \leq (v_1, v_1) \leq C_2 \|v_1\|^2$ . Поэтому если  $\|v_1\| < \delta$ , то  $(v_1, v_1) \leq C_2 \delta^2$ , а значит,  $C_1 \|f(v_1)\|^2 \leq (f(v_1), f(v_1)) \leq n A^2 (v_1, v_1) \leq n A^2 \delta^2 C_2$ , т.е.  $\|f(v_1)\| \leq A \delta \sqrt{n C_2 / C_1}$ . Таким образом, для данного  $\varepsilon > 0$  мы выбираем  $\delta > 0$  так, чтобы выполнялось неравенство  $\delta \sqrt{n C_2 / C_1} < \varepsilon$ .

$\square$

Две нормы  $\|\cdot\|_1$  и  $\|\cdot\|_2$  на одном и том же пространстве  $V$  называют *эквивалентными*, если существуют положительные константы  $k$  и  $K$ , для которых  $k\|v\|_1 \leq \|v\|_2 \leq K\|v\|_1$  для любого  $v \in V$ . Ясно, что непрерывность отображения не изменяется при замене нормы на эквивалентную ей норму.

Понятие эквивалентности норм важно для бесконечномерных пространств, но на конечномерном пространстве  $V$  все нормы эквивалентны. Действительно, введём на  $V$  какое-нибудь скалярное произведение. Согласно теореме 2.8.2.1 можно выбрать положительные константы  $C_1, C_2, C'_1, C'_2$  так, что  $C_1\|v\|_1^2 \leq (v, v) \leq C_2\|v\|_1^2$  и  $C'_1\|v\|_2^2 \leq (v, v) \leq C'_2\|v\|_2^2$ . Тогда  $\|v\|_2^2 \leq \frac{1}{C'_1}(v, v) \leq \frac{C_2}{C'_1}\|v\|_1^2$  и  $\|v\|_2^2 \geq \frac{1}{C'_2}(v, v) \geq \frac{C_1}{C'_2}\|v\|_1^2$ .

Таким образом, для конечномерных линейных пространств понятие непрерывности отображения не зависит от выбора нормы.

### 2.8.3. Метрическое определение середины отрезка

Множество  $M$  называют *метрическим пространством*, если для любых его двух элементов  $x, y$  задано *расстояние*  $d(x, y) \in \mathbb{R}$ , обладающее следующими свойствами:

- 1)  $d(x, y) \geq 0$  для любых  $x, y \in M$ ;
- 2)  $d(x, y) = 0$  тогда и только тогда, когда  $x = y$ ;
- 3)  $d(x, y) \leq d(x, z) + d(y, z)$  для любых  $x, y, z \in M$  (*неравенство треугольника*).

Функцию  $d(x, y)$  называют также *метрикой*.

В метрическом пространстве для любого подмножества  $X$  можно определить его *диаметр*  $\delta(X)$  как точную верхнюю грань расстояний между его элементами.

Легко видеть, что в любом нормированном пространстве метрику можно задать формулой  $d(u, v) = \|u - v\|$ .

В нормированном пространстве  $V$  вектор  $\frac{1}{2}(v_1 + v_2)$  можно определить непосредственно через метрику, не обращаясь к линейной структуре. Это делается следующим образом (см. [Б1]). Рассмотрим сначала множество  $H_1$ , состоящее из тех  $v \in V$ , для которых  $d(v, v_1) = d(v, v_2) = \frac{1}{2}d(v_1, v_2)$ . Ясно, что если  $v, v' \in H_1$ , то

$$d(v, v') \leq d(v, v_1) + d(v', v_1) = d(v_1, v_2),$$

поэтому  $\delta(H_1) \leq d(v_1, v_2)$ . Определим множество  $H_n$  по индукции следующим образом:  $H_n$  состоит из тех  $w \in H_{n-1}$ , для которых  $d(w, v) \leq \frac{1}{2}\delta(H_{n-1})$  для всех  $v \in H_{n-1}$ .

Легко проверить, что  $\lim_{n \rightarrow \infty} \delta(H_n) = 0$ . Действительно, предположим, что множество  $H_n$  не пусто. Пусть  $v, v' \in H_n$ . Тогда по определению  $v \in H_{n-1}$ , а значит,  $d(v, v') \leq \frac{1}{2}\delta(H_{n-1})$ , поэтому  $\delta(H_n) \leq \frac{1}{2}\delta(H_{n-1})$ . Таким образом,  $\delta(H_n) \leq \frac{1}{2^{n-1}}d(v_1, v_2)$ .

Итак, если пересечение всех множеств  $H_n$  непусто, то оно состоит из одного элемента. Покажем, что  $v = \frac{1}{2}(v_1 + v_2) \in H_n$  для всех  $n$ . Прежде всего заметим, что  $\|v - v_1\| = \|v - v_2\| = \frac{1}{2}\|v_1 - v_2\|$ , поэтому  $v \in H_1$ .

**Лемма.** Если  $w \in H_n$ , то  $\bar{w} = v_1 + v_2 - w \in H_n$ .

**Доказательство.** Применим индукцию по  $n$ . Пусть сначала  $n = 1$ . Если  $w \in H_1$ , то

$$\|\bar{w} - v_1\| = \|v_2 - w\| = \frac{1}{2}\|v_1 - v_2\| = \|v_1 - w\| = \|\bar{w} - v_2\|,$$

поэтому  $\bar{w} \in H_1$ . Предположим, что требуемое утверждение доказано для  $n - 1$ . Тогда если  $u \in H_{n-1}$ , то  $v_1 + v_2 - u \in H_{n-1}$ . Поэтому если  $w \in H_n$ , то для любого  $u \in H_{n-1}$

$$\|\bar{w} - u\| = \|v_1 + v_2 - w - u\| = \|\bar{u} - w\| \leq \frac{1}{2}\delta(H_{n-1}),$$

поэтому  $\bar{w} \in H_n$ .  $\square$

Мы уже доказали, что  $v = \frac{1}{2}(v_1 + v_2) \in H_1$ . Предположим, что  $v \in H_{n-1}$ . Пусть  $w \in H_{n-1}$ . Тогда согласно лемме  $\bar{w} = v_1 + v_2 - w \in H_{n-1}$ . Поэтому

$$2\|v - w\| = \|v_1 + v_2 - w\| = \|\bar{w} - w\| \leq \delta(H_{n-1}).$$

Следовательно,  $v \in H_n$ .

#### 2.8.4. Изометрические отображения

Пусть  $V$  и  $W$  — нормированные пространства. Отображение  $f: V \rightarrow W$  называют *изометрическим*, если  $\|u - v\| = \|f(u) - f(v)\|$  для любых  $u, v \in V$ . Ясно, что изометрическое отображение непрерывно.

Два нормированных пространства  $V$  и  $W$  называют *изометричными*, если существует изоморфизм  $f: V \rightarrow W$ , который является изометрическим отображением.

**Теорема 2.8.4.1 (Мазур—Улам).** Если  $f$  — изометрическое отображение, причём  $f(0) = 0$ , то  $f$  — линейное отображение.

**Доказательство.** Основная часть доказательства этой теоремы содержится в § 2.8.3. Там доказано, что для любых  $u, v \in V$  вектор  $\frac{1}{2}(u + v)$  полностью определяется метрикой. Поэтому при изометрическом отображении  $f$  вектор  $\frac{1}{2}(u + v)$  переходит в  $\frac{1}{2}(f(u) + f(v))$ . Положим  $v = 0$  и воспользуемся тем, что  $f(0) = 0$ . В результате получим  $f\left(\frac{1}{2}\right) = \frac{1}{2}f(u)$ . Следовательно, для любых  $u, v \in V$

$$f(u + v) = f\left(\frac{1}{2}(2u + 2v)\right) = \frac{1}{2}f(2u) + \frac{1}{2}f(2v) = f(u) + f(v).$$

Отображение  $f$  непрерывно, поэтому оно линейно (это доказывается точно так же, как в конце доказательства теоремы 2.8.1.1: сначала доказывается линейность над полем рациональных чисел, а затем используется непрерывность).  $\square$

### 2.8.5. Неравенства для норм

**Теорема 2.8.5.1.** [Ди] Пусть  $u$  и  $v$  — ненулевые векторы нормированного пространства. Тогда

$$\|u - v\| \geq \frac{1}{4}(\|u\| + \|v\|) \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\|,$$

причём константу  $1/4$  нельзя увеличить.

**Доказательство.** Ясно, что

$$\begin{aligned} \|u\| \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\| &\leq \|u\| \left\| \frac{u}{\|u\|} - \frac{v}{\|u\|} \right\| + \|u\| \left\| \frac{v}{\|u\|} - \frac{v}{\|v\|} \right\| \leq \\ &\leq \|u - v\| + \frac{\|(\|v\| - \|u\|)v\|}{\|v\|} \leq \\ &\leq \|u - v\| + |(\|v\| - \|u\|)v| \leq 2\|u - v\|. \end{aligned}$$

Аналогично

$$\|v\| \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\| \leq 2\|u - v\|.$$

Сложив эти два неравенства, получаем требуемое.

Предположим теперь, что неравенство

$$\|u - v\| \geq C(\|u\| + \|v\|) \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\| \quad (2.1)$$

имеет место для любых ненулевых векторов в любом нормированном пространстве. Рассмотрим пространство  $\mathbb{R}^2$  с нормой  $\|(x, y)\| = |x| + |y|$ . Выберем в нём векторы  $u = (1, \varepsilon)$  и  $v = (1, 0)$ , где  $\varepsilon > 0$ . Для этих векторов неравенство (2.1) принимает вид

$$\varepsilon > C(2 + \varepsilon) \frac{2\varepsilon}{1 + \varepsilon},$$

т.е.  $C \leq \frac{1+\varepsilon}{4+2\varepsilon}$ . Следовательно,  $C \leq 1/4$ .  $\square$

**Теорема 2.8.5.2.** [Du] В нормированном пространстве со скалярным произведением константу  $1/4$  в теореме 2.8.5.1 можно заменить на  $1/2$ .

**Доказательство.** В нормированном пространстве со скалярным произведением

$$\begin{aligned} \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\|^2 &= \left( \frac{u}{\|u\|} - \frac{v}{\|v\|}, \frac{u}{\|u\|} - \frac{v}{\|v\|} \right) = \\ &= 2 - \frac{2(u, v)}{\|u\| \cdot \|v\|} = \\ &= \frac{1}{2\|u\| \cdot \|v\|} (2\|u\| \cdot \|v\| - 2(u, v)) = \\ &= \frac{1}{2\|u\| \cdot \|v\|} (2\|u\| \cdot \|v\| - (\|u\|^2 + \|v\|^2 - \|u - v\|^2)) = \\ &= \frac{\|u - v\|^2 - (\|u\| - \|v\|)^2}{\|u\| \cdot \|v\|}. \end{aligned}$$

Поэтому

$$\begin{aligned} \|u - v\|^2 - \left( \frac{\|u\| + \|v\|}{2} \right)^2 \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\|^2 &= \\ = \|u - v\|^2 - \left( \frac{\|u\| + \|v\|}{2} \right)^2 \frac{\|u - v\|^2 - (\|u\| - \|v\|)^2}{\|u\| \cdot \|v\|} &= \\ \frac{(\|u\| - \|v\|)^2}{4\|u\| \cdot \|v\|} ((\|u\| + \|v\|)^2 - \|u - v\|^2); \end{aligned}$$

при записи последнего неравенства мы воспользовались тем, что

$$\|u - v\|^2 \left( 1 - \frac{(\|u\| + \|v\|)^2}{4\|u\| \cdot \|v\|} \right) = \|u - v\|^2 \frac{(\|u\| - \|v\|)^2}{4\|u\| \cdot \|v\|}.$$

Остаётся заметить, что  $(\|u\| + \|v\|)^2 \geq \|u - v\|^2$ , поскольку  $\|u\| \cdot \|v\| \geq -(u, v)$ .  $\square$

### 2.8.6. Двойственная норма

Пусть  $V$  — пространство со скалярным произведением. Тогда для любой нормы  $\|\cdot\|$  в этом пространстве можно определить двойственную норму  $\|\cdot\|^*$  следующим образом:

$$\|v\|^* = \max_{\|w\|=1} |(v, w)|.$$

Легко видеть, что при этом мы действительно получим норму. Например, неравенство  $\|u + v\|^* \leq \|u\|^* + \|v\|^*$  следует из того, что  $|(u + v, w)| \leq |(u, w)| + |(v, w)|$ .

**Теорема 2.8.6.1.** *Норма  $\|\cdot\|$  двойственна сама себе тогда и только тогда, когда  $\|v\| = \sqrt{(v, v)}$ .*

**Доказательство.** Предположим сначала, что  $\|v\| = \sqrt{(v, v)}$ . Тогда  $(\|v\|^*)^2 = \max_{\|w\|=1} (v, w)^2$ . Но  $(v, w)^2 \leq (v, v) \cdot (w, w) = (v, v)$ , причём равенство достигается для некоторого вектора  $w$ , пропорционального  $v$ . Поэтому  $\|v\|^* = \sqrt{(v, v)} = \|v\|$ .

Предположим теперь, что  $\|v\|^* = \|v\|$  для любого вектора  $v$ . Если  $w = v/\|v\|$ , то  $\|w\| = 1$ . Поэтому  $\|v\|^* = \max_{\|w\|=1} |(v, w)| \geq (v, v)/\|v\|$ , т.е.  $(v, v) \leq \|v\|^* \|v\| = \|v\|^2$ . Остаётся доказать, что выполняется и противоположное неравенство  $\|v\|^2 \leq (v, v)$ . Воспользовавшись уже доказанным неравенством  $(w, w) \leq \|w\|^2$ , получаем

$$\begin{aligned} \|v\| = \|v\|^* &= \max_{\|w\|=1} |(v, w)| = \max_{\|w\| \neq 0} \left| \left( v, \frac{w}{\|w\|} \right) \right| = \\ &= \max_{\|w\| \neq 0} \left| \left( v, \frac{w}{\sqrt{(w, w)}} \right) \right| \cdot \frac{\sqrt{(w, w)}}{\|w\|} \leq \max_{\|w\| \neq 0} \frac{|(v, w)|}{\sqrt{(w, w)}}. \end{aligned}$$

Последний максимум достигается при  $w = \pm \frac{v}{\sqrt{(v, v)}}$ ; он равен  $\sqrt{(v, v)}$ .  $\square$

**Пример 2.8.6.1.** *Нормы из примеров 2.8.0.2 и 2.8.0.3 двойственны друг другу.*

**Доказательство.** Пусть  $x = (x_1, \dots, x_n)$  и  $\|x\|_1 = \max |x_i|$ ,  $\|x\|_2 = \sum |x_i|$ . Тогда  $\|y\|_1^*$  — это максимальное из чисел  $|\sum x_i y_i|$ , где  $\max |x_i| = 1$ . Ясно, что максимум достигается для некоторого  $x = (\pm 1, \dots, \pm 1)$ , и он равен  $\sum |y_i| = \|y\|_2$ . Далее,  $\|y\|_2^*$  — это максимальное из чисел  $|\sum x_i y_i|$ , где  $\sum |x_i| = 1$ . Ясно, что если наибольшую абсолютную величину имеет координата  $y_k$ , то максимум достигается для  $x = (0, \dots, 1, \dots, 0)$ , где 1 стоит на  $k$ -м месте, и он равен  $|y_k| = \|y\|_1$ .  $\square$

**Пример 2.8.6.2.** Нормы  $l_p$  и  $l_q$  из примера 2.8.0.4 двойственны друг другу, если  $\frac{1}{p} + \frac{1}{q} = 1$ .

**Доказательство.** Требуется доказать, что если  $(y_1, \dots, y_n)$  — фиксированный набор (действительных) чисел, то для любого набора  $(x_1, \dots, x_n)$ , для которого  $(\sum_{i=1}^n |x_i|^p)^{1/p} = 1$ , имеет место неравенство  $|\sum x_i y_i| \leq (\sum_{i=1}^n |y_i|^q)^{1/q}$ , причём для некоторого набора достигается равенство. Это следует из известного *неравенства Гёльдера*

$$|\sum x_i y_i| \leq \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \left( \sum_{i=1}^n |y_i|^q \right)^{1/q}.$$

Равенство здесь достигается, когда числа  $x_i$  и  $y_i$  одного знака и  $|y_i|^q = \lambda |x_i|^p$  для некоторой константы  $\lambda$ .  $\square$

**Теорема 2.8.6.2.** Пусть  $\|\cdot\|_1 = \|\cdot\|^*$  и  $\|\cdot\|_2 = \|\cdot\|_1^*$ . Тогда  $\|\cdot\|_2 = \|\cdot\|$ , т.е. норма, двойственная двойственной норме, совпадает с исходной нормой.

**Доказательство.** Пусть  $D^n$  и  $D_2^n$  — единичные шары для норм  $\|\cdot\|$  и  $\|\cdot\|_2$ . Достаточно доказать, что  $D^n = D_2^n$ .

Согласно определению двойственной нормы

$$\|v\|^* = \max_{\|u\|=1} |(v, u)| \geq \frac{(v, w)}{\|w\|}$$

для любого  $w \neq 0$ . Поэтому для любых векторов  $v$  и  $w$  имеет место неравенство  $(v, w) \leq \|w\| \cdot \|v\|^*$ . Следовательно,

$$\|v\|_2 = \max_{\|w\|^*=1} |(v, w)| \leq \max_{\|w\|^*=1} (\|w\|^* \cdot \|v\|) = \|v\|.$$

Из неравенства  $\|v\|_2 \leq \|v\|$  следует, что  $D^n \subset D_2^n$ .

Докажем теперь, что  $D_2^n \subset D^n$ . Пусть  $u \in D_2^n$ . Тогда  $|(u, v)| \leq 1$  для любого  $v$  с  $\|v\|^* = 1$ , т.е. для любого  $v$ , для которого  $|(w, v)| \leq 1$  для всех  $w \in D^n$ . Для фиксированного  $v$  неравенство  $|(x, v)| \leq 1$  задаёт полосу  $\Pi_v$  между двумя гиперплоскостями, симметричными относительно начала координат. Мы доказали, что если  $u \in \Pi_v$ , то  $D^n \subset \Pi_v$ . Поэтому  $D_2^n$  содержится в пересечении всех полос, содержащих  $D^n$ . Но  $D^n$  — выпуклое тело, симметричное относительно начала координат, поэтому указанное пересечение полос совпадает с  $D^n$ .  $\square$

### Задачи

**8.1.** Докажите, что при  $n = 2$  нормированные пространства из примеров 2.8.0.2 и 2.8.0.3 изометричны, а при  $n \geq 3$  не изометричны.

**8.2.** Докажите, что нормированное пространство является пространством со скалярным произведением тогда и только тогда, когда любое его двумерное подпространство является пространством со скалярным произведением.

**8.3.** [Ке] Пусть  $x, y, z$  — линейно зависимые векторы в (вещественном) нормированном пространстве. Докажите, что

$$\|x\| + \|y\| + \|z\| + \|x + y + z\| \geq \|x + y\| + \|y + z\| + \|z + x\|.$$

**8.4.** [Ви] Пусть  $0 < a < 1$  и  $x, y, z$  — векторы в нормированном пространстве с метрикой  $d$ . Докажите, что если для любого вектора  $w$  из этого нормированного пространства имеет место неравенство

$$d(w, z) \leq ad(w, x) + (1 - a)d(w, y),$$

то  $z = ax + (1 - a)y$ .

### Решения

**1.1.** Пусть  $V$  — пространство всех матриц порядка  $n$ ,  $W \subset V$  — подпространство матриц с нулевым следом. Тогда  $A^T \in W^\perp$ . Кроме того,  $I \in W^\perp$  и  $\dim W^\perp = 1$ . Поэтому  $A^T = \lambda I$ .

**1.2.** Пусть  $A_1, \dots, A_m$  и  $B_1, \dots, B_k$  — строки матриц  $A$  и  $B$ . Из равенств  $A_1X = \dots = A_mX = 0$  следуют равенства  $B_1X = \dots = B_kX = 0$ , т.е. если  $X \in \langle A_1, \dots, A_m \rangle^\perp$ , то  $X \in \langle B_1, \dots, B_k \rangle^\perp$ . Поэтому  $\langle A_1, \dots, A_m \rangle^\perp \subset \langle B_1, \dots, B_k \rangle^\perp$ , а значит,  $\langle B_1, \dots, B_k \rangle \subset \langle A_1, \dots, A_m \rangle$ . Следовательно, строки матрицы  $B$  являются линейными комбинациями строк матрицы  $A$ , т.е.  $b_{ij} = \sum c_{ip} a_{pj}$ .

**1.3.** Пусть  $v = (a_1, \dots, a_n)$ . Если  $w = (b_1, \dots, b_n)$  — вектор из ортанта, не содержащего ни  $v$ , ни  $-v$ , то  $a_i b_i > 0$  и  $a_j b_j < 0$  для некоторых  $i$  и  $j$ . К вектору  $w$  можно прибавить вектор  $w_1 = (x_1, \dots, x_n)$ , где  $a_i x_i > 0$  и  $a_j x_j < 0$ , а все остальные координаты равны нулю. Числа  $x_i$  и  $x_j$  можно подобрать так, что  $a_i x_i + a_j x_j = -(v, w)$ , т.е.  $(v, w + w_1) = 0$ .

**1.4.** Точка  $X$  принадлежит прямой  $a$  тогда и только тогда, когда  $a^\perp(\overrightarrow{AX}) = 0$ , т.е.  $a^\perp(\overrightarrow{OX}) = a^\perp(\overrightarrow{OA})$ . Пусть  $X$  — точка пересечения прямых  $a$  и  $b$ . Тогда  $a^\perp(\overrightarrow{OX}) = a^\perp(\overrightarrow{OA})$  и  $b^\perp(\overrightarrow{OX}) = b^\perp(\overrightarrow{OB})$ . Поэтому  $c^\perp(\overrightarrow{OX}) = -(a^\perp + b^\perp)(\overrightarrow{OX}) =$



$-a^\perp(\overrightarrow{OA}) - b^\perp(\overrightarrow{OB})$ . Точка  $X$  принадлежит прямой  $c$  тогда и только тогда, когда  $c^\perp(\overrightarrow{OC}) = c^\perp(\overrightarrow{OX}) = -a^\perp(\overrightarrow{OA}) - b^\perp(\overrightarrow{OB})$ .

**1.5.** Пусть  $B(x, y) = x^*(y)$ . Рассмотрим для фиксированного  $x$  две линейные функции  $f_1(y) = x^*(y)$  и  $f_2(y) = y^*(x)$ . По условию равенства  $f_1(y) = 0$  и  $f_2(y) = 0$  эквивалентны. Поэтому  $\text{Ker } f_1 = \text{Ker } f_2$ , а значит,  $B(x, y) = k(x)B(y, x)$ . Докажем, что  $k(x)$  — постоянное число. С одной стороны,  $B(x + x', y) = x^*(y) + x'^*(y) = k(x)y^*(x) + k(x')y^*(x')$ . С другой стороны,  $B(x + x', y) = k(x + x')B(y, x + x') = k(x + x')y^*(x) + k(x + x')y^*(x')$ . Поэтому  $(k(x + x') - k(x))y^*(x) = (k(x') - k(x + x'))y^*(x')$ . Если  $x' = \lambda x$ , то  $k(x') = k(x)$ . Поэтому можно считать, что векторы  $x$  и  $x'$  неколлинеарны. Но тогда вектор  $y$  можно подобрать так, что  $y^*(x) \neq 0$  и  $y^*(x') \neq 0$ . Следовательно,  $k(x + x') = k(x)$ . Таким образом,  $B(x, y) = kB(y, x) = k^2 B(x, y)$ . Можно подобрать такие векторы  $x$  и  $y$ , что  $B(x, y) \neq 0$ . Поэтому  $k^2 = 1$ , т.е.  $k = \pm 1$ . При  $k = 1$  получаем симметрическую функцию, при  $k = -1$  — кососимметрическую.

**1.6.** Ответ:  $(0 : x_1 : \dots : x_n)$ . Если точка лежит в гиперплоскости  $A_1 \dots A_n$ , то она имеет барицентрические координаты вида  $(0 : y_1 : \dots : y_n)$ , поскольку является центром масс точек  $A_1, \dots, A_n$  с некоторыми массами. Остаётся доказать, что любая точка  $X$  прямой  $PA_0$  имеет барицентрические координаты вида  $(\alpha : x_1 : \dots : x_n)$ . Для этого достаточно заметить, что  $\overrightarrow{A_0 X} = \lambda \overrightarrow{A_0 P} = \lambda x_1 \overrightarrow{A_0 A_1} + \dots + \lambda x_n \overrightarrow{A_0 A_n}$ .

**1.7.** Достаточно заметить, что если  $X$  — точка прямой  $AB$ , а  $O$  — произвольна точка, то  $\overrightarrow{OX} = \lambda \overrightarrow{OA} + (1 - \lambda) \overrightarrow{OB}$  для некоторого  $\lambda$ .

**1.8.** Воспользуемся результатом задачи 1.7. Бесконечно удалённая точка получается при  $\lambda \rightarrow \infty$ . При этом мы получаем барицентрические координаты  $(a_0 - b_0 : \dots : a_n - b_n)$ . Остаётся заметить, что  $\sum a_i = \sum b_i = 1$ .

**1.9.** Центром масс точек  $B_0 = \sum y_i^0 A_i, \dots, B_n = \sum y_i^n A_i$  с массами  $t_0, \dots, t_n$  является точка  $t_0 \sum y_i^0 A_i + \dots + t_n \sum y_i^n A_i = (t_0 y_0^0 + \dots + t_n y_0^n) A_0 + \dots + (t_0 y_n^0 + \dots + t_n y_n^n) A_n$ . Поэтому  $(t_0 : \dots : t_n)$  — барицентрические координаты точки  $X$  относительно точек  $B_0, \dots, B_n$ , если

$$(t_0, \dots, t_n) \begin{pmatrix} y_0^0 & \dots & y_n^0 \\ \dots & \dots & \dots \\ y_0^n & \dots & y_n^n \end{pmatrix} = (x_0, \dots, x_n),$$

т.е.

$$(t_0, \dots, t_n) = (x_0, \dots, x_n) \begin{pmatrix} y_0^0 & \dots & y_n^0 \\ \dots & \dots & \dots \\ y_0^n & \dots & y_n^n \end{pmatrix}^{-1}.$$

**1.10.** Отношение ориентированных объёмов симплексов  $P_0 \dots P_n$  и  $A_0 \dots A_n$  равно определителю матрицы, составленной из координат векторов  $\overrightarrow{P_0 P_1}, \dots, \overrightarrow{P_0 P_n}$  относительно базиса  $\overrightarrow{A_0 A_1}, \dots, \overrightarrow{A_0 A_n}$ . При доказательстве теоремы 2.1.10.1 было доказано, что определитель этой матрицы равен указанному определителю.

**1.11.** Задачу а) можно получить из задачи б) при  $r = 0$ , поэтому решим сразу задачу б). Параллельные прямые проходят через бесконечно удалённую точку  $P$  с барицентрическими координатами  $(x_0 : x_1 : \dots : x_n)$ , сумма которых равна нулю (задача 1.8). Для бесконечно удалённой точки  $P$  результат задачи 1.6 остаётся в силе, поэтому точка  $B_i$  имеет барицентрические координаты  $(x_0 : \dots : 0 : \dots : x_n)$ , где 0 стоит на  $i$ -м месте. Учитывая, что  $x_0 + \dots + x_n = 0$ , получаем, что абсолютные барицентрические координаты точки  $B_i$  равны

$$\left( -\frac{x_0}{x_i} : \dots : 0 : \dots : -\frac{x_n}{x_i} \right).$$

Значит, согласно задаче 1.10 отношение ориентированного объёма симплекса  $A_0 A_1 \dots A_{r-1} B_r B_{r+1} \dots B_n$  к объёму исходного симплекса равно

$$\begin{vmatrix} 0 & -\frac{x_{r+1}}{x_r} & \dots & -\frac{x_n}{x_r} \\ -\frac{x_r}{x_{r+1}} & 0 & \dots & -\frac{x_n}{x_{r+1}} \\ \dots & \dots & \dots & \dots \\ -\frac{x_r}{x_n} & -\frac{x_{r+1}}{x_n} & \dots & 0 \end{vmatrix};$$

мы воспользовались тем, что  $\begin{vmatrix} I & 0 \\ X_1 & X_2 \end{vmatrix} = |X_2|$ . Полученный определитель легко преобразуется к виду

$$\frac{1}{x_r \dots x_n} \begin{vmatrix} 0 & -x_{r+1} & \dots & -x_n \\ -x_r & 0 & \dots & -x_n \\ \dots & \dots & \dots & \dots \\ -x_r & -x_{r+1} & \dots & 0 \end{vmatrix}.$$

Прибавим к первой строке все остальные строки, вынесем для первой строки множитель  $n - r$ , а затем вычтем первую строку из всех остальных. В результате получим верхнюю треугольную матрицу с элементами  $(-x_r, x_{r+1}, \dots, x_n)$  на диагонали. Таким образом, рассматриваемый определитель равен  $r - n$ .

**2.1.** Пусть  $x$  — столбец такого же размера, как и матрица  $A$ , и  $y = Ax$ . Тогда  $y^T y = x^T A^T A x$ . Ясно, что равенство  $y^T y = 0$  выполняется только для нулевого столбца, поэтому если  $A^T A x = 0$ , то  $Ax = y = 0$ . Следовательно,  $\text{Ker}(A^T A) = \text{Ker } A$ . Матрицы  $A$  и  $A^T A$  задают отображения одного и того же линейного пространства (хотя и в пространства разной размерности, если матрица  $A$  не квадратная), поэтому  $\dim \text{Im } A = \dim \text{Im}(A^T A)$ .

**2.2.** Так как для любых отображений  $A : U \rightarrow V$  и  $B : V \rightarrow W$  справедлива формула  $\dim(\text{Im } A \cap \text{Ker } B) = \dim \text{Ker } BA - \dim \text{Ker } A$  (теорема 2.2.1.1), то  $\dim(\text{Im } A^k \cap \text{Ker } A) = \dim \text{Ker } A^{k+1} - \dim \text{Ker } A^k$  для любого  $k$ . Следовательно,

$$\sum_{k=1}^n \dim(\text{Im } A^k \cap \text{Ker } A) = \dim \text{Ker } A^{n+1} - \dim \text{Ker } A.$$

Для доказательства второго равенства достаточно заметить, что  $\dim \text{Im } A^p = \dim V - \dim \text{Ker } A^p$ , где  $V$  — пространство, в котором действует оператор  $A$ .

**2.3.** Можно считать, что первые  $r$  функций  $f_1, \dots, f_r$  линейно независимы, а функции  $f_{r+1}, \dots, f_k$  через них линейно выражаются. Дополним первые  $r$  функций до базиса, и пусть  $e_1, \dots, e_n$  — двойственный ему базис. Тогда  $\bigcap_{i=1}^k \text{Ker } f_i = \bigcap_{i=1}^r \text{Ker } f_i = \langle e_{r+1}, \dots, e_n \rangle$ . Предположим, что  $\text{Ker } f$  содержит  $e_{r+1}, \dots, e_n$ . Покажем, что тогда  $f = \sum_{i=1}^r f(e_i)f_i$ . Действительно, если  $j \leq r$ , то  $\sum_{i=1}^r f(e_i)f_i(e_j) = \sum_{i=1}^r f(e_i)\delta_{ij} = f(e_j)$ , а если  $j > r$ , то  $\sum_{i=1}^r f(e_i)f_i(e_j) = 0 = f(e_j)$ . В обратную сторону утверждение очевидно.

**2.4.** Линейные функции на  $\text{Ker } B$  — это линейные функции на  $V$ , рассматриваемые с точностью до функций вида  $g(Bx)$ ; функции вида  $g(Bx)$  образуют пространство  $\text{Im } B^*$ . Линейные функции на  $\text{Ker } B / \text{Im } A$  — это линейные функции на  $\text{Ker } B$ , для которых  $f(Ax) = 0$  для всех  $x$ ; функции, для которых  $f(Ax) = 0$ , образуют пространство  $\text{Ker } A^*$ .

**3.1.** Пусть у одного из векторов  $e_1, \dots, e_{m+1}$  первая координата отлична от нуля. Без ограничения общности можно считать, что это — вектор  $e_{m+1}$ . Рассмотрим систему векторов  $e'_1 = e_1 - x_1 e_{m+1}, \dots, e'_m = e_m - x_m e_{m+1}$ , где  $x_i$  — первая координата вектора  $e_i$ , делённая на первую координату вектора  $e_{m+1}$ . Если же у всех векторов первая координата нулевая, то просто отбросим вектор  $e_{m+1}$ . В обоих случаях получим систему  $m$  линейно независимых векторов, у которых первая координата нулевая. Применим к этой системе векторов такую же конструкцию для второй координаты и т.д.

**3.2.** Рассмотрим все подмножества данного множества векторов, состоящие из линейно независимых векторов, и выберем среди них наибольшее по числу элементов. Пусть для определённости это будут векторы  $e_1, \dots, e_k$  ( $k \leq n$ ). Тогда существуют такие числа  $\lambda_i$  и  $\mu_i$ , что

$$\lambda_1 e_1 + \dots + \lambda_k e_k + \lambda_{k+1} e_{k+1} = \mu_1 e_1 + \dots + \mu_k e_k + \mu_{k+2} e_{k+2} = 0,$$

причём  $\lambda_{k+1} \neq 0$  и  $\mu_{k+2} \neq 0$ . Если  $\sum \lambda_i = 0$  или  $\sum \mu_i = 0$ , то доказательство завершено. Поэтому можно считать, что  $\sum \lambda_i = 1$  и  $\sum \mu_i = -1$ . Тогда числа  $\alpha_i = \lambda_i + \mu_i$  ( $i \leq k$ ),  $\alpha_{k+1} = \lambda_{k+1}$ ,  $\alpha_{k+2} = \mu_{k+2}$  и  $\alpha_i = 0$  при  $i > k+2$  — искомые.

**3.3.** Доказательство проведём индукцией по  $m$ . При  $m \leq n+1$  утверждение очевидно. Пусть  $m \geq n+2$ . Тогда существуют числа  $\alpha_1, \dots, \alpha_n$ , не все равные нулю, для которых  $\sum \alpha_i v_i = 0$  и  $\sum \alpha_i v_i = 0$  (см. задачу 3.2). Поэтому  $x = \sum t_i v_i + \lambda \sum \alpha_i v_i = \sum t'_i v_i$ , где  $t'_i = t_i + \lambda \alpha_i$  и  $\sum t'_i = \sum t_i$ . Остаётся подобрать число  $\lambda$  так, чтобы все числа  $t_i + \lambda \alpha_i$  были неотрицательны и по крайней мере одно из них равнялось нулю. Множество

$$\{\lambda \in \mathbb{R} \mid t_i + \lambda \alpha_i \geq 0 \text{ при } i = 1, \dots, m\}$$

замкнуто, непусто (оно содержит нуль) и ограничено снизу (и сверху), так как среди чисел  $\alpha_i$  есть положительные (и отрицательные); минимальное число  $\lambda$  из этого множества искомое.

**3.4.** Предположим, что матрица  $A$  вырожденная. Тогда существуют такие числа  $\lambda_1, \dots, \lambda_n$ , не все равные нулю, что  $\sum \lambda_i a_{ik} = 0$  при  $k = 1, \dots, n$ . Пусть

для определённости  $\lambda_1$  — наибольшее по модулю среди чисел  $\lambda_1, \dots, \lambda_n$ . Так как  $\lambda_1 a_{11} + \lambda_2 a_{12} + \dots + \lambda_n a_{1n} = 0$ , то  $|\lambda_1 a_{11}| = |\lambda_2 a_{12} + \dots + \lambda_n a_{1n}| \leq |\lambda_1|(|a_{12}| + \dots + |a_{1n}|) \leq |\lambda_1| \cdot |a_{12}|$ . Получено противоречие.

**3.5.** Нет, не обязательно. Можно, например, взять векторы  $x_1 = e_1 - e_2$  и  $x_2 = e_2 - e_3 - e_1$ . Тогда  $x_1 + x_2 + e_3 = 0$ .

**3.6.** Меняя на каждом шаге базис  $v_1, \dots, e_i, \dots, v_n$  на базис  $v_1, \dots, \varepsilon_i, \dots, v_n$ , можно заменить базис  $v_1, \dots, v_n$  на положительно ориентированный базис  $e_1, \dots, e_n$ , причём количество шагов равно количеству векторов  $\varepsilon_i$  среди  $v_1, \dots, v_n$ . Поэтому свойство (1) эквивалентно тому, что на каждом шаге ориентация базиса изменяется, т.е. эквивалентно свойству (2).

**3.7. а)** Докажем сначала, что если  $\lambda_1 e_1 + \dots + \lambda_{n+1} e_{n+1} = 0$ , то все числа  $\lambda_i$  одного знака. Пусть это не так. Можно считать, что все числа  $\lambda_i$  ненулевые и

$$x_1 e_1 + \dots + x_k e_k = y_{k+1} e_{k+1} + \dots + y_{n+1} e_{n+1}, \quad (1)$$

где  $x_i, y_i \geq 0$ . Домножим скалярно равенство (1) на вектор, стоящий в левой части. Тогда в левой части получим неотрицательную величину, а в правой отрицательную. Приходим к противоречию.

Выберем теперь минимальный набор  $e_{i_1}, \dots, e_{i_k}$  линейно зависимых векторов. Нужно доказать, что в него входят все векторы  $e_1, \dots, e_{n+1}$ . Предположим, что один из векторов, например  $e_1$ , не входит в этот набор. Так как  $\sum_p \lambda_{i_p} e_{i_p} = 0$ , где  $\lambda_{i_p} > 0$ , то  $0 = (e_1, \sum_p \lambda_{i_p} e_{i_p}) = \sum_p \lambda_{i_p} (e_1, e_{i_p}) < 0$ . Получено противоречие.

**б) Первое решение.** Предположим, что векторы  $e_1, \dots, e_{n+2}$  в  $\mathbb{R}^n$  таковы, что  $(e_i, e_j) < 0$  при  $i \neq j$ . С одной стороны, если  $\alpha_1 e_1 + \dots + \alpha_{n+2} e_{n+2} = 0$ , то все числа  $\alpha_i$  одного знака (см. решение задачи а). С другой стороны, числа  $\alpha_1, \dots, \alpha_{n+2}$  можно выбрать так, что  $\sum \alpha_i = 0$  (см. задачу 3.2). Получено противоречие.

**Второе решение.** Докажем, что не существует векторов  $e_1, \dots, e_{n+2}$  в  $\mathbb{R}^n$ , которые образуют попарно тупые углы, т.е.  $(e_i, e_j) < 0$  при  $i \neq j$ . При  $n = 1$  это очевидно: если векторы  $e_1$  и  $e_2$  противоположно направлены, то вектор  $e_3$  сонаправлен с одним из них.

Пусть теперь  $n \geq 2$ . Достаточно доказать, что если в  $\mathbb{R}^n$  существуют векторы  $e_1, \dots, e_{n+2}$ , образующие попарно тупые углы, то в  $\mathbb{R}^{n-1}$  существуют векторы  $e'_1, \dots, e'_{n+1}$ , образующие попарно тупые углы. Действительно, повторяя это рассуждение мы дойдём до трёх векторов в  $\mathbb{R}^1$  и получим противоречие. Рассмотрим векторы

$$e'_i = e_{n+2} - \frac{(e_{n+2}, e_{n+2})}{(e_{n+2}, e_i)} e_i \text{ для } i = 1, \dots, n+1.$$

Легко проверить, что

$$(e'_i, e'_j) = -(e_{n+2}, e_{n+2}) + \frac{(e_{n+2}, e_{n+2})^2 (e_i, e_j)}{(e_{n+2}, e_i)(e_{n+2}, e_j)} < 0$$

при  $i \neq j$ , поскольку числа  $(e_i, e_j)$ ,  $(e_{n+2}, e_i)$  и  $(e_{n+2}, e_j)$  отрицательны. Кроме того,  $(e'_i, e_{n+2}) = 0$ , т.е. векторы  $e'_1, \dots, e'_{n+1}$  лежат в одной гиперплоскости.

**3.8.** Применим индукцию по  $n$ . При  $n = 1$  утверждение очевидно. Можно считать, что  $e_m = (1, 0, \dots, 0)$ . Тогда  $e_i = (a_i, \dots)$ , где  $a_i \leq 0$  при  $i < m$ . Запишем каждый из векторов  $e_1, \dots, e_{m-1}$  в виде  $e_i = (a_i, 0, \dots, 0) + \varepsilon_i$ . Векторы  $\varepsilon_1, \dots, \varepsilon_{m-1}$  лежат в  $(n-1)$ -мерном пространстве и при этом  $(\varepsilon_i, \varepsilon_j) = (e_i, e_j) - a_i a_j \leq 0$ . Кроме того, среди векторов  $\varepsilon_1, \dots, \varepsilon_{m-1}$  есть по крайней мере  $m-2$  ненулевых вектора, поскольку среди векторов  $e_1, \dots, e_{m-1}$  не может быть более одного вектора вида  $\lambda e_m$ , где  $\lambda < 0$ . Следовательно, согласно предположению индукции  $m-2 \leq 2n-2$ , т.е.  $m \leq 2n$ .

**Замечание.** Такие же рассуждения позволяют получить ещё одно решение задачи 3.7.

**3.9.** Применим индукцию по  $n$ . При  $n = 1$  утверждение очевидно. Предположим, что оно верно для  $n-1$ , и докажем, что тогда оно верно и для  $n$ . Выберем в  $\mathbb{R}^{n-1}$  векторы  $x'_i = (x_{i1}, \dots, x_{in})$ ,  $i = 1, \dots, n$ , для которых  $(x'_i, x'_j) = a_{ij}$  при  $1 \leq i < j \leq n$ , причём векторы  $x'_1, \dots, x'_{n-1}$  линейно независимы. Положим  $x_i = (x_{i1}, \dots, x_{in}, 0)$  для  $i = 1, \dots, n-1$  и  $x_n = (x_{1n}, \dots, x_{nn}, 1)$ . Тогда векторы  $x_1, \dots, x_n$  линейно независимы. Положим  $x_{n+1} = (t_1, \dots, t_{n+1})$  и запишем для  $t_1, \dots, t_{n+1}$  линейные уравнения  $(x_i, x_{n+1}) = a_{i,n+1}$  для  $i = 1, \dots, n$ . Матрица этой системы уравнений имеет размер  $n \times (n+1)$  и её ранг равен  $n$ , поскольку векторы  $x_1, \dots, x_n$  линейно независимы. Поэтому по теореме Кронекера–Капелли (теорема 2.2.3.1) эта система уравнений имеет решение.

**3.10.** Пусть  $n_i$  — единичный вектор нормали к грани  $F_i$  ( $i = 1, \dots, m$ ) многогранника  $M$  в  $\mathbb{R}^n$ . Рассмотрим пространство  $\mathbb{R}^{m-1}$  — ортогональное дополнение в  $\mathbb{R}^{n+m-1}$  к  $\mathbb{R}^n$ . Согласно задаче 3.9 в  $\mathbb{R}^{m-1}$  можно выбрать векторы  $x_1, \dots, x_m$ , для которых  $(x_i, x_j) = -(n_i, n_j)$  при  $i \neq j$ . Положим  $e_i = x_i + n_i$ . Вектор  $e_i$  ортогонален грани  $F_i$ , причём он не ортогонален нашему подпространству  $\mathbb{R}^n$ , поскольку  $(e_i, n_i) = (n_i, n_i) = 1$ . Кроме того,  $(e_i, e_j) = 0$  при  $i \neq j$ . Искомый  $m$ -гранный угол образуют (аффинные) гиперплоскости,  $i$ -я из которых проходит через грань  $F_i$  ортогонально  $e_i$ .

**4.1.** Пусть  $X = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$  и  $Y = \begin{pmatrix} 1 & \dots & 1 \\ y_1 & \dots & x_n \end{pmatrix}$ . Тогда  $\|a_{ij}\|_1^n = XY$  и

$$\text{rk } \|a_{ij}\|_1^n \leq \text{rk } X \leq 2.$$

**4.2.** Пусть  $e_1$  — вектор, порождающий  $\text{Im } A$ ; дополним его до базиса  $e_1, \dots, e_n$ . Матрица  $A$  относительно этого базиса имеет вид  $\begin{pmatrix} a_1 & \dots & a_n \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{pmatrix}$ . Поэтому

$$\text{tr } A = a_1 \text{ и } |A + I| = 1 + a_1.$$

**4.3.** Достаточно доказать, что  $\text{Ker } A^* \cap \text{Im } A = 0$ . Если  $A^*v = 0$  и  $v = Aw$ , то  $(v, v) = (Aw, v) = (w, A^*v) = 0$ , поэтому  $v = 0$ .

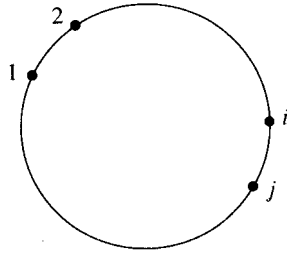


Рис. 2.3.

**4.4.** Если столбцы матрицы  $B$  являются линейными комбинациями столбцов матрицы  $A$ , то  $B = AX$  для некоторой матрицы  $X$ . Если строки матрицы  $(C, D)$  являются линейными комбинациями строк матрицы  $(A, B)$ , то  $(C, D) = Y(A, B)$ , т.е.  $C = YA$  и  $D = YB = YAX$ . А так как  $X = A^{-1}B$  и  $Y = CA^{-1}$ , то  $D = (CA^{-1})A(A^{-1}B) = CB$ . В обратную сторону рассуждения аналогичны.

**4.5.** Пусть  $v$  и  $w$  — столбцы координат векторов из  $V$  и  $W$ . Тогда  $B(v, w) = v^T Bw$ . Размерности пространств, порождённых векторами вида  $v^T B$  и  $Bw$ , равны  $\text{rk } B$ . Следовательно,  $\dim N_V = \dim V - \text{rk } B$  и  $\dim N_W = \dim W - \text{rk } B$ .

**4.6.** Пусть  $r_i = \text{rk } A_i$  и  $r = \text{rk}(A_1 + A_2)$ . Тогда  $\dim V_i = \dim W_i = r_i$  и  $\dim(V_1 + V_2) = \dim(W_1 + W_2) = r$ . Равенство  $r_1 + r_2 = r$  означает, что  $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ , т.е.  $V_1 \cap V_2 = 0$ . Аналогично  $W_1 \cap W_2 = 0$ .

**4.7.** Равенство  $B^T A = 0$  означает, что столбцы матриц  $A$  и  $B$  попарно ортогональны. Следовательно, пространства, порождённые столбцами матриц  $A$  и  $B$ , пересекаются только по нулю. Остаётся воспользоваться результатом задачи 4.6.

**4.8.** Пусть порядок матриц  $A$  и  $B$  равен  $2m + 1$ . Согласно неравенству Сильвестра  $\text{rk } A + \text{rk } B \leq \text{rk } AB + 2m + 1 = 2m + 1$ . Поэтому  $\text{rk } A \leq m$  или  $\text{rk } B \leq m$ . Если  $\text{rk } A \leq m$ , то  $\text{rk } A^T = \text{rk } A \leq m$ , а значит,  $\text{rk}(A + A^T) \leq \text{rk } A + \text{rk } A^T \leq 2m < 2m + 1$ .

**4.9.** Можно считать, что  $a_{12} \neq 0$ . Пусть  $A_i$  —  $i$ -я строка матрицы  $A$ . Докажем, что  $a_{21}A_i = a_{2i}A_1 + a_{1i}A_2$ , т.е.

$$a_{12}a_{ij} + a_{1j}a_{2i} + a_{1i}a_{j2} = 0. \quad (1)$$

Равенство (1) антисимметрично по  $i$  и  $j$ , поэтому можно считать, что  $i < j$  (рис. 2.3). В равенстве (1) отрицателен только множитель  $a_{j2}$ , поэтому оно эквивалентно теореме Птолемея для четырёхугольника  $X_1 X_2 X_i X_j$ .

**5.1.** Для каждого вектора  $\varepsilon_j$  справедливо равенство  $\sum_{i=1}^n \cos^2 \alpha_{ij} = 1$ , поэтому  $\sum_{j=1}^k \sum_{i=1}^n \cos^2 \alpha_{ij} = k$ .

**5.2.** Пусть  $e_1, \dots, e_n$  — ортонормированный базис пространства  $V$ . Тогда  $v' = \sum_{i=1}^n (v, e_i)e_i$  и  $w' = \sum_{i=1}^n (w, e_i)e_i$ . Поэтому  $(v', w) = \sum_{i=1}^n (v, e_i)(w, e_i) = (v, w')$ .

5.3. Пусть  $U_1$  — ортогональное дополнение вектора  $u$  в  $U$ . Так как  $\dim U_1^\perp + \dim W = n - (m - 1) + m = n + 1$ , то  $\dim(U_1^\perp \cap W) \geq 1$ . Если  $w \in U_1^\perp \cap W$ , то  $w \perp U_1$  и  $w \perp u$ , поэтому  $w \perp U$ .

5.4. Применим к векторам  $x_1, \dots, x_k$  процесс ортогонализации с последующей нормировкой. Векторы  $e_1, \dots, e_k$  полученной ортонормированной системы линейно выражаются через  $x_1, \dots, x_k$ , причём коэффициенты зависят только от скалярных произведений  $(x_i, x_j)$ . Следовательно, векторы  $e_1, \dots, e_k$  линейно выражаются через  $x_1, \dots, x_k$ , причём коэффициенты зависят только от скалярных произведений  $(x_i, x_j)$ . Аналогичное утверждение верно и для выражений  $x_1, \dots, x_k$  через  $e_1, \dots, e_k$ . Пусть  $\varepsilon_1, \dots, \varepsilon_k$  — ортонормированная система, полученная при ортогонализации векторов  $y_1, \dots, y_k$ . Так как  $(x_i, x_j) = (y_i, y_j)$ , то векторы  $y_1, \dots, y_k$  выражаются через  $\varepsilon_1, \dots, \varepsilon_k$  с такими же коэффициентами, с какими векторы  $x_1, \dots, x_k$  выражаются через  $e_1, \dots, e_k$ . Поэтому ортогональный оператор, отображающий  $e_i$  в  $\varepsilon_i$  ( $i = 1, \dots, k$ ), отображает  $x_i$  в  $y_i$ .

5.5. Можно считать, что все векторы  $e_1, \dots, e_n$  имеют единичную длину. Тогда

$$0 \leq (e_1 + \dots + e_n)^2 < n - n(n-1) \sin \alpha,$$

поскольку  $(e_i, e_j) < \cos(90^\circ + \alpha) = -\sin \alpha$  при  $i \neq j$ . Поэтому  $0 < 1 - (n-1) \sin \alpha$ , т.е.  $n-1 < \frac{1}{\sin \alpha}$ .

5.6. Предположим, что  $\sum_{i=1}^m p_i(x_1, \dots, x_n) = \sum_{j=1}^n x_j^2$ , где  $p_i(x_1, \dots, x_n)$  — многочлен с вещественными коэффициентами от переменных  $x_1, \dots, x_n$ . Пусть  $p_i = c_{0i} + c_{1i} + \dots + c_{ri}$ , где  $c_{ki}$  — однородный многочлен степени  $k$  (здесь  $r$  — максимальная степень мономов, входящих в  $p_1, \dots, p_m$ ). Тогда  $\sum_{i=1}^m p_i^2 = d_0 + d_1 + \dots + d_{2r}$ , где  $d_0, \dots, d_{2r}$  — однородные многочлены степени  $0, \dots, 2r$ , причём  $d_0 = \sum_{i=1}^m c_{0i}^2$  и  $d_{2r} = \sum_{i=1}^m c_{ri}^2$ . Поэтому  $c_{0i} = 0$  при всех  $i$  и  $r = 1$ , т.е.  $p_i$  — однородная линейная функция  $\sum_{s=1}^n a_{si} x_s$ . Ясно, что  $\sum_{i=1}^m p_i^2 = \sum_{i=1}^m \sum_{s,t=1}^n a_{si} a_{ti} x_s x_t$ . Таким образом,  $\sum_{s,t=1}^n a_{si} a_{ti} = \delta_{st}$ . Это означает, что векторы  $e_j = (a_{j1}, \dots, a_{jm})$ ,  $j = 1, \dots, n$ , образуют ортонормированную систему. Количество ортонормированных векторов в  $m$ -мерном пространстве не превосходит  $m$ , поэтому  $n \leq m$ .

5.7. а) Пусть  $\hat{C}$  — объединение лучей, выходящих из точки  $O$  и идущих в точки  $C$ . Множество  $C$  выпукло и не содержит точку  $O$ , поэтому множества  $\hat{C}$  и  $-\hat{C}$  пересекаются только в точке  $O$ . Ясно также, что эти множества замкнуты. При  $n \geq 2$  топологическое пространство  $\mathbb{R}^n \setminus \{O\}$  связно, поэтому объединение его замкнутых непересекающихся подпространств  $\hat{C} \setminus \{O\}$  и  $-\hat{C} \setminus \{O\}$  не может совпадать со всем пространством  $\mathbb{R}^n \setminus \{O\}$ . Прямая, проходящая через точку  $O$  и точку, не принадлежащую этому объединению, — искомая.

б) Применим индукцию по  $n$ . При  $n = 2$  утверждение уже доказано (задача а). Для доказательства шага индукции мы снова воспользуемся задачей а). Проведём через точку  $O$  прямую  $l$ , не пересекающую  $C$ , и рассмотрим ортогональную проекцию на гиперплоскость  $V^{n-1} \subset \mathbb{R}^n$ , ортогональную прямой  $l$ . Пусть  $O'$  и  $C'$  — проекции на эту гиперплоскость прямой  $l$  и множества  $C$ . Согласно предположению индукции через точку  $O'$  можно провести гиперплоскость  $W^{n-2} \subset V^{n-1}$ , не пересекающую  $C'$ . Гиперплоскость в  $\mathbb{R}^n$ , натянутая на

$W^{n-2}$  и  $l$ , — искомая.

**5.8.** Выберем ортонормированный базис так, чтобы матрица данного ортогонального преобразования имела блочно диагональный вид с блоками  $\pm 1$  и  $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$  (теорема 3.1.3.3). По условию определитель этого преобразования положительный, поэтому элементов  $-1$  на диагонали чётное число. Эти элементы можно разбить на пары и рассматривать их как матрицы, соответствующие углу  $\varphi = \pi$ .

Если у нас есть два блока, соответствующие углам  $\varphi_1$  и  $\varphi_2$ , то рассмотрим соответствующие им двумерные подпространства (плоскости) и выберем в них прямые  $l_1, m_1$  и  $l_2, m_2$  так, что композиция симметрий первой плоскости относительно прямых  $l_1$  и  $m_1$  — это поворот на угол  $\varphi_1$ , а композиция симметрий второй плоскости относительно прямых  $l_2$  и  $m_2$  — это поворот на угол  $\varphi_2$ . Тогда в четырёхмерном пространстве (прямой сумме этих двух плоскостей) композиция симметрии относительно плоскости, натянутой на  $l_1$  и  $l_2$ , и симметрии относительно плоскости, натянутой на  $m_1$  и  $m_2$ , — это данное преобразование (на ортогональном дополнении к этому четырёхмерному пространству тождественно).

После того как мы последовательно представим требуемым образом такие пары блоков, у нас может остаться один блок. Но тогда мы должны представить требуемым способом преобразование, у которого на диагонали есть хотя бы одна единица. Рассмотрим соответствующее трёхмерное пространство, которое естественным образом представляется в виде прямой суммы плоскости и прямой. В этой плоскости мы выберем прямые  $l$  и  $m$  так, чтобы композиция симметрий относительно этих прямых была поворотом на данный угол. Тогда композиция симметрий относительно прямых  $l$  и  $m$  — это данное преобразование.

**5.9. Первое решение.** [Ne2] Пусть  $A$  — множество векторов в  $\mathbb{R}^n$ , имеющих неотрицательные координаты и длину не меньше 1;  $B$  — ортогональная проекция  $A$  на  $V$ . Множество  $A$  замкнуто, поэтому в  $A$  можно выбрать вектор  $v$  так, что его проекция  $v' = (v'_1, \dots, v'_n)$  имеет наименьшую длину. Докажем, что все координаты вектора  $v'$  неотрицательны. Предположим, что  $v'_i < 0$ . Рассмотрим вектор  $u = (0, \dots, -v'_i, \dots, 0)$  с ненулевой  $i$ -й координатой. Ясно, что  $\|v' + u\| < \|v'\|$ . Кроме того, если  $u'$  — ортогональная проекция на  $V$  вектора  $u$ , то  $\|v' + u'\| \leq \|v' + u\|$  (следствие 2 теоремы 2.5.3.2). Поэтому  $\|(v + u)'\| \leq \|v' + u'\| \leq \|v' + u\| < \|v'\|$ . Координаты векторов  $v$  и  $u$  неотрицательны, поэтому  $v + u \in A$ . Приходим к противоречию.

Итак, координаты вектора  $v'$  неотрицательны. Если  $v' = 0$ , то ненулевой вектор  $v$  с неотрицательными координатами лежит в  $V^\perp$ , а если  $v' \neq 0$ , то  $v'$  — ненулевой вектор с неотрицательными координатами, который лежит в  $V$ .

**Второе решение.** Пусть  $\Delta^{n-1}$  — множество векторов в  $\mathbb{R}^n$  с неотрицательными координатами, сумма которых равна 1. Предположим, что в пространстве  $V^\perp$  нет ненулевых векторов с неотрицательными координатами. Тогда, в частности, в  $V^\perp$  нет векторов из  $\Delta^{n-1}$ . Рассмотрим вектор  $v \in \Delta^{n-1}$ . Пусть  $v'$  и



$v''$  — его ортогональные проекции на  $V$  и  $V^\perp$ . Из того, что  $v \notin V^\perp$ , следует, что  $v' \neq 0$ . Это означает, что проекция выпуклого множества  $\Delta^{n-1}$  на подпространство  $V$  представляет собой выпуклое множество, не содержащее точку 0. Такое выпуклое множество лежит по одну сторону от некоторой гиперплоскости (задача 5.7), поэтому существует вектор  $w \in V$ , для которого  $(w, v') > 0$  для всех  $v \in \Delta^{n-1}$ . Ясно также, что  $(w, v) = (w, v')$ . Поэтому  $(w, v) > 0$  для всех  $v \in \Delta^{n-1}$ . В качестве  $v$  можно взять любой из векторов  $(1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ , поэтому все координаты вектора  $w \in V$  положительны.

**5.10.** а) Рассмотрим факторпространство  $\mathbb{R}^n/V = \mathbb{R}^k$ , где  $k \geq 2$ . Пусть  $D$  — образ в  $\mathbb{R}^k$  множества  $\Delta^{n-1}$ , состоящего из векторов в  $\mathbb{R}^n$  с неотрицательными координатами, сумма которых равна 1. Множество выпукло, замкнуто и не содержит начала координат. Поэтому согласно задаче 5.7 оно лежит по одну сторону от некоторой гиперплоскости в  $\mathbb{R}^k$ . Прообраз этой гиперплоскости в  $\mathbb{R}^n$  — искомая гиперплоскость.

б) Пусть  $V$  — линейное подпространство в  $\mathbb{R}^n$ , которое не содержит ненулевых векторов с неотрицательными координатами. Согласно задаче а) подпространство  $V$  содержится в некоторой гиперплоскости, обладающей тем же свойством. Такая гиперплоскость задаётся уравнением  $\sum a_i x_i = 0$ , где  $a_i \geq 0$ . Вектор  $(a_1, \dots, a_n)$  лежит в ортогональном дополнении к этой гиперплоскости, поэтому он лежит в  $V^\perp$ .

**5.11.** Рассмотрим в  $\mathbb{R}^{n+m}$  подпространство, состоящее из векторов вида  $(x, Ax^T)$ , и подпространство, состоящее из векторов вида  $(-A^T y^T, y)$ . Эти пространства ортогональны, поскольку скалярное произведение пары векторов из разных пространств равно  $-xA^T y^T + yAx^T = 0$ . Ясно также, что размерности этих подпространств равны  $n$  и  $m$ , поэтому они являются ортогональными дополнениями друг друга. Таким образом, можно воспользоваться результатом задачи 5.9 и получить либо ненулевой вектор  $(x, Ax^T)$  с неотрицательными координатами, либо ненулевой вектор  $(-A^T y^T, y)$  с неотрицательными координатами, что и требовалось.

**5.12.** Пусть  $X$  — множество тех  $\lambda$ , для которых все координаты вектора  $(A - \lambda J)x^T$  неотрицательны для некоторого ненулевого вектора  $x = (x_1, \dots, x_n)$  с неотрицательными координатами, а  $Y$  — множество тех  $\lambda$ , для которых все координаты вектора  $(-A^T + \lambda J^T)y^T$  неотрицательны для некоторого ненулевого вектора  $y = (y_1, \dots, y_m)$  с неотрицательными координатами. Ясно, что множества  $X$  и  $Y$  замкнуты. Кроме того, эти множества непустые, поскольку  $X$  содержит все отрицательные числа с достаточно большой абсолютной величиной, а  $Y$  содержит все достаточно большие положительные числа.

Согласно задаче 5.11, применённой к матрице  $A - \lambda J$ , любое число  $\lambda$  принадлежит  $X \cup Y$ . Поэтому из связности прямой  $\mathbb{R}$  следует, что  $X \cap Y \neq \emptyset$ , что и требовалось.

**5.13.** а) Пусть  $(a_i, b_i)$ ,  $i = 1, 2, 3, 4$ , — векторы сторон четырёхугольника, обладающего требуемым свойством. Как и при доказательстве теоремы 2.5.6.1, получаем  $a^2 = b^2$  и  $(a, b) = 0$ , где  $a = (a_1, \dots, a_4)$  и  $b = (b_1, \dots, b_4)$ . (И на-

оборот, если векторы  $a$  и  $b$  обладают этим свойством, то четырёхугольник обладает требуемым свойством.) Без ограничения общности можно считать, что

$a^2 = b^2 = 1$ . Тогда существует ортогональная матрица  $\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix}$ . Век-

торы  $e_i = (a_i, b_i, c_i, d_i)$  образуют ортонормированный базис. Рассмотрим вектор  $e = e_1 + e_2 + e_3 + e_4$ . При проекции на трёхмерное подпространство, ортогональное  $e$ , векторы  $e_1, \dots, e_4$  переходят в векторы, идущие из центра правильного тетраэдра в его вершины. Векторы  $(a_i, b_i)$  являются проекциями векторов  $e_i$  на плоскость, соответствующую координатам  $a$  и  $b$ . У вектора  $e$  координаты  $a$  и  $b$  нулевые, поскольку  $a_1 + \dots + a_4 = 0$  и  $b_1 + \dots + b_4 = 0$ . Поэтому указанная плоскость тоже лежит в трёхмерном пространстве, ортогональном  $e$ .

б) Решение аналогично.

**6.1.** а) Рекуррентное соотношение из §1.1.11 для определителя трёхдиагональной матрицы показывает, что определители указанных матриц удовлетворяют тому же самому рекуррентному соотношению, которому удовлетворяют многочлены Лежандра. Кроме того, при  $n = 1$  и  $n = 2$  определители равны многочленам.

б) и в) решаются аналогично.

**7.1.** Матрица  $A = B + iC$ , где  $B$  и  $C$  — вещественные матрицы, эрмитова тогда и только тогда, когда матрица  $B$  симметрическая, а матрица  $C$  кососимметрическая. В таком случае  $\begin{pmatrix} B & -C \\ C & B \end{pmatrix}^T = \begin{pmatrix} B^T & C^T \\ -C^T & B^T \end{pmatrix} = \begin{pmatrix} B & -C \\ C & B \end{pmatrix}$ .

**7.2.** Пусть  $p_A(\lambda)$  — характеристический многочлен матрицы  $A = B + iC$ , где  $B$  и  $C$  — вещественные матрицы. Легко проверить, что

$$\begin{vmatrix} B - \lambda I & -C \\ C & B - \lambda I \end{vmatrix} = |B - \lambda I + iC| \cdot |B - \lambda I - iC| = \\ = p_A(\lambda) \overline{p_A(\lambda)} = |p_A(\lambda)|^2$$

(см. §2.7.2).

**7.3.** Будем последовательно переставлять векторы  $ie_2, \dots, ie_n$  в базисе  $e_1, ie_1, \dots, e_n, ie_n$  так, чтобы получить базис  $e_1, \dots, e_n, ie_1, \dots, ie_n$ . Для этого нужно сделать  $1 + 2 + \dots + (n-1) = \frac{n(n-1)}{2}$  транспозиций.

**7.4.** Если  $V$  — вещественная форма  $W$ , то ясно, что  $V \cap iV = 0$ . Пусть  $V \cap iV = 0$ . Тогда  $W = V \oplus iV$ , т.е. любой вектор  $w \in W$  можно представить в виде суммы  $w = v_1 + iv_2$ , где  $v_1, v_2 \in V$ . Поэтому множество пар  $(v_1, v_2)$ , где  $v_1, v_2 \in V$ , с естественной операцией умножения на комплексные числа изоморфно  $W$ .

**7.5.** Пусть  $x_1 + iy_1, \dots, x_n + iy_n$  — базис  $W$ . Из равенства  $V^{\mathbb{C}} = W \oplus \bar{W}$  следует, что  $x_1, y_1, \dots, x_n, y_n$  — комплексный базис  $V^{\mathbb{C}}$ , т.е. вещественный базис  $V$ . Требуемая комплексная структура  $J: V \rightarrow V$  задаётся формулами  $Jx_k = -y_k$  и  $Jy_k = x_k$ ; ясно, что тогда  $J^{\mathbb{C}}(x_k + iy_k) = -y_k + ix_k = i(x_k + iy_k)$ .

**7.6.** Пусть  $a = a_1 + ia_2$ ,  $b = b_1 + ib_2$ , где  $a_k, b_k \in \mathbb{R}$ . Матрица указанного отображения относительно базиса  $1, i$  имеет вид  $\begin{pmatrix} a_1 + b_1 & -a_2 + b_2 \\ a_2 + b_2 & a_1 - b_1 \end{pmatrix}$ ; её определитель равен  $|a^2| - |b^2|$ .

**7.7.** Пусть  $e_1, e_2 = ie_1$  — базис  $V_{\mathbb{R}}$ . Если  $f(e_1) = \varepsilon$ , то  $f(e_2) = -i\varepsilon$ . Запишем вектор  $\varepsilon$  в виде  $\varepsilon = k(\cos \varphi e_1 + \sin \varphi e_2)$ . Тогда  $-i\varepsilon = k(\sin \varphi e_1 - \cos \varphi e_2)$ . Поэтому  $f$  является композицией гомотетии  $v \mapsto kv$  и преобразования, заданного матрицей  $\begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$ . Это преобразование является движением, изменяющим ориентацию, т.е. симметрией относительно некоторой прямой.

**7.8.** Пусть  $p = [n/2]$ . Комплексное подпространство, порождённое векторами  $e_1 + ie_2, e_3 + ie_4, \dots, e_{2p-1} + ie_{2p}$ , обладает требуемым свойством, так как если  $x = (a_1, ia_1, a_2, ia_2, \dots)$  и  $y = (b_1, ib_1, b_2, ib_2, \dots)$ , то  $B(x, y) = (a_1 b_1 - a_1 b_1) + \dots = 0$ .

**7.9.** Пусть  $J$  — ортогональная матрица оператора комплексной структуры. Тогда  $J^T = J^{-1}$  и  $J^2 = -I$ , т.е.  $J = -J^{-1}$ . Поэтому  $J^T = -J$ .

**8.1.** Пусть  $y_1 = x_1 + x_2$  и  $y_2 = x_1 - x_2$ . Тогда  $\max_{1 \leq i \leq 2} |y_i| = |x_1| + |x_2|$ . Это даёт требуемую изометрию в двумерном случае.

Докажем теперь, что при  $n \geq 3$  рассматриваемые пространства не изометричны. У изометричных пространств единичные шары переводятся друг в друга изоморфизмом пространств. В рассматриваемых пространствах единичные шары являются выпуклыми многогранниками. Ясно, что выпуклые многогранники, которые переводятся друг в друга изоморфизмом, имеют одинаковое число граней размерности  $n - 1$ . Для нормированного пространства из примера 2.8.0.2 такие грани задаются уравнениями  $\pm x_1 \pm x_2 \pm \dots \pm x_n = 1$ , поэтому их число равно  $2^n$ . Для нормированного пространства из примера 2.8.0.3 такие грани задаются уравнениями  $x_i = \pm 1$ , поэтому их число равно  $2n$ .

**8.2.** Воспользуемся теоремой 2.8.1.1. Векторы  $u + v$  и  $u - v$  лежат в подпространстве, порождённом векторами  $u$  и  $v$ . Поэтому тождество параллелограмма выполняется тогда и только тогда, когда оно выполняется в любом двумерном подпространстве.

**8.3.** Пусть  $\lambda x + \mu y + \nu z = 0$ , где  $\lambda, \mu, \nu$  — вещественные числа, хотя бы одно из которых отлично от нуля. Достаточно рассмотреть два случая.

**Случай 1.**  $\lambda, \mu, \nu \geq 0$ .

Если  $\mu = \nu = 0$ , то  $x = 0$ . В этом случае требуемое неравенство очевидно. Поэтому можно считать, что  $\nu > 0$  и  $\lambda \geq \mu, \lambda \geq \nu$ . Тогда  $z = \alpha x + \beta y$ , где  $\alpha = -\lambda/\nu \leq -1$  и  $\beta = -\mu/\nu \leq 0$  (при этом  $\alpha \leq \beta$ ). Положим  $l = -1/\alpha = \nu/\lambda$  и  $k = (\alpha - \beta)/\alpha = (\lambda - \mu)/\lambda$ . Тогда  $0 \leq k, l \leq 1$  и выполняются равенства  $ky - lz = x + y$  и  $(1 - k)y - (1 - l)z = -(z + x)$ . Поэтому

$$\begin{aligned} \|x + y\| &\leq k\|y\| + l\|z\|, \\ \|z + x\| &\leq (1 - k)\|y\| + (1 - l)\|z\|. \end{aligned}$$

Следовательно,  $\|x + y\| + \|z + x\| \leq \|y\| + \|z\|$ . Остаётся заметить, что  $\|y + z\| \leq \|x\| + \|x + y + z\|$ .

**Случай 2.**  $\mu \geq \lambda \geq 0 > \nu$ .

В этом случае  $z = \alpha x + \beta y$ , где  $\alpha = -\lambda/\nu \geq 0$  и  $\beta = -\mu/\nu \geq 0$  (при этом  $\alpha \leq \beta$ ). Положим  $l = 1/(1 + \beta)$  и  $k = (\beta - \alpha)/(1 + \beta)$ . Тогда  $0 \leq k, l \leq 1$  и выполняются равенства  $kx + l(x + y + z) = x + y$  и  $(1 - k)x + (1 - l)(x + y + z) = x + z$ . Поэтому  $\|x + y\| + \|z + x\| \leq \|x\| + \|x + y + z\|$ . Остаётся заметить, что  $\|y + z\| \leq \|y\| + \|z\|$ .

**8.4.** Пусть  $x' = x - z$ ,  $y' = y - z$ ,  $w' = w - z$ . Тогда  $d(w, z) = d(w', 0)$ ,  $d(w, x) = d(w', x')$  и  $d(w, y) = d(w', y')$ . Поэтому достаточно рассмотреть случай, когда  $z = 0$ .

Рассмотрим сначала случай, когда  $a = 1/2$ , т.е. для любого  $w$  имеет место неравенство  $2d(w, 0) \leq d(w, x) + d(w, y)$ . Пусть  $T$  — множество всех целых неотрицательных  $t$ , для которых

$$(2m + t)d(x + y, 0) \leq d(mx + my, x) + d(mx + my, y) \quad (2.1)$$

для всех натуральных  $m$ . Докажем, что  $T$  содержит все целые неотрицательные числа  $t$ . При  $t = 0$  достаточно применить исходное неравенство для  $w = mx + my$ . Остаётся доказать, что если  $t \in T$ , то  $t + 1 \in T$ . Ясно, что если число  $m$  натуральное, то число  $2m$  тоже натуральное, поэтому

$$(4m + t)d(x + y, 0) \leq d(2mx + 2my, x) + d(2mx + 2my, y). \quad (2.2)$$

Подставим в исходное неравенство  $w = (2m - 1)x + 2my$ . После несложных преобразований оно перепишется в виде

$$d(2mx + 2my, x) \leq d(mx + my, x) + \left(m - \frac{1}{2}\right) d(x + y, 0). \quad (2.3)$$

Аналогично

$$d(2mx + 2my, y) \leq d(mx + my, y) + \left(m - \frac{1}{2}\right) d(x + y, 0). \quad (2.4)$$

Подставляя неравенства (2.3) и (2.4) в неравенство (2.2), получаем

$$(4m + t)d(x + y, 0) \leq d(mx + my, x) + d(mx + my, y) + (2m - 1)d(x + y, 0),$$

т.е.

$$(2m + t + 1)d(x + y, 0) \leq d(mx + my, x) + d(mx + my, y).$$

Таким образом,  $t + 1 \in T$ .

При  $m = 1$  неравенство (2.1) принимает вид

$$(2 + t)d(x + y, 0) \leq d(y, 0) + d(x, 0).$$

Мы доказали, что такое неравенство имеет место для всех целых неотрицательных  $t$ . При  $t \rightarrow \infty$  получаем  $d(x + y, 0) \leq 0$ , т.е.  $x + y = 0$ , что и требовалось.

Пусть теперь  $a$  — произвольное число между 0 и 1. При необходимости меняя местами  $x$  и  $y$ , можно добиться, чтобы число  $a$  удовлетворяло неравенствам  $0 < a < 1/2$ . Тогда если  $d(w, 0) \leq ad(w, x) + (1 - a)d(w, y)$ , то

$$\begin{aligned} d(w, 0) &\leq d(aw, ax) + d(w - aw, y - ay) \leq \\ &\leq d(aw, ax) + d(w - aw, aw) + d(aw, y - ay) = \\ &\leq d(aw, ax) + (1 - 2a)d(w, 0) + d(aw, y - ay), \end{aligned}$$

т.е.  $2d(aw, 0) \leq d(aw, ax) + d(aw, y - ay)$ . Это неравенство имеет место для всех  $w$ . Поэтому, как было доказано выше,  $ax + y - ay = 0$ , что и требовалось.

## Глава 3.

# Канонические формы матриц и линейных операторов

### 3.1. След и собственные значения оператора

#### 3.1.1. След

Следом квадратной матрицы  $A$  называют сумму её диагональных элементов. След матрицы  $A$  обозначают  $\text{tr } A$  (от английского trace) или  $\text{Sp } A$  (от немецкого Spur). Легко проверить, что  $\text{tr } AB = \sum_{i,j} a_{ij}b_{ji} = \text{tr } BA$ . Поэтому  $\text{tr } P^{-1}AP = \text{tr } PP^{-1}A = \text{tr } A$ , т.е. след матрицы линейного оператора не зависит от выбора базиса.

Равенство  $\text{tr } ABC = \text{tr } ACB$  выполняется не всегда. Например, если  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  и  $C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ , то  $ABC = 0$  и  $ACB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ .

Для следа оператора  $A$  в евклидовом пространстве имеется следующая полезная формула.

**Теорема 3.1.1.1.** Пусть  $e_1, \dots, e_n$  — ортонормированный базис. Тогда

$$\text{tr } A = \sum_{i=1}^n (Ae_i, e_i).$$

**Доказательство.** Так как  $Ae_i = \sum_j a_{ji}e_j$ , то  $(Ae_i, e_i) = a_{ii}$ .  $\square$

**Замечание.** След оператора инвариантен, но приведённое выше определение следа использует базис, т.е. оно не инвариантно. Можно, однако, дать и инвариантное определение следа оператора (теорема 5.1.3.2).

### 3.1.2. Собственные векторы и собственные значения

Ненулевой вектор  $v \in V$  называют *собственным вектором* линейного оператора  $A : V \rightarrow V$ , если  $Av = \lambda v$ ; число  $\lambda$  называют при этом *собственным значением* оператора  $A$ . Фиксируем  $\lambda$  и рассмотрим уравнение  $Av = \lambda v$ , т.е.  $(A - \lambda I)v = 0$ . Это уравнение имеет ненулевое решение  $v$  тогда и только тогда, когда  $|A - \lambda I| = 0$ . Поэтому собственными значениями оператора  $A$  являются корни многочлена  $p(\lambda) = |A - \lambda I|$ ; этот многочлен имеет степень  $n$ , где  $n$  — порядок матрицы  $A$ . Любой оператор над полем  $\mathbb{C}$  имеет собственный вектор, и любой оператор в нечётномерном пространстве над полем  $\mathbb{R}$  тоже имеет собственный вектор.

Набор собственных значений оператора  $A$  называют его *спектром*.

Многочлен  $p(\lambda) = |A - \lambda I|$  называют *характеристическим многочленом* оператора  $A$ ; этот многочлен зависит лишь от самого оператора и не зависит от выбора базиса (см. §2.3.1). Иногда характеристическим многочленом называют многочлен  $|\lambda I - A|$ .

**Теорема 3.1.2.1.** Если  $Ae_1 = \lambda_1 e_1, \dots, Ae_k = \lambda_k e_k$ , причём числа  $\lambda_1, \dots, \lambda_k$  попарно различны, то векторы  $e_1, \dots, e_k$  линейно независимы.

**Доказательство.** Предположим противное. Выбрав минимальную линейно зависимую систему векторов, будем считать, что  $e_k = \alpha_1 e_1 + \dots + \alpha_{k-1} e_{k-1}$ , причём  $\alpha_1 \dots \alpha_{k-1} \neq 0$  и векторы  $e_1, \dots, e_{k-1}$  линейно независимы. Тогда  $Ae_k = \alpha_1 \lambda_1 e_1 + \dots + \alpha_{k-1} \lambda_{k-1} e_{k-1}$  и  $Ae_k = \lambda_k e_k = \alpha_1 \lambda_k e_1 + \dots + \alpha_{k-1} \lambda_k e_{k-1}$ . Следовательно,  $\lambda_1 = \lambda_k$ . Получено противоречие.  $\square$

**Замечание.** Если характеристический многочлен оператора  $A$  над полем  $\mathbb{C}$  не имеет кратных корней, то собственные векторы  $A$  образуют базис.

### 3.1.3. Диагонализируемые операторы

Линейный оператор  $A$ , имеющий базис собственных векторов, называют *диагонализируемым* или *полупростым*. Если  $X$  — матрица, составленная из столбцов координат собственных векторов  $x_1, \dots, x_n$ , а  $\lambda_i$  — собственное значение, соответствующее  $x_i$ , то  $AX = X\Lambda$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Поэтому  $X^{-1}AX = \Lambda$ . Верно и обратное: если  $X^{-1}AX =$

$\text{diag}(\lambda_1, \dots, \lambda_n)$ , то  $\lambda_1, \dots, \lambda_n$  — собственные значения оператора  $A$ , а столбцы матрицы  $X$  — соответствующие им собственные векторы.

Над полем  $\mathbb{C}$  недиагонализуемым может быть лишь оператор, имеющий кратные собственные значения, а такие операторы образуют множество меры нуль.

Докажем, что над полем  $\mathbb{C}$  диагонализуемы следующие важные классы операторов: унитарные и эрмитовы.

**Замечание.** Более общее утверждение доказано в §3.7.1.

Собственные значения эрмитова оператора  $A$  вещественны. В самом деле, если  $Ax = \lambda x$ , то  $(Ax, x) = \lambda(x, x)$ , поэтому  $\lambda = (Ax, x)/(x, x)$ . А так как  $(Ax, x) = (x, Ax) = \overline{(Ax, x)}$ , то  $(Ax, x) \in \mathbb{R}$ . Собственные значения унитарного оператора  $A$  равны по модулю 1, так как  $\|Ax\| = \|x\|$ .

Рассматривая комплексификации симметрического и ортогонального оператора над полем  $\mathbb{R}$ , получаем, что для них справедливы аналогичные утверждения.

**Теорема 3.1.3.1.** Если  $A$  — унитарный или эрмитов оператор и  $AW \subset W$ , то  $AW^\perp \subset W^\perp$ .

**Доказательство.** Если  $A$  — унитарный оператор, то  $AW = W$ , а значит,  $A^{-1}W = W$ . Поэтому  $A^*W \subset W$  как для унитарного, так и для эрмитова оператора  $A$ . Пусть  $y \in W^\perp$  и  $w \in W$ . Тогда  $(Ay, w) = (y, A^*w) = 0$ , так как  $A^*w \in W$ . Следовательно,  $Ay \in W^\perp$ .  $\square$

Доказательство теоремы 3.1.3.1 почти без изменений проходит для симметрического или ортогонального оператора  $A$  над полем  $\mathbb{R}$ .

**Теорема 3.1.3.2.** Для эрмитова или унитарного оператора  $A$  существует ортонормированный базис собственных векторов.

**Доказательство.** Применим индукцию по размерности  $n$  пространства  $V$ . При  $n = 1$  утверждение очевидно. Пусть  $e$  — собственный вектор оператора  $A$ . Тогда подпространство  $\langle e \rangle^\perp$  инвариантно относительно  $A$ , поэтому можно рассмотреть ограничение  $A$  на это подпространство. Оператор эрмитов (унитарен) тогда и только тогда, когда в каком-либо ортонормированном базисе его матрица эрмитова (унитарна). Поэтому, дополнив вектор  $e$  до ортонормированного базиса пространства  $V$ , получим, что ограничение оператора  $A$  на подпространство  $\langle e \rangle^\perp$  тоже является эрмитовым (унитарным) оператором. Следовательно, по предположению индукции для оператора  $A$  существует ортонормированный базис собственных векторов.  $\square$



**Замечание 1.** Собственные значения симметрического оператора над полем  $\mathbb{R}$  вещественны, поэтому для симметрического оператора доказательство теоремы 3.1.3.2 проходит без изменений.

**Замечание 2.** Если для некоторого оператора существует ортонормированный базис  $e_1, \dots, e_n$  с собственными значениями  $\lambda_1, \dots, \lambda_n$ , то матрицу  $A$  можно представить в виде  $A = U\Lambda U^*$ , где  $U$  — унитарная матрица, столбцами которой служат координаты векторов  $e_1, \dots, e_n$ , а  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Действительно,  $(U^*AU)_{ij} = e_i^* A e_j = \lambda_j e_i^* e_j = \lambda_j \delta_{ij}$ , поэтому  $U^*AU = \Lambda$ , т.е.  $A = U\Lambda U^*$ .

**Теорема 3.1.3.3.** Для ортогонального оператора  $A$  существует ортонормированный базис, в котором его матрица имеет блочно диагональный вид с блоками  $\pm 1$  и  $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ .

**Доказательство.** В случае, когда есть собственный вектор (соответствующий собственному значению  $\pm 1$ ), доказательство аналогично доказательству теоремы 3.1.3.2. Предположим поэтому, что векторы  $x$  и  $Ax$  не параллельны для всех  $x$ . Функция  $\varphi(x) = \angle(x, Ax)$  (угол между векторами  $x$  и  $Ax$ ) непрерывна на компакте — единичной сфере. Пусть  $\varphi_0 = \angle(x_0, Ax_0)$  — минимум функции  $\varphi(x)$ ,  $e$  — вектор, параллельный биссектрисе угла между векторами  $x_0$  и  $Ax_0$ . Тогда  $\varphi \leq \angle(e, Ae) \leq \angle(e, Ax_0) + \angle(Ax_0, Ae) = \frac{\varphi_0}{2} + \frac{\varphi_0}{2}$ , а значит, вектор  $Ae$  лежит в плоскости  $\langle Ax_0, e \rangle = \langle x_0, e \rangle$ . Она инвариантна относительно  $A$ , так как  $Ax_0, Ae \in \langle x_0, e \rangle$ .

Ортогональное преобразование плоскости является поворотом или симметрией относительно прямой, но симметрия имеет собственные значения 1 и  $-1$ , поэтому матрица оператора  $A$  на пространстве  $\langle x_0, e \rangle$  имеет вид  $\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ , где  $\sin \varphi \neq 0$ .  $\square$

### 3.1.4. Спектр матрицы Якоби

Интересными спектральными свойствами обладает матрица Якоби

$$J = (a_{ij}) = \begin{pmatrix} a_1 & -b_1 & 0 & \dots & 0 \\ -c_1 & a_2 & -b_2 & \dots & 0 \\ 0 & -c_2 & a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix}, \text{ где } b_i c_i > 0.$$

Собственные значения этой матрицы вещественны и не кратны. Рассмотрим для матрицы  $J$  последовательность многочленов  $D_k(\lambda) = |a_{ij} - \lambda \delta_{ij}|_1^k$  ( $D_0(\lambda) = 1$ ,  $D_n(\lambda)$  — характеристический многочлен  $J$ ). Эти многочлены удовлетворяют рекуррентному соотношению

$$D_k(\lambda) = (a_k - \lambda)D_{k-1}(\lambda) - b_{k-1}c_{k-1}D_{k-2}(\lambda) \quad (1)$$

(см. § 1.1.11), поэтому характеристический многочлен  $D_n(\lambda)$  зависит не от самих чисел  $b_k, c_k$ , а лишь от их произведений. Заменяя в матрице  $J$  элементы  $b_k$  и  $c_k$  на  $\sqrt{b_k c_k}$ , получим симметрическую матрицу  $J'$  с тем же самым характеристическим многочленом. Следовательно, собственные значения матрицы  $J$  вещественны.

Симметрическая матрица имеет базис собственных векторов, поэтому остаётся доказать, что каждому собственному значению  $\lambda$  матрицы  $J'$  соответствует не более одного собственного вектора  $(x_1, \dots, x_n)$ . Это верно даже для матрицы  $J$ . Так как

$$\begin{aligned} (a_1 - \lambda)x_1 - b_1x_2 &= 0, \\ -c_1x_1 + (a_2 - \lambda)x_2 - b_2x_3 &= 0, \\ &\dots\dots\dots \\ -c_{n-2}x_{n-2} + (a_{n-1} - \lambda)x_{n-1} - b_{n-1}x_n &= 0, \\ -c_{n-1}x_{n-1} + (a_n - \lambda)x_n &= 0, \end{aligned}$$

то, сделав замену  $y_1 = x_1, \dots, y_2 = b_1x_2, \dots, y_k = b_1 \dots b_{k-1}x_k$ , получим

$$\begin{aligned} y_2 &= (a_1 - \lambda)y_1, \\ y_3 &= (a_2 - \lambda)y_2 - c_1b_1y_1, \\ &\dots\dots\dots \\ y_n &= (a_{n-1} - \lambda)y_{n-1} - c_{n-2}b_{n-2}y_{n-2}. \end{aligned}$$

Эти соотношения для  $y_k$  совпадают с соотношениями (1) для  $D_k$ , поэтому если  $y_1 = c = cD_0(\lambda)$ , то  $y_k = cD_k(\lambda)$ . Следовательно, вектор  $(x_1, \dots, x_k)$  определен однозначно (с точностью до пропорциональности).

### 3.1.5. Собственные значения полинома от матрицы

Для вычисления собственных значений матрицы может оказаться полезным следующее утверждение.

**Теорема 3.1.5.1.** Если  $\lambda$  — собственное значение матрицы  $A$ , а  $f$  — произвольный многочлен, то  $f(\lambda)$  — собственное значение матрицы  $f(A)$ .

**Доказательство.** Требуется доказать, что если  $|A - \lambda I| = 0$ , то  $|f(A) - f(\lambda)I| = 0$ . Для этого достаточно проверить, что  $f(A) - f(\lambda)I = (A - \lambda I)g(A)$ , где  $g$  — некоторый многочлен. Это легко проверяется почленно:  $A^n - (\lambda I)^n = (A - \lambda I)(A^{n-1} + \lambda A^{n-2} + \dots + \lambda^{n-1}I)$ .  $\square$

**Следствие.** Если  $\lambda$  — собственное значение матрицы  $A$  и  $f(A) = 0$ , то  $f(\lambda) = 0$ .

**Доказательство.** Число  $f(\lambda)$  является собственным значением матрицы  $f(A) = 0$ , поэтому  $f(\lambda) = 0$ .  $\square$

В качестве примера применения теоремы 3.1.5.1, вычислим собственные значения циркулянта.

Рассмотрим сначала простейший циркулянт

$$C_n = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Так как  $C_n e_k = e_{k+1}$ , то  $C_n^s e_k = e_{k+s}$ . Поэтому  $C_n^n = I$ . Следовательно, собственные значения являются корнями уравнения  $x^n = 1$ . Пусть  $\varepsilon = \exp(2\pi i/n)$ . Докажем, что векторы  $u_s = \sum_{k=1}^n \varepsilon^{ks} e_k$ , ( $s = 1, \dots, n$ ) являются собственными векторами  $C_n$ , соответствующими собственным значениям  $\varepsilon^{-s}$ . В самом деле,  $C_n u_s = \sum \varepsilon^{ks} C_n e_k = \sum \varepsilon^{ks} e_{k+1} = \sum \varepsilon^{-s} \varepsilon^{s(k+1)} e_{k+1} = \varepsilon^{-s} u_s$ .

**Теорема 3.1.5.2.** Пусть  $\varepsilon_1, \dots, \varepsilon_n$  — попарно различные корни степени  $n$  из единицы и  $f(x) = b_0 + b_1 x + b_2 x^2 + \dots + b_{n-1} x^{n-1}$ . Тогда собственные значения циркулянта  $A = b_0 I + b_1 C_n + b_2 C_n^2 + \dots + b_{n-1} C_n^{n-1}$ , равны  $f(\varepsilon_1), \dots, f(\varepsilon_n)$ .

**Доказательство.** Мы уже убедились, что собственные значения матрицы  $C_n$  равны  $\varepsilon_1, \dots, \varepsilon_n$ . Воспользовавшись теоремой 3.1.5.1, получаем требуемое.  $\square$

В качестве другого примера вычислим собственные значения и собственные векторы сопровождающей матрицы

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ p_0 & p_1 & p_2 & \dots & p_{n-1} \end{pmatrix}$$

многочлена  $\lambda^n - p_{n-1}\lambda^{n-1} - \dots - p_0$ . Пусть  $x$  — столбец  $(x_1, \dots, x_n)$ . Уравнение  $Ax = \lambda x$  можно переписать в виде  $x_2 = \lambda x_1$ ,  $x_3 = \lambda x_2$ ,  $\dots$ ,  $x_n = \lambda x_{n-1}$ . Поэтому собственные векторы матрицы  $A$  имеют вид  $(\alpha, \lambda\alpha, \lambda^2\alpha, \dots, \lambda^{n-1}\alpha)$ , где  $p_0 + p_1\lambda + \dots + p_{n-1}\lambda^{n-1} = \lambda^n$ .

### 3.1.6. Собственные значения матриц $AB$ и $BA$

Если  $A$  — матрица размером  $m \times n$ , а  $B$  — матрица размером  $n \times m$ , то след матрицы  $AB$  равен следу матрицы  $BA$ . Оказывается, что справедливо более сильное утверждение: у этих матриц одинаковые ненулевые собственные значения. Мы сначала рассмотрим случай квадратных матриц, а затем — общий случай.

**Теорема 3.1.6.1.** Пусть  $A$  и  $B$  — квадратные матрицы одного порядка. Тогда характеристические многочлены матриц  $AB$  и  $BA$  совпадают.

**Доказательство.** Предположим сначала, что матрица  $A$  невырожденная. Тогда  $|AB - \lambda I| = |A^{-1}(AB - \lambda I)A| = |BA - \lambda I|$ . Ясно также, что многочлен  $P(\lambda, A, B) = |AB - \lambda I|$  непрерывно зависит от элементов матриц  $A$  и  $B$ , поэтому равенство  $P(\lambda, A, B) = P(\lambda, B, A)$  выполняется и для вырожденных матриц. Действительно, любая вырожденная матрица является пределом невырожденных матриц, поскольку матрица  $A + \lambda I$  невырожденная для почти всех  $\lambda$ .  $\square$

Обратимся теперь к случаю прямоугольных матриц.

**Теорема 3.1.6.2.** Пусть  $A$  — матрица размером  $m \times n$ , а  $B$  — матрица размером  $n \times m$ , причём  $m \leq n$ . Тогда  $(-\lambda)^{n-m}|AB - \lambda I_m| = |BA - \lambda I_n|$ .

**Первое доказательство.** Дополним матрицы  $A$  и  $B$  нулями до квадратных матриц порядка  $n$ . К полученным квадратным матрицам можно применить теорему 3.1.6.1. В результате получим требуемое.  $\square$

**Второе доказательство.** Если  $\lambda \neq 0$ , то  $|AB - \lambda I_m| = (-\lambda)^m |I_m - \frac{1}{\lambda}AB|$  и  $|BA - \lambda I_n| = (-\lambda)^n |I_n - \frac{1}{\lambda}BA|$ . Поэтому согласно задаче 3.2  $(-\lambda)^{n-m}|AB - \lambda I_m| = |BA - \lambda I_n|$ . По непрерывности последнее равенство выполняется и при  $\lambda = 0$ .  $\square$

### 3.1.7. Собственные значения и суммы элементов столбцов

**Теорема 3.1.7.1.** Предположим, что сумма элементов каждого столбца квадратной матрицы  $A$  равна 1. Пусть столбец  $(x_1, \dots, x_n)^T$  является собственным вектором матрицы  $A$ , причём  $x_1 + \dots + x_n \neq 0$ . Тогда этому собственному вектору соответствует собственное значение 1.

**Доказательство.** Складывая равенства  $\sum a_{1j}x_j = \lambda x_1, \dots, \sum a_{nj}x_j = \lambda x_n$ , получаем  $\sum_{i,j} a_{ij}x_j = \lambda \sum_j x_j$ . Но

$$\sum_{i,j} a_{ij}x_j = \sum_j \left( x_j \sum_i a_{ij} \right) = \sum_j x_j,$$

поскольку  $\sum_i a_{ij} = 1$ . Таким образом,  $\lambda \sum_j x_j = \sum_j x_j$ , причём  $\sum_j x_j \neq 0$ . Следовательно,  $\lambda = 1$ .  $\square$

**Теорема 3.1.7.2.** Если сумма абсолютных величин элементов каждого столбца квадратной матрицы  $A$  не превосходит 1, то все её собственные значения не превосходят 1.

**Доказательство.** Пусть  $(x_1, \dots, x_n)$  — собственный вектор, соответствующий собственному значению  $\lambda$ . Тогда

$$|\lambda x_i| = \left| \sum_j a_{ij}x_j \right| \leq \sum_j |a_{ij}| \cdot |x_j|, \quad i = 1, \dots, n.$$

Складывая эти неравенства, получаем

$$|\lambda| \sum_i |x_i| \leq \sum_{ij} |a_{ij}| \cdot |x_j| \leq \sum_j \left( |x_j| \sum_i |a_{ij}| \right) \leq \sum_j |x_j|,$$

поскольку  $\sum_i |a_{ij}| \leq 1$ . Поделив обе части этого неравенства на ненулевое число  $\sum |x_j|$ , получим  $|\lambda| \leq 1$ .  $\square$

**Замечание.** Теорема 3.1.7.2 остаётся верной и в том случае, когда некоторые столбцы матрицы  $A$  нулевые.

**Теорема 3.1.7.3.** Пусть  $A = \|a_{ij}\|_1^n$ ,  $S_j = \sum_{i=1}^n |a_{ij}|$ . Тогда  $\sum_{j=1}^n S_j^{-1} |a_{jj}| \leq \text{rk } A$ , где слагаемые, соответствующие нулевым значениям  $S_j$ , можно заменить нулями.

**Доказательство.** Умножая столбцы матрицы  $A$  на некоторые ненулевые элементы, мы всегда можем добиться, чтобы числа  $S_j$  для новой матрицы были равны 0 или 1, причём  $a_{jj} \geq 0$ . Ранг матрицы при таких преобразованиях не изменяется. Применяя теорему 3.1.7.2 к новой матрице, получаем

$$\sum |a_{jj}| = \sum a_{jj} = \text{tr } A = \sum \lambda_i \leq \sum |\lambda_i| \leq \text{rk } A.$$

$\square$

### 3.1.8. Матрица Каца

Матрицей Каца называют трёхдиагональную матрицу

$$\text{Кас}_n = \begin{pmatrix} 0 & n & & & \\ 1 & 0 & n-1 & & \\ & 2 & 0 & n-2 & \\ \dots & \dots & \dots & \dots & \dots \\ & & n-1 & 0 & 1 \\ & & & n & 0 \end{pmatrix}.$$

**Теорема 3.1.8.1.** Собственные значения матрицы  $\text{Кас}_n$  равны  $-n$ ,  $-n+2$ ,  $-n+4$ ,  $\dots$ ,  $n-2$ ,  $n$ .

**Доказательство.** [Ed] Положим  $f_k(x) = \text{sh}^k x \text{ch}^{n-k} x$  для  $k = 0, \dots, n$  и рассмотрим пространство  $V$ , состоящее из линейных комбинаций функций  $f_k(x)$ . Ясно, что функции

$$g_k(x) = (\text{sh } x + \text{ch } x)^k (\text{ch } x - \text{sh } x)^{n-k}, \quad k = 0, \dots, n,$$

лежат в пространстве  $V$ .

Равенство

$$\frac{d}{dx} f_k(x) = k f_{k-1}(x) + (n-k) f_{k+1}(x)$$

показывает, что матрица  $\text{Кас}_n$  является матрицей оператора  $\frac{d}{dx}$  в пространстве  $V$  относительно базиса  $\{f_k\}$ . Легко проверить, что  $g_k(x) = e^{(2k-n)x}$ , поэтому  $g_k(x)$  — собственная функция оператора  $\frac{d}{dx}$  с собственным значением  $2k - n$ .  $\square$

### 3.1.9. Теорема Годдарда—Шнайдера

Пусть  $f(x, y)$  — многочлен от некоммутирующих переменных  $x$  и  $y$ ;  $f_0(x) = f(x, 0)$ . Пусть, далее, матрицы  $A$ ,  $B$  и  $X$  размера  $n \times n$ ,  $m \times m$  и  $n \times m$  соответственно таковы, что  $AX = XB$ , причём  $\text{rk } X = r$ . Легко проверить, что тогда  $f(A, XK)X = Xf(B, KX)$  для любой матрицы  $K$  размера  $m \times n$ .

**Теорема 3.1.9.1.** Характеристические многочлены матриц  $f(A, XK)$  и  $f(B, KX)$  имеют вид  $\pm \theta(\lambda) p(\lambda)$  и  $\pm \theta(\lambda) q(\lambda)$ , где многочлены  $\theta$ ,  $p$  и  $q$  имеют степени  $r$ ,  $n - r$  и  $m - r$  соответственно, причём многочлены  $p$  и  $q$  не зависят от  $K$  и являются делителями характеристических многочленов матриц  $f_0(A)$  и  $f_0(B)$ .

**Доказательство.** [Go] Для  $r = 0$  утверждение очевидно. Пусть  $0 < r < \min(n, m)$ . Существуют такие невырожденные матрицы  $P$  и  $Q$  порядка  $n$  и  $m$ , что  $PXQ^{-1} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$  (теорема 2.2.3.2). Пусть  $C = PAP^{-1}$ ,  $D = QBQ^{-1}$  и  $L = QKP^{-1}$ . Запишем матрицу  $C$  в блочном виде  $\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$ , где  $C_{11}$  — матрица порядка  $r$ ; матрицы  $D$  и  $L$  разобьём на блоки аналогичным образом. Из равенства  $AX = XB$  следует, что  $CY = YD$ , т.е.  $C_{11} = D_{11}$ ,  $C_{21} = 0$  и  $D_{12} = 0$ . Поэтому  $PXKP^{-1} = YL = \begin{pmatrix} L_{11} & L_{12} \\ 0 & 0 \end{pmatrix}$  и  $QKXQ^{-1} = LY = \begin{pmatrix} L_{11} & 0 \\ L_{21} & 0 \end{pmatrix}$ , следовательно,  $Pf(A, XK)P^{-1} = f(C, YL) = \begin{pmatrix} f(C_{11}L_{11}) & * \\ 0 & f(C_{22}, 0) \end{pmatrix}$  и  $Qf(B, KX)Q^{-1} = f(D, LY) = \begin{pmatrix} f(D_{11}L_{11}) & 0 \\ * & f(D_{22}, 0) \end{pmatrix}$ . В качестве  $\theta$ ,  $p$  и  $q$  возьмём характеристические многочлены матриц  $f(C_{11}L_{11})$ ,  $f(C_{22}, 0)$  и  $f(D_{22}, 0)$ . Ясно, что

$$Pf_0(A)P^{-1} = f_0(C) = \begin{pmatrix} f_0(C_{11}) & * \\ 0 & f_0(C_{22}) \end{pmatrix}$$

и

$$Qf_0(B)Q^{-1} = f_0(D) = \begin{pmatrix} f_0(D_{11}) & 0 \\ * & f_0(D_{22}) \end{pmatrix}.$$

При  $r = \min(n, m)$  блочная запись имеет вид  $(C_{11}, C_{12})$  или  $\begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix}$  и один из многочленов  $p$  и  $q$  равен  $\pm 1$ .  $\square$

**Следствие 1.** Если  $AX = XB$  и  $\text{rk } X = r$ , то характеристические многочлены матриц  $A$  и  $B$  имеют общий делитель степени  $r$ .

Пусть  $X$  — матрица размера  $n \times m$ , составленная из линейно независимых собственных векторов матрицы  $A$ . Тогда  $AX = X\Lambda$ , где  $\Lambda$  — диагональная матрица порядка  $m$ , составленная из собственных значений матрицы  $A$ .

**Следствие 2.** Для любой матрицы  $K$  размера  $m \times n$  собственные значения матрицы  $\Lambda + KX$  являются собственными значениями матрицы  $A + XK$ .

**Доказательство.** Достаточно рассмотреть многочлен  $f(x, y) = x + y$ .  $\square$

### Задачи

**1.1.** Пусть  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_n$  — ортонормированные базисы пространства  $V$ ,  $A : V \rightarrow V$  — линейный оператор. Докажите, что  $\sum_i (Ae_i, e_i) = \sum_i (A\varepsilon_i, \varepsilon_i)$ .

**1.2.** а) Существуют ли такие матрицы  $A$  и  $B$ , что  $AB - BA = I$ ?

б) Докажите, что если  $AB - BA = A$ , то  $A$  — вырожденная матрица.

**1.3.** Докажите, что если  $\text{rk } A = 1$  и  $\text{tr } AB = 0$ , то  $ABA = 0$ .

**1.4.** Найдите собственные значения и собственные векторы матрицы  $A = \|a_{ij}\|$ , где  $a_{ij} = \lambda_i/\lambda_j$ .

**1.5.** Докажите, что любая квадратная матрица  $A$  является суммой двух невырожденных матриц.

**1.6.** Докажите, что собственные значения матрицы непрерывно зависят от её элементов, точнее говоря, для любого  $\varepsilon > 0$  существует такое  $\delta > 0$ , что если  $|a_{ij} - b_{ij}| < \delta$  для всех  $i, j$  и  $\lambda$  — собственное значение данной матрицы  $A$ , то у матрицы  $B$  есть такое собственное значение  $\mu$ , что  $|\lambda - \mu| < \varepsilon$ .

**1.7.** Сумма элементов каждой строки невырожденной матрицы  $A$  равна  $s$ . Докажите, что сумма элементов каждой строки матрицы  $A^{-1}$  равна  $1/s$ .

**1.8.** а) Докажите, что если первый столбец матрицы  $S^{-1}AS$  имеет вид  $(\lambda, 0, 0, \dots, 0)$ , то первый столбец матрицы  $S$  является собственным вектором матрицы  $A$ , соответствующим собственному значению  $\lambda$ .

б) Докажите, что если первая строка матрицы  $S^{-1}AS$  имеет вид  $(\lambda, 0, 0, \dots, 0)$ , то первая строка матрицы  $S^{-1}$  является собственным вектором матрицы  $A^T$ , соответствующим собственному значению  $\lambda$ .

**1.9.** Пусть  $f(\lambda) = |A - \lambda I|$ , где  $A$  — матрица порядка  $n$ . Докажите, что  $f'(\lambda) = -\sum_{i=1}^n |A_i - \lambda I|$ , где  $A_i$  — матрица, полученная из матрицы  $A$  вычёркиванием  $i$ -й строки и  $i$ -го столбца.

**1.10.** Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения матрицы  $A$ . Докажите, что собственные значения матрицы  $\text{adj } A$  равны  $\prod_{i \neq 1} \lambda_i, \dots, \prod_{i \neq n} \lambda_i$ .

**1.11.** Матрица  $A = \|a_{ij}\|_i^n$  центрально симметрична, т.е.  $a_{ij} = a_{n-i, n-j}$ . Докажите, что среди собственных векторов матрицы  $A$ , соответствующих любому собственному значению, есть ненулевой симметричный ( $x_i = x_{n-i}$ ) или антисимметричный ( $x_i = -x_{n-i}$ ) вектор.

**1.12.** Докажите, что собственные значения матриц

а)  $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$ ; б)  $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$ ; в)  $\begin{pmatrix} A & B \\ B & -A \end{pmatrix}$  являются соответственно: а) соб-



ственными значениями матриц  $A \pm B$ ; б) собственными значениями матриц  $A \pm iB$ ; в) квадратными корнями из собственных значений матриц  $A^2 + B^2 \pm i(AB - BA)$ .

**1.13.** Элементы  $a_{i,n-i+1} = x_i$  комплексной матрицы  $A = \|a_{ij}\|_1^n$  могут быть ненулевыми, а все остальные её элементы равны нулю. Какому условию должен удовлетворять набор чисел  $(x_1, \dots, x_n)$ , чтобы матрица  $A$  была диагонализируемой?

**1.14.** [Drl] а) Докажите, что матрица  $A$  имеет  $m$  линейно независимых собственных векторов, соответствующих вещественным собственным значениям, тогда и только тогда, когда существует неотрицательно определённая матрица  $S$  ранга  $m$ , для которой  $AS = SA^*$ .

б) Докажите, что матрица  $A$  имеет  $m$  линейно независимых собственных векторов, соответствующих равным по модулю 1 собственным значениям, тогда и только тогда, когда существует такая неотрицательно определённая матрица  $S$  ранга  $m$ , что  $ASA^* = S$ .

**1.15.** Пусть  $A$  — матрица порядка  $n$  с собственными значениями  $x_1, \dots, x_n$ . Докажите, что для любого натурального  $m$  имеет место равенство

$$A^{n+m} = \sum_{k=0}^{n-1} (-1)^k s_{(m|k)}(x_1, \dots, x_n) A^{n-k-1},$$

где  $s_{(m|k)}$  — функция Шура, соответствующая разбиению  $(m|k)$ .

## 3.2. Жорданова нормальная форма

### 3.2.1. Подобные матрицы

Пусть  $A$  — матрица некоторого линейного оператора относительно базиса  $e$  пространства  $V$ ; тогда  $P^{-1}AP$  — матрица того же оператора относительно базиса  $Pe$ . Матрицы  $A$  и  $P^{-1}AP$  называют *подобными* или *сопряжёнными*. Выбрав подходящим образом базис, матрицу оператора можно привести к более простому виду: к нормальной жордановой форме, циклической форме, к матрице с равными элементами на диагонали или всеми нулевыми, кроме одного, элементами на диагонали и т.д.

Может показаться, что класс матриц, сопряжённых с матрицей  $A$  посредством комплексных матриц, шире класса матриц, сопряжённых с  $A$  посредством вещественных матриц. Но это не так.

**Теорема 3.2.1.1.** Пусть  $A$  и  $B$  — вещественные матрицы и  $A = PBP^{-1}$ , где  $P$  — комплексная матрица. Тогда  $A = QBQ^{-1}$ , где  $Q$  — вещественная матрица.

**Доказательство.** Требуется доказать, что если среди решений уравнения

$$AX = XB \quad (1)$$

есть невырожденная комплексная матрица  $P$ , то среди решений есть и невырожденная вещественная матрица  $Q$ . Решения над  $\mathbb{C}$  линейного уравнения (1) образуют линейное пространство  $W$  над  $\mathbb{C}$  с базисом  $C_1, \dots, C_n$ . Матрицу  $C_j$  можно представить в виде суммы  $C_j = X_j + iY_j$ , где  $X_j$  и  $Y_j$  — вещественные матрицы. Так как  $A$  и  $B$  — вещественные матрицы, то из равенства  $AC_j = C_jB$  следуют равенства  $AX_j = X_jB$  и  $AY_j = Y_jB$ . Поэтому пространство  $W$  порождено над  $\mathbb{C}$  вещественными матрицами  $X_1, \dots, X_n, Y_1, \dots, Y_n$ , а значит, в нём можно выбрать вещественный базис  $D_1, \dots, D_n$ .

Пусть  $P(t_1, \dots, t_n) = |t_1D_1 + \dots + t_nD_n|$ . По условию многочлен  $P(t_1, \dots, t_n)$  не равен тождественно нулю над  $\mathbb{C}$ , поэтому он не равен тождественно нулю и над  $\mathbb{R}$ , т.е. матричное уравнение (1) имеет невырожденное вещественное решение  $t_1D_1 + \dots + t_nD_n$ .  $\square$

### 3.2.2. Существование и единственность жордановой формы

*Жордановой клеткой* называют квадратную матрицу

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

порядка  $r$ .

*Жордановой матрицей* называют матрицу, состоящую из диагональных блоков  $J_{r_i}(\lambda_i)$  и нулей вне этих блоков.

*Жордановым базисом* для оператора  $A : V \rightarrow V$  называют базис пространства  $V$ , в котором матрица  $A$  является жордановой.

**Теорема 3.2.2.1.** *Для любого линейного оператора  $A : V \rightarrow V$  над алгебраически замкнутым полем существует жорданов базис, причём жорданова матрица оператора  $A$  определена однозначно с точностью до перестановки жордановых клеток.*

**Доказательство.** [Vä1] Докажем сначала существование жорданова базиса. Доказательство проведем индукцией по  $n = \dim V$ . При  $n = 1$  утверждение очевидно. Пусть  $\lambda$  — собственное значение оператора  $A$ . Рассмотрим вырожденный оператор  $B = A - \lambda I$ . Жорданов базис для оператора  $B$  будет также жордановым базисом для оператора  $A = B + \lambda I$ .

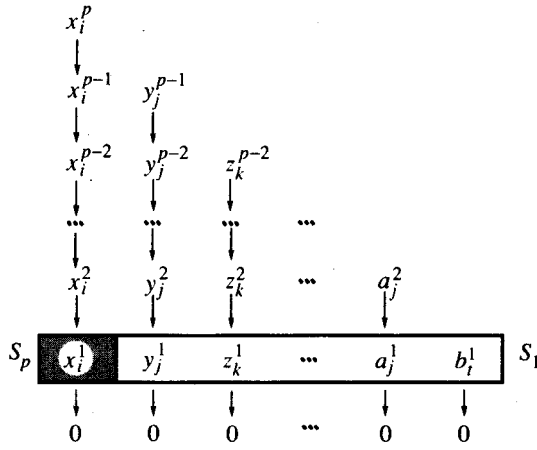


Рис. 3.1.

Последовательность  $\text{Im } B^0 \supset \text{Im } B^1 \supset \text{Im } B^2 \supset \dots$  стабилизируется, поэтому можно выбрать натуральное число  $p$  так, что  $\text{Im } B^{p+1} = \text{Im } B^p \neq \text{Im } B^{p-1}$ . Тогда  $\text{Im } B^p \cap \text{Ker } B = 0$  и  $\text{Im } B^{p-1} \cap \text{Ker } B \neq 0$ . Следовательно,  $B^p(\text{Im } B^p) = \text{Im } B^p$ .

Пусть  $S_i = \text{Im } B^{i-1} \cap \text{Ker } B$ . Тогда  $\text{Ker } B = S_1 \supset S_2 \supset \dots \supset S_p \neq 0$  и  $S_{p+1} = 0$ . Следить за дальнейшим ходом доказательства поможет рис. 3.1. Выберем в пространстве  $S_p$  базис  $x_i^1$  ( $i = 1, \dots, n_p$ ). Так как  $x_i^1 \in \text{Im } B^{p-1}$ , то  $x_i^1 = B^p x_i^p$  для некоторого вектора  $x_i^p$ . Рассмотрим векторы  $x_i^k = B^{p-k} x_i^p$  ( $k = 1, \dots, p$ ). Векторы  $x_i^1$  дополним до базиса пространства  $S_{p-1}$  векторами  $y_j^1$ , найдём вектор  $y_j^{p-1}$ , для которого  $y_j^1 = B^{p-2} y_j^{p-1}$ , и рассмотрим векторы  $y_j^l = B^{p-l-1} y_j^{p-1}$  ( $l = 1, \dots, p-1$ ). Затем дополним векторы  $x_i^1$  и  $y_j^1$  до базиса пространства  $S_{p-2}$  векторами  $z_k^1$  и т.д. Если  $i$  изменяется от 1 до  $I$ ,  $\dots$ ,  $t$  изменяется от 1 до  $T$ , то количество всех выбранных векторов равно  $pI + (p-1)J + \dots + 2S + T = I + (I+J) + \dots + (I+J+\dots+S+T) = \dim S_p + \dim S_{p-1} + \dots + \dim S_1$ . А так как  $\dim(\text{Im } B^{i-1} \cap \text{Ker } B) = \dim \text{Ker } B^i - \dim \text{Ker } B^{i-1}$  (см. §2.2.1), то  $\sum_{i=1}^p \dim S_i = \dim \text{Ker } B^p$ .

Дополним выбранные векторы базисом пространства  $\text{Im } B^p$  и докажем, что получится базис пространства  $V$ . Количество этих векторов показывает, что достаточно проверить их линейную независимость. Пред-

положим, что

$$f + \sum \alpha_i x_i^p + \sum \beta_i x_i^{p-1} + \cdots + \sum \gamma_j y_j^{p-1} + \cdots + \sum \delta_t b_t^1 = 0, \quad (1)$$

где  $f \in \text{Im } B^p$ . Применив к (1) оператор  $B^p$ , получим  $B^p(f) = 0$ , а значит,  $f = 0$ , так как  $B^p(\text{Im } B^p) = \text{Im } B^p$ . Применив теперь к (1) оператор  $B^{p-1}$ , получим  $\sum \alpha_i x_i^1 = 0$ , а значит, все числа  $\alpha_i$  нулевые. Применив теперь к (1) оператор  $B^{p-2}$ , получим  $\beta_i x_i^1 + \sum \gamma_j y_j^1 = 0$ , а значит, все числа  $\beta_i$  и  $\gamma_j$  нулевые, и т.д.

По предположению индукции в пространстве  $\text{Im } B^p \neq V$  можно выбрать жорданов базис для оператора  $B$ ; дополнив его выбранными векторами, получим жорданов базис пространства  $V$ .

Для доказательства однозначности жордановой формы достаточно проверить, что количество жордановых клеток матрицы  $B$ , соответствующих нулевому собственному значению, определено однозначно. Этим клеткам можно сопоставить диаграмму, изображенную на рис. 3.1, поэтому количество клеток порядка  $k$  равно

$$\begin{aligned} \dim S_k - \dim S_{k+1} &= (\dim \text{Ker } B^k - \dim \text{Ker } B^{k-1}) - \\ &\quad - (\dim \text{Ker } B^{k+1} - \dim \text{Ker } B^k) = \\ &= 2 \dim \text{Ker } B^k - \dim \text{Ker } B^{k-1} - \dim \text{Ker } B^{k+1} = \\ &= \text{rk } B^{k-1} - 2 \text{rk } B^k + \text{rk } B^{k+1}; \end{aligned}$$

это число определено инвариантно.  $\square$

### 3.2.3. Возведение матрицы в степень

Жорданову нормальную форму удобно использовать при возведении матрицы в степень. В самом деле, если  $A = P^{-1}JP$ , то  $A^s = P^{-1}J^sP$ . Для возведения в степень жордановой клетки  $\lambda I + N$  можно воспользоваться формулой бинома Ньютона  $(\lambda I + N)^s = \sum_{k=0}^s \binom{s}{k} \lambda^k N^{s-k}$  (эта формула верна, так как  $IN = NI$ ). Единственными ненулевыми элементами матрицы  $N^m$  являются единицы на местах  $(1, m+1), (2, m+2), \dots, (n-m, n)$ , где  $n$  — порядок матрицы  $N$ ; если  $m \geq n$ , то  $N^m = 0$ .

### 3.2.4. Жорданова форма над $\mathbb{R}$

Жорданов базис всегда существует лишь над алгебраически замкнутым полем; над полем  $\mathbb{R}$  жорданов базис существует не всегда. Но над полем  $\mathbb{R}$  тоже имеется жорданова форма, являющаяся овеществлением

жордановой формы над  $\mathbb{C}$ . Объясним, как она устроена. Заметим сначала, что жорданов базис, соответствующий вещественным собственным значениям оператора  $A$ , над  $\mathbb{R}$  строится точно так же, как над  $\mathbb{C}$ . Поэтому интересен лишь случай невещественных собственных значений.

Пусть  $A^{\mathbb{C}}$  — комплексификация вещественного оператора  $A$  (см. §2.7.1).

**Теорема 3.2.4.1.** *Между жордановыми клетками оператора  $A^{\mathbb{C}}$ , соответствующими собственным значениям  $\lambda$  и  $\bar{\lambda}$ , существует взаимно однозначное соответствие.*

**Доказательство.** Пусть  $B = P + iQ$ , где  $P$  и  $Q$  — вещественные операторы. Если  $x$  и  $y$  — вещественные векторы,  $z = x + iy$ ,  $\bar{z} = x - iy$ , то равенства  $Bz = 0$  и  $\bar{B}\bar{z} = 0$  эквивалентны. А так как  $(A - \bar{\lambda}I)^k = \overline{(A - \lambda I)^k}$ , то отображение  $z \mapsto \bar{z}$  задаёт взаимно однозначное соответствие между  $\text{Ker}(A - \lambda I)^k$  и  $\text{Ker}(A - \bar{\lambda}I)^k$ . Размерности этих пространств определяют количество и размеры жордановых клеток.  $\square$

Пусть  $J_n^*(\lambda)$  — матрица порядка  $2n$ , полученная из жордановой клетки  $J_n(\lambda)$  заменой каждого элемента  $x + iy$  матрицей  $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ .

**Теорема 3.2.4.2.** *Для оператора  $A$  существует базис, в котором его матрица имеет блочно диагональный вид с блоками  $J_{m_1}(t_1), \dots, J_{m_k}(t_k)$  для вещественных собственных значений  $t_i$  и блоками  $J_{n_1}^*(\lambda_1), \dots, J_{n_s}^*(\lambda_s)$  для невещественных собственных значений  $\lambda_i$  и  $\bar{\lambda}_i$ .*

**Доказательство.** Если  $\lambda$  — собственное значение оператора  $A$ , то согласно теореме 3.2.4.1  $\bar{\lambda}$  — тоже собственное значение оператора  $A$ , причём каждой жордановой клетке  $J_n(\lambda)$  оператора  $A$  соответствует жорданова клетка  $J_n(\bar{\lambda})$ . Кроме того, если  $e_1, \dots, e_n$  — жорданов базис для  $J_n(\lambda)$ , то  $\bar{e}_1, \dots, \bar{e}_n$  — жорданов базис для  $J_n(\bar{\lambda})$ . Поэтому вещественные векторы  $x_1, y_1, \dots, x_n, y_n$ , где  $e_k = x_k + iy_k$ , линейно независимы. Матрица ограничения оператора  $A$  на подпространство  $\langle x_1, y_1, \dots, x_n, y_n \rangle$  в базисе  $x_1, y_1, \dots, x_n, y_n$  имеет вид  $J_n^*(\lambda)$ .  $\square$

### 3.2.5. Аддитивная и мультипликативная запись жордановой формы

Жорданово разложение показывает, что любой линейный оператор  $A$  можно представить в виде суммы  $A = A_s + A_n$ , где  $A_s$  — диагонализируемый оператор,  $A_n$  — нильпотентный оператор, причём  $A_s A_n = A_n A_s$ .

**Теорема 3.2.5.1.** Операторы  $A_s$  и  $A_n$  определены однозначно, причём  $A_s = S(A)$  и  $A_n = N(A)$ , где  $S$  и  $N$  — некоторые многочлены.

**Доказательство.** Рассмотрим сначала одну жорданову клетку  $A = \lambda I + N_k$  порядка  $k$ . Пусть  $S(t) = \sum_{i=1}^m s_i t^i$ . Тогда  $S(A) = \sum_{i=1}^m s_i \sum_{j=0}^i \binom{i}{j} \lambda^j N_k^{i-j}$ , так как  $IN_k = N_k I$ . Коэффициент при  $N_k^p$  равен  $\sum_i s_i \binom{i-p}{i} \lambda^{i-p} = \frac{1}{p!} S^{(p)}(\lambda)$ , где  $S^{(p)}$  —  $p$ -я производная многочлена  $S$ . Поэтому нужно подобрать многочлен  $S$  так, что  $S(\lambda) = \lambda$  и  $S(\lambda) = \dots = S^{(k-1)}(\lambda) = 0$ , где  $k$  — порядок жордановой клетки. Если  $\lambda_1, \dots, \lambda_n$  — различные собственные значения оператора  $A$ ,  $k_1, \dots, k_n$  — порядки максимальных соответствующих им жордановых клеток, то многочлен  $S$  должен принимать в точке  $\lambda_i$  значение  $\lambda_i$  и иметь нулевые производные до порядка  $k_i - 1$  включительно. Такой многочлен всегда можно построить. Ясно также, что если  $A_s = S(A)$ , то  $A_n = A - S(A)$ , т.е.  $N(A) = A - S(A)$ .

Докажем теперь единственность разложения. Пусть  $A_s + A_n = A = A'_s + A'_n$ , причём  $A_s A_n = A_n A_s$  и  $A'_s A'_n = A'_n A'_s$ . Если  $A X = X A$ , то  $S(A) X = X S(A)$  и  $N(A) X = X N(A)$ . Поэтому  $A_s A'_s = A'_s A_s$  и  $A_n A'_n = A'_n A_n$ . Оператор  $B = A'_s - A_s = A'_n - A_n$  является разностью коммутирующих диагонализированных операторов, поэтому он диагонализирован (см. задачу 1.7 б). С другой стороны, оператор  $B$  является разностью коммутирующих нильпотентных операторов, поэтому он нильпотентен (см. задачу 1.7 а). Диагонализированный нильпотентный оператор равен нулю.  $\square$

**Следствие.** Если  $Ax = 0$ , то  $A_n x = 0$  и  $A_s x = 0$ .

**Доказательство.** Если  $Ax = 0$ , то  $A_n x = N(A)x = N(0)x = \lambda x$ . Но  $A_n^k = 0$  для некоторого натурального  $k$ , поэтому  $\lambda = 0$ , т.е.  $A_n x = 0$ . Следовательно,  $A_s x = (A - A_n)x = 0$ .  $\square$

Аддитивное жорданово разложение  $A = A_s + A_n$  позволяет получить для невырожденного оператора  $A$  разложение  $A = A_s A_u$ , где  $A_u$  — унитарный оператор, т.е. сумма единичного оператора и нильпотентного.

**Теорема 3.2.5.2.** Пусть  $A$  — невырожденный оператор над  $\mathbb{C}$ . Тогда  $A$  можно представить в виде  $A = A_s A_u = A_u A_s$  где  $A_s$  — полупростой оператор,  $A_u$  — унитарный оператор, причём такое разложение единственно.

**Доказательство.** Если оператор  $A$  невырожден, то оператор  $A_s$  тоже невырожден. Тогда  $A = A_s + A_n = A_s A_u$ , где  $A_u = A_s^{-1}(A_s + A_n) = I + A_s^{-1} A_n$ . Так как операторы  $A_s^{-1}$  и  $A_n$  коммутируют (см. §1.2.5), то

$A_s^{-1}A_n$  — нильпотентный оператор, коммутирующий с  $A_s$ . Докажем теперь единственность разложения. Если  $A = A_sA_u = A_uA_s$  и  $A_u = I + N$ , где  $N$  — нильпотентный оператор, то  $A = A_s(I + N) = A_s + A_sN$ , где  $A_sN$  — нильпотентный оператор, коммутирующий с  $A$ . Такой оператор  $A_sN = A_n$  единствен.  $\square$

### 3.2.6. Классификация Кронекера пар линейных операторов

Будем считать линейные операторы  $A_1, A_2 : V \rightarrow V$  эквивалентными, если один оператор получается из другого заменой базиса в пространстве  $V$ , т.е.  $PA_1 = A_2P$ , где  $P$  — невырожденный оператор. Теорема Жордана позволяет выбрать в каждом классе эквивалентности один оператор и тем самым решить задачу классификации линейных операторов.

Более сложная задача *классификации пар линейных отображений* была решена Кронекером. При этом пара линейных отображений  $A_1, B_1 : V \rightarrow W$  считается эквивалентной паре отображений  $A_2, B_2 : V \rightarrow W$ , если одна пара получается из другой заменой базисов в  $V$  и  $W$ , т.е.  $PA_1 = A_2Q$  и  $PB_1 = B_2Q$ , где  $P$  и  $Q$  — невырожденные линейные операторы. Чтобы решить задачу классификации, нужно привести к каноническому виду пару линейных отображений.

В том случае, когда хотя бы одно из отображений  $A$  и  $B$  обратимо, классификация Кронекера получается непосредственно из классификации Жордана. В самом деле, с помощью обратимого преобразования можно отождествить пространства  $V$  и  $W$  (т.е. в качестве базиса пространства  $W$  выбрать образ базиса пространства  $V$ ). Тогда матрица обратимого отображения будет единичной матрицей, а матрицу второго отображения можно будет привести к жордановой нормальной форме, подходящим образом выбрав базис пространства  $V$ .

Несмотря на столь обнадеживающее начало, классификация пар линейных отображений оказывается достаточно сложной в том случае, когда оба отображения вырожденные.

Назовём пару линейных отображений  $A, B : V \rightarrow W$  *разложимой*, если существуют разложения пространств  $V = V_1 \oplus V_2$  и  $W = W_1 \oplus W_2$ , для которых  $AV_i \subset W_i$  и  $BV_i \subset W_i$ , причём хотя бы одно из этих разложений нетривиально. В случае разложимой пары отображений мы фактически получаем две пары отображений  $A_1, B_1 : V_1 \rightarrow W_1$  и  $A_2, B_2 : V_2 \rightarrow W_2$ , причём каждую пару можно приводить к каноническому виду независимо от другой. В дальнейшем нас будут интересовать лишь неразложимые пары линейных отображений.

Пусть  $I_k$  — единичная матрица порядка  $k$ ,  $P_k = (I_k, 0)$  — единичная матрица, к которой справа приписан столбец нулей,  $Q_k = (0, I_k)$ .

Несложно показать, что пара линейных отображений  $P_k, Q_k : V^{k+1} \rightarrow W^k$  неразложима.

**Теорема 3.2.6.1 (Кронекер).** Любая неразложимая пара линейных отображений  $A, B : V \rightarrow W$  над полем  $\mathbb{C}$  эквивалентна одной из следующих пар:

- 1)  $(I_k, J_k(\lambda)), k \geq 1;$
- 2)  $(J_k(\lambda), I_k), k \geq 1;$
- 3)  $(P_k, Q_k), k \geq 0;$
- 4)  $(P_k^T, Q_k^T), k \geq 0.$

**Доказательство.** ([Bu]) Перейдя при необходимости к транспонированным матрицам (т.е. к отображениям двойственных пространств), можно считать, что  $\dim V \leq \dim W$ . Пусть  $\dim V = n, \dim W = n + l$ , где  $l \geq 0$ .

В том случае, когда  $\text{Ker } A \cap \text{Ker } B = U \neq 0$ , неразложимая пара отображений имеет вид  $A, B : U \rightarrow 0$ , причём  $\dim U = 1$ . А в том случае, когда  $\text{Im } A + \text{Im } B = U \neq W$ , неразложимая пара отображений имеет вид  $A, B : 0 \rightarrow U^\perp$ , где  $U^\perp$  — дополнение к  $U$  в  $W$ ; при этом  $\dim U = 1$ . Поэтому в дальнейшем будем считать, что  $\text{Ker } A \cap \text{Ker } B = 0$  и  $\text{Im } A + \text{Im } B = W$ .

Рассмотрим подпространства  $K = \text{Im } A \cap \text{Im } B$  и  $L = A^{-1}(K) \cap B^{-1}(K)$ . Если  $\dim \text{Ker } A = a$  и  $\dim \text{Ker } B = b$ , то  $\dim \text{Im } A = n - a$  и  $\dim \text{Im } B = n - b$ . Поэтому  $\dim K + (n + l) = (n - a) + (n - b)$ , т.е.  $\dim K = n - l - a - b$ . Ясно также, что

$$\dim A^{-1}(K) = \dim \text{Ker } A + \dim K = a + (n - l - a - b) = n - l - b$$

и  $\dim B^{-1}(K) = n - l - a$ . Поэтому

$$\dim L \geq \dim A^{-1}(K) + \dim B^{-1}(K) - \dim V = n - a - b - 2l.$$

Положим  $\dim L = n - a - b - 2l + r$ , где  $r \geq 0$ .

Рассмотрим подпространство  $(\text{Ker } A)_L = L \cap \text{Ker } A$  и возьмём его дополнение  $(\text{Ker } A)_N$  в пространстве  $\text{Ker } A$ . Аналогично определим подпространства  $(\text{Ker } B)_L$  и  $(\text{Ker } B)_N$ .

В пространстве  $A^{-1}(K)$  возьмём дополнение  $A^{-1}(K)_N$  к пространству  $L \oplus (\text{Ker } A)_N$ . Аналогично определим подпространство  $B^{-1}(K)_N$ . Ясно, что

$$\begin{aligned} \dim A^{-1}(K)_N + \dim(\text{Ker } A)_N &= \dim A^{-1}(K) - \dim L = \\ &= (n - l - b) - (n - a - b - 2l + r) = l + a - r. \end{aligned}$$

Аналогично  $\dim B^{-1}(K)_N + \dim(\text{Ker } B)_N = l + b - r$ .



Покажем, что сумма подпространств  $L \oplus (\text{Ker } A)_N \oplus A^{-1}(K)_N$  и  $(\text{Ker } B)_N \oplus B^{-1}(K)_N$  прямая. Предположим, что вектор  $x$  лежит в первом из этих подпространств. Тогда  $Ax \in A(L) + K \subset K$ . А если вектор  $x$  лежит во втором из этих подпространств, то  $Bx \in K$ . Таким образом,  $x \in A^{-1}(K) \cap B^{-1}(K) = L$ . Остаётся заметить, что  $(\text{Ker } B)_N \oplus B^{-1}(K)_N \cap L = 0$ .

Выберем в пространстве  $V$  подпространство  $X$  так, что

$$V = L \oplus (\text{Ker } A)_N \oplus A^{-1}(K)_N \oplus (\text{Ker } B)_N \oplus B^{-1}(K)_N \oplus X.$$

При этом  $\dim X = n - (n - a - b - 2l + r) - (l + a - r) - (l + b - r) = r$ . Рассмотрим подпространство  $Y = AX + BX \subset W$ . Ясно, что  $\dim Y \leq 2r$ . Учитывая, что  $AL \subset K$ ,  $A(\text{Ker } A)_N = 0$ ,  $A(A^{-1}(K)_N) \subset K$  и т.д., получаем, что пространство  $W = \text{Im } A + \text{Im } B$  можно представить в виде

$$K + B(\text{Ker } A)_N + B(A^{-1}(K)_N) + A(\text{Ker } B)_N + A(B^{-1}(K)_N) + Y.$$

Сумма размерностей слагаемых при этом не превосходит

$$(n - l - a - b) + (l + a - r) + (l + b - r) + 2r = n + l = \dim W.$$

Это означает, что рассматриваемая сумма подпространств прямая и, кроме того, ограничения отображения  $A$  на  $(\text{Ker } B)_N$  и  $B^{-1}(K)_N$  обратимы, ограничения отображение  $B$  на  $(\text{Ker } A)_N$  и  $A^{-1}(K)_N$  обратимы и  $\dim Y = 2r$ .

Отображения  $A$  и  $B$  переводят  $L \oplus A^{-1}(K)_N \oplus B^{-1}(K)_N$  в  $K \oplus B(A^{-1}(K)_N) \oplus A(B^{-1}(K)_N)$ ,  $(\text{Ker } A)_N$  в  $B(\text{Ker } A)_N$ ,  $(\text{Ker } B)_N$  в  $A(\text{Ker } B)_N$ ,  $X$  в  $Y$ . Неразложимая пара отображений получается лишь в том случае, когда пространства  $V$  и  $W$  сводятся к одному из этих четырёх слагаемых.

Легко проверить, что неразложимые отображения

$$(\text{Ker } A)_N \rightarrow B(\text{Ker } A)_N, \quad (\text{Ker } B)_N \rightarrow A(\text{Ker } B)_N, \quad X \rightarrow Y$$

имеют вид  $(0, I_1)$ ,  $(I_1, 0)$ ,  $(P_1^T, Q_1^T)$  соответственно.

Остаётся разобрать случай неразложимой пары отображений

$$A, B : V = L \oplus A^{-1}(K)_N \oplus B^{-1}(K)_N \rightarrow K \oplus B(A^{-1}(K)_N) \oplus A(B^{-1}(K)_N) = W.$$

**Лемма.** Пусть  $A_L, B_L : L \rightarrow K$  — ограничения отображений  $A$  и  $B$ . Тогда пара отображений  $A_L, B_L$  неразложима.

**Доказательство.** Предположим, что пара отображений  $A_L, B_L$  разложима, и пусть  $L = L' \oplus L''$  и  $K = K' \oplus K''$  — соответствующие разложения пространств  $L$  и  $K$ .

Для отображения  $A : U \rightarrow K' \oplus K''$  можно построить разложение  $U = U' \oplus U''$  так, что  $AU' \subset K'$  и  $AU'' \subset K''$ . В самом деле,  $A^{-1}(K') = \text{Ker } A \oplus U'_1$  и  $A^{-1}(K'') = \text{Ker } A \oplus U''_1$ , причём  $U = \text{Ker } A \oplus U'_1 \oplus U''_1$ . Положим  $U' = \text{Ker } A \oplus U'_1$  и  $U'' = U''_1$ . Тогда  $U = U' \oplus U''$  — требуемое разложение.

Построим таким способом разложение  $A^{-1}(K)_N = A^{-1}(K)'_N \oplus A^{-1}(K)''_N$  для отображения  $A$  и разложение  $B^{-1}(K)_N = B^{-1}(K)'_N \oplus B^{-1}(K)''_N$  для отображения  $B$ . Тогда образ пространства  $L' \oplus A^{-1}(K)'_N \oplus B^{-1}(K)'_N$  при отображении  $A$  содержится в  $K' \oplus A(B^{-1}(K)'_N)$ , а при отображении  $B$  — в  $K' \oplus B(A^{-1}(K)'_N)$ .

Рассмотрим разложения пространств  $V = V' \oplus V''$  и  $W = W' \oplus W''$ , где  $V' = L' \oplus A^{-1}(K)'_N \oplus B^{-1}(K)'_N$  и  $W' = K' \oplus B(A^{-1}(K)'_N) \oplus A(B^{-1}(K)'_N)$ ; пространства  $V''$  и  $W''$  определяются аналогично. Это разложение показывает, что пара отображений  $A, B$  разложима.  $\square$

Напомним, что в рассматриваемой ситуации  $r = \dim X = 0$ , поэтому  $\dim L = m - n - 2l$  и  $\dim K = m - n - l$ , где  $m = a + b$ . Разберём отдельно три случая: (1)  $l = 0$ ; (2)  $l = 1$ ; (3)  $l \geq 2$ .

(1)  $l = 0$ . Покажем, что в этом случае одно из отображений  $A$  и  $B$  взаимно однозначно, т.е. пара  $A, B$  эквивалентна паре  $I, J$  или  $J, I$ , где  $I$  — единичная матрица,  $J$  — жорданова матрица. Ясно, что для неразложимой пары отображений жорданова матрица  $J$  состоит из одного жорданова блока  $J_k(\lambda)$ .

Доказательство проведём индукцией по  $n$ . Напомним, что при  $l = 0$  имеют место равенства  $\dim V = \dim W = n$  и  $\dim L = \dim K = n - m$ .

База индукции:  $n = 1$ . Из того, что  $W = \text{Im } A + \text{Im } B \neq 0$ , следует, что одно из отображений  $A$  и  $B$  ненулевое. А для отображений одномерных пространств ненулевое линейное отображение взаимно однозначно.

Для доказательства шага индукции придётся воспользоваться индукцией по  $m$ . При  $m = 0$ , т.е. при  $\dim \text{Ker } A + \dim \text{Ker } B = 0$ , оба отображения  $A$  и  $B$  взаимно однозначны. В случае  $m > 0$  рассмотрим ограничения  $A_L$  и  $B_L$  отображений  $A$  и  $B$  на  $L$ . Согласно лемме полученная пара отображений  $A_L$  и  $B_L$  неразложима. Но тогда по предположению индукции одно из отображений  $A_L$  и  $B_L$  отображает  $L$  на  $K$  взаимно однозначно. Пусть, например, отображение  $A_L : L \rightarrow K$  взаимно однозначно. Тогда

$$L = A^{-1}(K) = L \oplus (\text{Ker } A)_N \oplus A^{-1}(K)_N,$$

поэтому  $A^{-1}(K)_N = 0$ . Следовательно, отображение

$$A : V = L \oplus B^{-1}(K)_N \rightarrow K \oplus A(B^{-1}(K)_N) = W$$

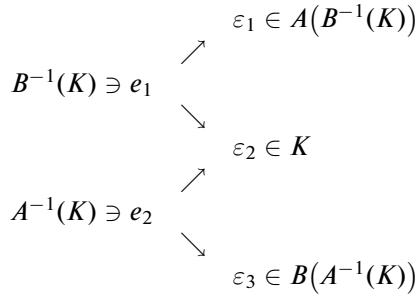
взаимно однозначное.

(2)  $l = 1$ . В этом случае пара  $A, B$  эквивалентна паре  $\begin{pmatrix} I_n \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ I_n \end{pmatrix}$ , т.е. паре  $P_n^T, Q_n^T$ .

Доказательство проведём индукцией по  $n$ . Предварительно напомним, что при  $l = 1$  имеем  $\dim V = n$ ,  $\dim W = n + 1$ ,  $\dim L = n - m - 2$ ,  $\dim K = n - m - 1$ .

При  $n = 1$  в качестве базиса пространства  $V$  выбираем произвольный ненулевой вектор  $e_1$ , а в качестве базиса в пространстве  $W$  выбираем векторы  $\varepsilon_1 = Ae_1$  и  $\varepsilon_2 = Be_1$ .

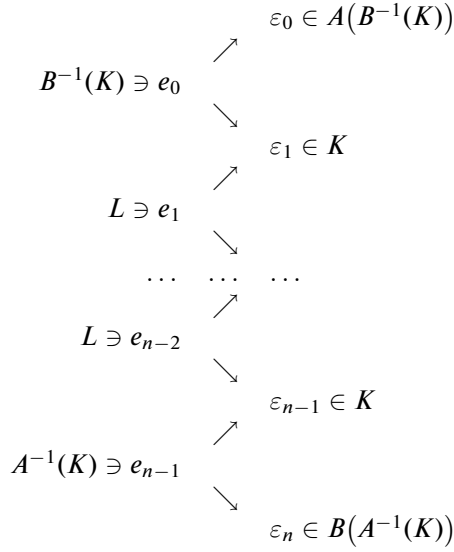
При  $n = 2$  имеем  $\dim L = n - m - 2 = -m$ , поэтому  $m = 0$ , т.е.  $\dim L = 0$  и  $\dim K = 1$ . Выберем в  $K$  ненулевой вектор  $\varepsilon_2$  и рассмотрим векторы  $e_1$  и  $e_2$ , для которых  $\varepsilon_2 = Be_1$  и  $\varepsilon_2 = Ae_2$ . Положим, далее,  $\varepsilon_1 = Ae_1$  и  $\varepsilon_3 = Be_2$ :



(стрелки, направленные вверх, соответствуют отображению  $A$ , а стрелки, направленные вниз, соответствуют отображению  $B$ ).

При  $n > 2$  рассмотрим ограничения  $A_L, B_L : L \rightarrow K$  отображений  $A$  и  $B$ . Согласно лемме пара отображений  $A_L, B_L$  неразложима. При этом  $\dim L = n - m - 2 = t$  и  $\dim K = t + 1$ , где  $t < n$ . По предположению индукции пара  $A_L, B_L$  эквивалентна паре  $\begin{pmatrix} I_t \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ I_t \end{pmatrix}$ . В частности,  $\text{Ker } A_L = \text{Ker } B_L = 0$ , а значит,  $t = n - 2$ . По предположению индукции можно построить базисы  $e_1, \dots, e_{n-2}$  и  $\varepsilon_1, \dots, \varepsilon_{n-1}$  в пространствах  $L$  и  $K$ , а с их помощью можно построить требуемые базисы  $e_0, e_1, \dots, e_{n-2}$ ,

$e_{n-1}$  и  $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{n-1}, \varepsilon_n$ :



(3)  $l \geq 2$ . В этом случае пара отображений  $A, B$  разложима. Доказательство проведём индукцией по  $n$ .

При  $n = 1$  выделяются прямые слагаемые  $V \rightarrow \text{Im } A + \text{Im } B$  и  $0 \rightarrow (\text{Im } A + \text{Im } B)^\perp \neq 0$ .

При  $n > 1$  рассмотрим ограничения  $A_L, B_L : L \rightarrow K$  отображений  $A$  и  $B$ . Так как  $\dim L = n - m - 2l = t < n$  и  $\dim K = t + l$ , то по предположению индукции пара  $A_L, B_L$  разложима. Но тогда согласно лемме разложима и пара  $A, B$ .  $\square$

### Задачи

**2.1.** Дан линейный оператор  $A : V \rightarrow V$ . Докажите, что пространство  $V$  можно разложить в прямую сумму таких подпространств  $N$  и  $S$ , что  $A|_N$  — нильпотентный оператор,  $A|_S$  — невырожденный оператор, причём это разложение единственно (*разложение Фиттинга*).

**2.2.** Докажите, что матрицы  $A$  и  $A^T$  подобны.

**2.3.** Пусть  $\sigma$  — произвольная перестановка символов  $1, \dots, n$ ;  $P = \|p_{ij}\|_1^n$ , где  $p_{ij} = \delta_{i, \sigma(j)}$ . Докажите, что матрица  $P^{-1}AP$  получается из матрицы  $A$  перестановкой строк, соответствующей  $\sigma$ , и такой же перестановкой столбцов.

**Замечание.** Матрицу  $P$  называют *матрицей перестановки*.

**2.4.** Докажите, что если  $A^2 = A$  и  $\text{tr } A = 0$ , то  $A = 0$ .

**2.5.** Количество попарно различных собственных значений матрицы  $A$  равно  $m$ , причём  $m > 1$ . Пусть  $b_{ij} = \text{tr}(A^{i+j})$ . Докажите, что  $|b_{ij}|_0^{m-1} \neq 0$  и  $|b_{ij}|_0^m = 0$ .

**2.6.** Докажите, что  $\text{rk } A = \text{rk } A^2$  тогда и только тогда, когда существует  $\lim_{\lambda \rightarrow 0} (A + \lambda I)^{-1} A$ .

### 3.3. Минимальный многочлен и характеристический многочлен

#### 3.3.1. Минимальный многочлен

Пусть  $p(t) = \sum_{k=0}^n a_k t^k$  — многочлен  $n$ -й степени. Для любой квадратной матрицы  $A$  можно рассмотреть матрицу  $p(A) = \sum_{k=0}^n a_k A^k$ . Многочлен  $p(t)$  называют *аннулирующим многочленом матрицы  $A$* , если  $p(A) = 0$  — нулевая матрица. Если  $A$  — матрица порядка  $n$ , то матрицы  $I, A, \dots, A^{n^2}$  линейно зависимы, так как размерность пространства матриц порядка  $n$  равна  $n^2$ . Поэтому для любой матрицы порядка  $n$  существует аннулирующий многочлен, причём его степень не превосходит  $n^2$ . Аннулирующий многочлен матрицы  $A$  минимальной степени с коэффициентом 1 при старшей степени называют *минимальным многочленом матрицы  $A$* . Докажем, что минимальный многочлен единствен. В самом деле, если  $p_1(A) = A^m + \dots = 0 = p_2(A) = A^m + \dots$ , то многочлен  $p_1 - p_2$  является аннулирующим для матрицы  $A$ , причём его степень меньше  $m$ . Следовательно,  $p_1 - p_2 = 0$ .

Легко проверить, что если  $B = X^{-1}AX$ , то  $B^n = X^{-1}A^nX$ , поэтому  $p(B) = X^{-1}p(A)X$ , а значит, минимальный многочлен однозначно определён не только для матрицы, но и для оператора.

**Теорема 3.3.1.1.** *Любой аннулирующий многочлен матрицы  $A$  делится на минимальный многочлен.*

**Доказательство.** Пусть  $p$  — минимальный многочлен матрицы  $A$ ,  $q$  — аннулирующий многочлен. Поделив  $q$  на  $p$  с остатком, получим  $q = pf + r$ , где  $\deg r < \deg p$ , причём  $r(A) = q(A) - p(A)f(A) = 0$ , т.е.  $r$  — аннулирующий многочлен. Следовательно,  $r = 0$ .  $\square$

Аннулирующим многочленом вектора  $v \in V$  (относительно оператора  $A: V \rightarrow V$ ) называют такой многочлен  $p$ , что  $p(A)v = 0$ . Аннулирующий многочлен минимальной степени с коэффициентом 1 при старшей степени называют *минимальным многочленом вектора  $v$* . Аналогично доказательству теоремы 3.3.1.1 можно доказать, что минимальный многочлен вектора является делителем минимального многочлена оператора  $A$ .

**Теорема 3.3.1.2.** *Существует вектор, минимальный многочлен которого совпадает с минимальным многочленом оператора  $A$ .*

**Доказательство.** Любой идеал  $I$  в кольце многочленов от одной переменной порождён многочленом  $f$  минимальной степени. В самом деле, если  $g \in I$ , то  $g = fh + r$ , а значит,  $r \in I$ , так как  $fh \in I$ .

Для любого вектора  $v \in V$  рассмотрим идеал  $I_v = \{p \mid p(A)v = 0\}$ ; этот идеал порождён некоторым многочленом  $p$  со старшим коэффициентом 1. Если  $p_A$  — минимальный многочлен оператора  $A$ , то  $p_A \in I_v$ , поэтому  $p_v$  — делитель  $p_A$ . Следовательно, когда  $v$  пробегает всё пространство  $V$ , получается лишь конечный набор многочленов  $p_v$ ; пусть это будут многочлены  $p_1, \dots, p_k$ . Пространство  $V$  содержится в объединении конечного набора своих подпространств  $V_i = \{x \in V \mid p_i(A)x = 0\}$  ( $i = 1, \dots, k$ ), поэтому  $V = V_i$  для некоторого  $i$ . Тогда  $p_i(A)V = 0$ , т.е.  $p_i$  делится на  $p_A$ , а значит,  $p_i = p_A$ .  $\square$

### 3.3.2. Теорема Гамильтона—Кэли

Простейшие соображения показывают, что степень минимального многочлена матрицы  $A$  порядка  $n$  не превосходит  $n^2$ . Оказывается, степень минимального многочлена не превосходит  $n$ , потому что характеристический многочлен матрицы  $A$  является аннулирующим многочленом.

**Теорема 3.3.2.1 (Гамильтон—Кэли).** *Пусть  $p(t) = |tI - A|$ . Тогда  $p(A) = 0$ .*

**Доказательство.** Для жордановой формы оператора доказательство очевидно, так как  $(\lambda - t)^n$  является аннулирующим многочленом жордановой клетки  $J_n(\lambda)$ . Приведём, однако, доказательство, не использующее теоремы Жордана. Можно считать, что  $A$  — матрица оператора над  $\mathbb{C}$ . Доказательство проведём индукцией по порядку  $n$  матрицы  $A$ . При  $n = 1$  утверждение очевидно. Пусть  $\lambda$  — собственное значение оператора  $A$ ,  $e_1$  — соответствующий ему собственный вектор. Дополним  $e_1$  до базиса  $e_1, \dots, e_n$ . В базисе  $e_1, \dots, e_n$  матрица оператора  $A$  имеет вид

$\begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix}$ , где  $A_1$  — матрица оператора  $A$  в факторпространстве  $V/\langle e_1 \rangle$ . Поэтому  $p(t) = (t - \lambda)|tI - A_1| = (t - \lambda)p_1(t)$ . По предположению индукции  $p_1(A_1) = 0$  в факторпространстве  $V/\langle e_1 \rangle$ , т.е.  $p_1(A_1)V \subset \langle e_1 \rangle$ . Остается заметить, что  $(A - \lambda I)e_1 = 0$ .  $\square$

**Замечание.** Используя жорданову нормальную форму, легко проверить, что минимальный многочлен матрицы  $A$  равен  $\prod_i (t - \lambda_i)^{n_i}$ , где произведение берётся по всем собственным значениям  $\lambda_i$  матрицы  $A$  и  $n_i$  — размер максимальной жордановой клетки, соответствующей  $\lambda_i$ . В частности, матрица  $A$  диагонализируема тогда и только тогда, когда её минимальный многочлен не имеет кратных корней и все его корни принадлежат основному полю.

### 3.3.3. Критерий совпадения характеристического многочлена и минимального многочлена

Согласно теореме Гамильтона—Кэли характеристический многочлен матрицы  $A$  порядка  $n$  совпадает с минимальным многочленом тогда и только тогда, когда степень минимального многочлена равна  $n$ . Минимальный многочлен матрицы  $A$  является минимальным многочленом для некоторого вектора  $v$  (см. теорему 3.3.1.2). Поэтому характеристический многочлен совпадает с минимальным тогда и только тогда, когда для некоторого вектора  $v$  векторы  $v, Av, \dots, A^{n-1}v$  линейно независимы.

**Теорема 3.3.3.1.** *Характеристический многочлен матрицы  $A$  порядка  $n$  совпадает с минимальным многочленом тогда и только тогда, когда для любого вектора  $(x_0, \dots, x_{n-1})$  найдутся столбцы  $P$  и  $Q$  длиной  $n$ , для которых  $x_k = Q^T A^k P$ .*

**Доказательство.** Предположим сначала, что степень минимального многочлена матрицы  $A$  равна  $n$ . Тогда найдётся такой столбец  $P$ , что столбцы  $P, AP, \dots, A^{n-1}P$  линейно независимы, т.е. матрица  $K$ , составленная из этих столбцов, невырождена. Любой вектор  $X = (x_1, \dots, x_n)$  можно представить в виде  $X = (XK^{-1})K = (Q^T P, \dots, Q^T A^{n-1}P)$ , где  $Q^T = XK^{-1}$ .

Предположим теперь, что для любого вектора  $(x_1, \dots, x_n)$  найдутся такие столбцы  $P$  и  $Q$ , что  $x_k = Q^T A^k P$ . Тогда найдутся такие столбцы  $P_1, \dots, P_n, Q_1, \dots, Q_n$ , что матрица

$$B = \begin{pmatrix} Q_1^T P_1 & \dots & Q_1^T A^{n-1} P_1 \\ \dots & \dots & \dots \\ Q_n^T P_n & \dots & Q_n^T A^{n-1} P_n \end{pmatrix}$$

невырождена. Матрицы  $I, A, \dots, A^{n-1}$  линейно независимы, так как иначе столбцы матрицы  $B$  были бы линейно зависимы.  $\square$

### 3.3.4. Обобщения теоремы Гамильтона–Кэли

Теорема Гамильтона–Кэли допускает различные обобщения.

**Теорема 3.3.4.1.** Пусть  $p_A(\lambda)$  — характеристический многочлен матрицы  $A$ , а матрица  $X$  коммутирует с  $A$ . Тогда  $p_A(X) = M(A - X)$ , где  $M$  — некоторая матрица, коммутирующая с  $A$  и  $X$ .

**Доказательство.** [Gr3] Поскольку, согласно §1.2.5,  $B \operatorname{adj} B = |B| \cdot I$ , получаем  $p_A(\lambda)I = [\operatorname{adj}(A - \lambda I)](A - \lambda I) = \left( \sum_{k=0}^{n-1} A_k \lambda^k \right) (A - \lambda I) = \sum_{k=0}^n \lambda^k A'_k$ . Все матрицы  $A'_k$  пропорциональны единичной, потому что матрица  $p(\lambda)I$  пропорциональна единичной. Следовательно,  $p_A(X) = \sum_{k=0}^n X^k A'_k$ . Если матрица  $X$  коммутирует с  $A$  и  $A_k$ , то  $p_A(X) = \left( \sum_{k=0}^{n-1} A_k X^k \right) (A - X)$ . Но матрицы  $A_k$  полиномиально выражаются через  $A$  (см. задачу 2.13), поэтому если  $X$  коммутирует с  $A$ , то  $X$  коммутирует с  $A_k$ .  $\square$

Аналог теоремы Гамильтона–Кэли можно доказать для квадратной блочной матрицы

$$A = \begin{pmatrix} A_{11} & \dots & A_{1m} \\ \dots & \dots & \dots \\ A_{m1} & \dots & A_{mm} \end{pmatrix},$$

где  $A_{ij}$  — матрица размера  $n_i \times n_j$  (в обычную теорему Гамильтона–Кэли эта теорема превращается при  $m = 1$ ). Предварительно введём некоторые обозначения.

Пусть  $\delta(i) = (0, \dots, 0, 1, 0, \dots, 0)$  — набор из  $m$  цифр с единицей на  $i$ -м месте;

$$A^{\delta(i)} = \begin{pmatrix} 0 & & \\ A_{i1} & \dots & A_{im} \\ 0 & & \end{pmatrix}.$$

Тогда  $A = \sum_{i=1}^m A^{\delta(i)}$ . Для мультииндекса  $z = (z_1, \dots, z_m)$  индукцией по  $|z| = z_1 + \dots + z_m$  определим  $A^z = \sum_{i=1}^m A^{\delta(i)} A^{z - \delta(i)}$ ; предполагается, что  $A^z = 0$ , если  $z_i < 0$  для некоторого  $i$ , и  $A^z = I$  для  $z = (0, 0, \dots, 0)$ .



Пусть, далее,

$$I^{\delta(i)} = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & I_{ii} & 0 \\ 0 & & & \ddots \\ & & & & 0 \end{pmatrix},$$

где  $I_{ii}$  — единичная матрица порядка  $n_i$ . Тогда  $I = \sum_{i=1}^m I^{\delta(i)}$ . Рассмотрим матрицу  $B(x_1, \dots, x_m) = A - \sum_{i=1}^m I^{\delta(i)}$ . Характеристическим многочленом блочной матрицы  $A$  называют многочлен

$$f(x_1, \dots, x_m) = |B(x_1, \dots, x_m)| = \sum b_{z_1 \dots z_m} x_1^{z_1} \dots x_m^{z_m} = \sum_z b_z x^z$$

(суммирование ведётся по значениям мультииндекса  $z = (z_1, \dots, z_m)$ , где  $0 \leq z_i \leq n_i$ ).

**Теорема 3.3.4.2.** Если в характеристическом многочлене  $f(x_1, \dots, x_m)$  заменить  $x^z$  на  $A^z$ , то получим нулевую матрицу.

**Доказательство.** [Wa] Пусть  $\text{adj } B(x_1, \dots, x_m) = \sum_z B_z x^z$ , где элементы матриц  $B_z$  не зависят от  $x$ . Из равенства  $f(x_1, \dots, x_m) \cdot I = || \cdot I = (\text{adj } B)B$  получаем  $\sum_z b_z x^z I = \left( \sum_z B_z x^z \right) \left( A - \sum_{i=1}^m x_i I^{\delta(i)} \right)$ . Так как  $\sum_{i=1}^m I^{\delta(i)} = I$ , то  $x^z A = \sum_{i=1}^m I^{\delta(i)} x^z A$ , а значит,  $\sum_z b_z x^z I = \sum_z B_z \sum_{i=1}^m I^{\delta(i)} (A x^z - x^{z+\delta(i)})$ . Поэтому  $b_z I = \sum_{i=1}^m (B_z I^{\delta(i)} A - B_{z-\delta(i)} I^{\delta(i)})$ . Следовательно, если  $x^z$  заменить на  $A^z$ , то получим матрицу  $\sum_z B_z \sum_{i=1}^m I^{\delta(i)} (A A^z - A^{z+\delta(i)})$ . Пусть  $J_z = \sum_{i=1}^m I^{\delta(i)} (A A^z - A^{z+\delta(i)})$ . Достаточно доказать, что  $J_z = 0$ . По определению  $A^{z+\delta(i)} = \sum_{j=1}^m A^{\delta(j)} A^{z+\delta(i)-\delta(j)}$ , а так как  $I^{\delta(i)} A^{\delta(j)} = \delta_{ij} A^{\delta(i)} = \delta_{ij} I^{\delta(i)} A$ , то получаем  $\sum_{i=1}^m I^{\delta(i)} A^{z+\delta(i)} = \sum_{i,j=1}^m \delta_{ij} I^{\delta(i)} A A^{z+\delta(i)-\delta(j)} = \sum_{i=1}^m I^{\delta(i)} A A^z$ .  $\square$

### Задачи

**3.1.** Пусть  $A$  — матрица порядка  $n$ ;  $f_1(A) = A - (\operatorname{tr} A)I$ ,  $f_{k+1}(A) = f_k(A)A - \frac{1}{k+1} \operatorname{tr}(f_k(A)A)I$ . Докажите, что  $f_n(A) = 0$ .

**3.2.** Пусть  $A$  и  $B$  — матрицы порядка  $n$ . Докажите, что если  $\operatorname{tr} A^m = \operatorname{tr} B^m$  при  $m = 1, \dots, n$ , то собственные значения матриц  $A$  и  $B$  совпадают.

**3.3.** Пусть  $p_A$  — минимальный многочлен матрицы  $A$ . Докажите, что если матрицы  $A$  и  $B$  не имеют общих собственных значений, то матрица  $p_A(B)$  невырожденная.

**3.4.** Матрица  $A$  невырождена и её минимальный многочлен  $p(\lambda)$  совпадает с характеристическим. Докажите, что минимальный многочлен матрицы  $A^{-1}$  равен  $p(0)^{-1} \lambda^n p(\lambda^{-1})$ .

**3.5.** Пусть  $p_A$  — минимальный многочлен матрицы  $A$ ,  $M_n$  — пространство матриц порядка  $n$ . Докажите, что  $B_k = \{A \in M_n \mid \deg p_A < k\}$  — замкнутое подмножество  $M_n$ .

## 3.4. Каноническая форма Фробениуса

### 3.4.1. Циклические клетки

Жорданова форма является лишь одной из канонических форм матриц линейных операторов. Примером другой канонической формы является *циклическое разложение* (каноническая форма Фробениуса). Циклической клеткой называют матрицу вида

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}.$$

Если  $A: V^n \rightarrow V^n$  и  $Ae_1 = e_2, \dots, Ae_{n-1} = e_n$ , то матрица оператора  $A$  относительно базиса  $e_1, \dots, e_n$  является циклической клеткой.

**Теорема 3.4.1.1.** Для любого линейного оператора  $A: V \rightarrow V$  (над полем  $\mathbb{C}$  или  $\mathbb{R}$ ) существует базис, в котором матрица  $A$  имеет блочно диагональный вид, причём блоками являются циклические клетки.

**Доказательство.** [J1] Применим индукцию по размерности пространства  $V$ . Если степень минимального многочлена  $A$  равна  $k$ , то существует вектор  $y \in V$ , степень минимального многочлена которого тоже равна  $k$  (теорема 3.3.1.2). Пусть  $y_i = A^{i-1}y$ . Дополним базис  $y_1, \dots, y_k$  пространства  $W = \langle y_1, \dots, y_k \rangle$  до базиса пространства  $V$  и рассмотрим  $W_1^* = \langle y_k^*, A^*y_k^*, \dots, A^{*k-1}y_k^* \rangle$ . Докажем, что  $V = W \oplus W_1^{*\perp} = A$ -инвариантное разложение пространства  $V$ .

Степень минимального многочлена оператора  $A^*$  тоже равна  $k$ , поэтому подпространство  $W_1^*$  инвариантно относительно  $A^*$ , а значит, подпространство  $(W_1^*)^\perp$  инвариантно относительно  $A$  (см. теорему 2.1.8.1). Остается доказать, что  $W_1^* \cap W^\perp = 0$  и  $\dim W_1^* = k$ . Предположим, что  $a_0y_k^* + \dots + a_sA^{*s}y_k^* \in W^\perp$  при  $0 \leq s \leq k-1$  и  $a_s \neq 0$ . Тогда  $A^{*k-s-1}(a_0y_k^* + \dots + a_sA^{*s}y_k^*) \in W^\perp$ , значит,  $0 = \langle a_0A^{*k-s-1}y_k^* + \dots + a_sA^{*k-1}y_k^*, y \rangle = a_0\langle y_k^*, A^{k-s-1}y \rangle + \dots + a_s\langle y_k^*, A^{k-1}y \rangle = a_0\langle y_k^*, y_{k-s} \rangle + \dots + a_s\langle y_k^*, y_k \rangle = a_s$ . Получено противоречие.

Матрица ограничения оператора  $A$  на подпространство  $W$  в базисе  $y_1, \dots, y_k$  является циклической клеткой. Ограничение оператора  $A$  на подпространство  $W_1^{*\perp}$  можно представить в требуемом виде по предположению индукции.  $\square$

**Замечание.** При доказательстве для оператора  $A$  найден базис, в котором его матрица имеет блочно диагональный вид. На диагонали стоят циклические клетки с характеристическими многочленами  $p_1, p_2, \dots, p_k$ , где  $p_1$  — минимальный многочлен  $A$ ,  $p_2$  — минимальный многочлен ограничения оператора  $A$  на некоторое подпространство, поэтому  $p_2$  — делитель  $p_1$ . Аналогично  $p_{i+1}$  — делитель  $p_i$ .

### 3.4.2. Характеристический многочлен циклической клетки

Докажем, что характеристический многочлен циклической клетки

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

равен  $\lambda^n + \sum_{k=0}^{n-1} a_k \lambda^k$ . Так как  $Ae_1 = e_2, \dots, Ae_{n-1} = e_n$  и  $Ae_n = -\sum_{k=0}^{n-1} a_k e_{k+1}$ ,

то  $(A^n + \sum_{k=0}^{n-1} a_k A^k)e_1 = 0$ . Учитывая, что  $e_i = A^{i-1}e_1$ , получаем, что  $\lambda^n +$

$\sum_{k=0}^{n-1} a_k \lambda^k$  — аннулирующий многочлен оператора  $A$ . Остается заметить,

что векторы  $e_1, Ae_1, \dots, A^{n-1}e_1$  линейно независимы, поэтому степень минимального многочлена  $A$  не меньше  $n$ .

Попутно мы доказали, что характеристический многочлен циклической клетки совпадает с минимальным многочленом.

### Задачи

**4.1.** Матрица оператора  $A$  является блочно диагональной и состоит из двух циклических клеток с взаимно простыми характеристическими многочленами  $p$  и  $q$ . Докажите, что базис можно выбрать так, чтобы матрица оператора стала циклической клеткой.

**4.2.** Пусть  $A$  — жорданова клетка, т.е. существует такой базис  $e_1, \dots, e_n$ , что  $Ae_1 = \lambda e_1$ ,  $Ae_k = e_{k-1} + \lambda e_k$  при  $k = 2, \dots, n$ . Докажите, что существует такой вектор  $v$ , что векторы  $v, Av, \dots, A^{n-1}v$  образуют базис (тогда матрица оператора  $A$  относительно базиса  $v, Av, \dots, A^{n-1}v$  является циклической клеткой).

**4.3.** Для циклической клетки  $A$  укажите такую симметрическую матрицу  $S$ , что  $A = SA^T S^{-1}$ .

## 3.5. Приведение диагонали к удобному виду

### 3.5.1. Все диагональные элементы, кроме одного, равны нулю

Замена матрицы  $A$  на  $XAX^{-1}$  сохраняет след, поэтому диагональные элементы матрицы  $XAX^{-1}$  не могут быть сделаны совсем произвольными. Можно, однако, диагональ матрицы  $A$  привести к более удобному виду, например, матрица  $A \neq \lambda I$  подобна матрице, у которой все диагональные элементы, кроме одного, равны нулю, и любая матрица подобна матрице, у которой все диагональные элементы равны.

**Теорема 3.5.1.1.** Пусть  $A \neq \lambda I$ . Тогда  $A$  подобна матрице, все диагональные элементы которой, кроме одного, равны нулю.

**Доказательство.** [Gi] Диагональ циклической клетки имеет требуемый вид. Поэтому утверждение верно для любой матрицы, характеристический многочлен которой совпадает с минимальным (см. §3.4.1).

Для матрицы порядка 2 характеристический многочлен не совпадает с минимальным только для матриц вида  $\lambda I$ . Пусть теперь  $A$  — матрица порядка 3, причём  $A \neq \lambda I$  и характеристический многочлен матрицы  $A$

не совпадает с минимальным многочленом. Тогда минимальный многочлен матрицы  $A$  имеет вид  $(x-\lambda)(x-\mu)$ , а характеристический многочлен равен  $(x-\lambda)^2(x-\mu)$ , причём случай  $\lambda = \mu$  не исключается. Следовательно, матрица  $A$  подобна матрице  $C = \begin{pmatrix} 0 & 1 & 0 \\ a & b & 0 \\ 0 & 0 & \lambda \end{pmatrix}$ , причём характеристический многочлен матрицы  $\begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}$  делится на  $x-\lambda$ , т.е.  $\lambda^2 - b\lambda - a = 0$ . Если  $b = \lambda = 0$ , то теорема верна. Пусть  $b = \lambda \neq 0$ . Так как  $b^2 - b^2 - a = 0$ , то  $a = 0$ . В этом случае

$$\begin{pmatrix} b & -1 & b \\ b & 0 & 0 \\ b & 0 & b \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{pmatrix} = \begin{pmatrix} 0 & 0 & b^2 \\ 0 & b & 0 \\ 0 & b & b^2 \end{pmatrix} = \\ = \begin{pmatrix} 0 & -b & b \\ -b & 0 & b \\ -b & -b & 2b \end{pmatrix} \begin{pmatrix} b & -1 & b \\ b & 0 & 0 \\ b & 0 & b \end{pmatrix},$$

причём  $\det \begin{pmatrix} b & -1 & b \\ b & 0 & 0 \\ b & 0 & b \end{pmatrix} \neq 0$ , поэтому матрица  $A$  подобна матрице  $\begin{pmatrix} 0 & -b & b \\ -b & 0 & b \\ -b & -b & 2b \end{pmatrix}$ .

Пусть, наконец,  $b \neq \lambda$ . Тогда для матрицы  $D = \text{diag}(b, \lambda)$  теорема верна, поэтому существует такая матрица  $P$ , что  $PD P^{-1} = \begin{pmatrix} 0 & * \\ * & * \end{pmatrix}$ . Матрица

$$\begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} C \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} 0 & * \\ * & D \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} = \begin{pmatrix} 0 & * \\ * & PD P^{-1} \end{pmatrix}$$

имеет требуемый вид.

Предположим теперь, что теорема верна для матриц порядка  $m$ , где  $m \geq 3$ . Матрица  $A$  порядка  $m+1$  имеет вид  $\begin{pmatrix} A_1 & * \\ * & * \end{pmatrix}$ , где  $A_1$  — матрица порядка  $m$ . Так как  $A \neq \lambda I$ , можно считать, что  $A_1 \neq \lambda I$  (для доказательства можно воспользоваться результатом задачи 2.3). По предположению индукции существует такая матрица  $P$ , что диагональ матрицы  $PA_1 P^{-1}$  имеет вид  $(0, 0, \dots, 0, \alpha)$ . Поэтому диагональ матрицы

$$X = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_1 & * \\ * & * \end{pmatrix} \begin{pmatrix} P^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} PA_1 P^{-1} & * \\ * & * \end{pmatrix}$$

имеет вид  $(0, \dots, 0, \alpha, \beta)$ . Если  $\alpha = 0$ , то доказательство завершено. Пусть  $\alpha \neq 0$ . Тогда  $X = \begin{pmatrix} 0 & * \\ * & C_1 \end{pmatrix}$ , где диагональ матрицы  $C_1$  порядка  $m$  имеет вид  $(0, 0, \dots, \alpha, \beta)$ , поэтому  $C_1 \neq \lambda I$ . Следовательно, существует такая матрица  $Q$ , что диагональ матрицы  $QCQ^{-1}$  имеет вид  $(0, \dots, 0, x)$ . Поэтому диагональ матрицы  $\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} 0 & * \\ * & C_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix}$  имеет требуемый вид.

□

**Замечание.** Доказательство остаётся справедливым для поля любой характеристики.

### 3.5.2. Все диагональные элементы равны

**Теорема 3.5.2.1.** Пусть  $A$  — произвольная комплексная (вещественная) матрица. Тогда существует такая унитарная (ортогональная) матрица  $U$ , что матрица  $UAU^{-1}$  имеет на диагонали равные числа.

**Доказательство.** Если унитарная матрица  $P$  соответствует перестановке  $\sigma$ , то диагональные элементы матрицы  $P^{-1}AP$  получаются из диагональных элементов матрицы  $A$  перестановкой  $\sigma$  (см. задачу 2.3). Поэтому любые два диагональных элемента матрицы  $A$  можно переставить на места  $(1, 1)$  и  $(2, 2)$ . Доказательство проведём индукцией по числу диагональных элементов, не равных  $m = \operatorname{tr} A/n$ ; число таких элементов не может быть равно 1. Для работы с такими элементами нам потребуется следующее утверждение.

**Лемма.** Пусть  $A$  — комплексная (вещественная) матрица, диагональ которой имеет вид  $(a_1, a_2, \dots, a_n)$ . Тогда для любого  $t$ , где  $0 \leq t \leq 1$ , существует такая унитарная (ортогональная) матрица  $U$ , что диагональ матрицы  $UAU^{-1}$  имеет вид  $(a'_1, a'_2, a_3, \dots, a_n)$ , где  $a'_1 = ta_1 + (1-t)a_2$  и  $a'_2 = (1-t)a_1 + ta_2$ .

**Доказательство.** Так как

$$\begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} U^{-1} & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} UPU^{-1} & UQ \\ RU^{-1} & S \end{pmatrix},$$

то можно ограничиться матрицами порядка 2. Унитарная матрица порядка 2 с определителем 1 имеет вид  $\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$ , где  $|u|^2 + |v|^2 = 1$  (см. задачу 4.2). Легко проверить, что

$$\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} \begin{pmatrix} a_1 & b \\ c & a_2 \end{pmatrix} \begin{pmatrix} \bar{u} & -v \\ \bar{v} & u \end{pmatrix} = \begin{pmatrix} a_1|u|^2 + b\bar{v}u + c\bar{u}v + a_2|v|^2 & * \\ * & * \end{pmatrix}.$$

В вещественном случае величина  $a_1 \cos^2 \alpha + (b+c) \cos \alpha \sin \alpha + a_2 \sin^2 \alpha$  при изменении  $\alpha$  от 0 до  $\pi/2$  изменяется от  $a_1$  до  $a_2$ , поэтому в вещественном случае доказательство леммы завершено.

Пусть  $u = \cos \alpha e^{i\varphi}$ ,  $v = \sin \alpha e^{i\psi}$ . Тогда  $f(u, v) = a_1|u|^2 + b\bar{v}u + c\bar{u}v + a_2|v|^2 = a_1 \cos^2 \alpha + a_2 \sin^2 \alpha + \sin \alpha \cos \alpha (be^{i\beta} + ce^{-i\beta})$ , где  $\beta = \varphi - \psi$ . Точки  $be^{i\beta} + ce^{-i\beta}$  при  $0 \leq \beta \leq 2\pi$  образуют на комплексной плоскости

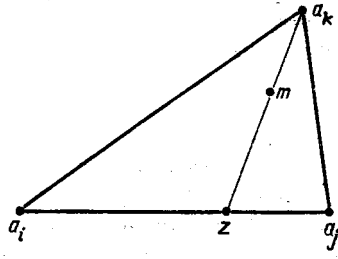


Рис. 3.2.

эллипс (или отрезок) с центром в нуле. В самом деле, точки  $e^{i\beta}$  лежат на единичной окружности, а отображение  $z \mapsto bz + c\bar{z}$  задаёт вещественно линейное отображение комплексной плоскости (это отображение может быть вырожденным). Поэтому величина  $(be^{i\beta} + ce^{-i\beta})/(a_1 - a_2)$  принимает вещественное значение  $p$  для некоторого  $\beta$ . При таком  $\beta$  получаем  $f(u, v) = ta_1 + (1 - t)a_2$ , где  $t = \cos^2 \alpha + p \sin \alpha \cos \alpha$ . При изменении  $\alpha$  от 0 до  $\pi/2$  величина  $t$  изменяется от 1 до 0. Доказательство леммы завершено.  $\square$

В случае двух диагональных элементов, не равных  $m = \operatorname{tr} A/n$ , доказательство теоремы непосредственно следует из леммы. Предположим теперь, что количество диагональных элементов матрицы  $A$ , не равных  $m$ , больше двух. Тогда существуют такие диагональные элементы  $a_i, a_j$  и  $a_k$ , что точка  $m = \operatorname{tr} A/n$  на комплексной плоскости лежит внутри треугольника  $a_i a_j a_k$ . Пусть  $z$  — точка пересечения прямых  $ma_k$  и  $a_i a_j$  (рис. 3.2). Согласно лемме матрицу с диагональю  $(a_i, a_j, \dots)$  можно заменить на матрицу с диагональю  $(z, \dots)$ , а матрицу с диагональю  $(z, a_k, \dots)$  можно заменить на матрицу с диагональю  $(m, \dots)$ ; при этом все диагональные элементы, кроме  $a_i, a_j$  и  $a_k$ , не изменятся. В результате количество диагональных элементов, равных  $m$ , увеличится по крайней мере на 1.  $\square$

### 3.5.3. Все диагональные элементы ненулевые

**Теорема 3.5.3.1.** *Любая ненулевая матрица  $A$  порядка  $n$  подобна матрице, все диагональные элементы которой отличны от нуля.*

**Доказательство.** [Ма4] Любая матрица  $A$  подобна матрице, диагональные элементы которой равны  $\operatorname{tr} A/n$  (теорема 3.5.2.1), поэтому остаётся

рассмотреть случай, когда  $\text{tr } A = 0$ . Можно считать, что матрица  $A$  жорданова. Рассмотрим сначала матрицу  $A = \|a_{ij}\|_1^n$ , где  $a_{ij} = \delta_{1i}\delta_{2j}$ . Если  $U = \|u_{ij}\|$  — унитарная матрица, то  $UAU^{-1} = UAU^* = B$ , где  $b_{ii} = u_{i1}\bar{u}_{i2}$ . Ясно, что матрицу  $U$  можно подобрать так, что  $b_{ii} \neq 0$ .

Дальнейшее доказательство проведём индукцией по  $n$ ; при  $n = 2$  утверждение доказано. Предположим сначала, что матрица  $A$  диагональная. Можно считать, что  $a_{11} \neq 0$ . Тогда  $A = \begin{pmatrix} a_{11} & 0 \\ 0 & \Lambda \end{pmatrix}$ , где  $\Lambda$  — ненулевая диагональная матрица. Пусть матрица  $U$  такова, что все диагональные элементы матрицы  $U\Lambda U^{-1}$  ненулевые. Тогда диагональные элементы матрицы

$$\begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & \Lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U^{-1} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & U\Lambda U^{-1} \end{pmatrix}$$

ненулевые.

Предположим теперь, что матрица  $A$  не диагональная. Можно считать, что  $a_{12} \neq 0$  и матрица  $C$ , полученная из матрицы  $A$  вычёркиванием первой строки и первого столбца, ненулевая. Пусть матрица  $U$  такова, что все диагональные элементы матрицы  $UCU^{-1}$  ненулевые. Рассмотрим матрицу

$$D = \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} A \begin{pmatrix} 1 & 0 \\ 0 & U^{-1} \end{pmatrix} = \begin{pmatrix} a_{11} & * \\ * & UCU^{-1} \end{pmatrix}.$$

Единственным ненулевым диагональным элементом матрицы  $D$  может быть элемент  $a_{11}$ . Если  $a_{11} = 0$ , то для матрицы  $\begin{pmatrix} 0 & * \\ * & d_{22} \end{pmatrix}$  подберём такую матрицу  $V$ , что диагональные элементы матрицы  $V \begin{pmatrix} 0 & * \\ * & d_{22} \end{pmatrix} V^{-1}$  ненулевые. Тогда все диагональные элементы матрицы  $\begin{pmatrix} V & 0 \\ 0 & I \end{pmatrix} D \begin{pmatrix} V^{-1} & 0 \\ 0 & I \end{pmatrix}$  ненулевые.  $\square$

### Задачи

**5.1.** Пусть  $A$  — вещественная матрица с нулевым следом. Докажите, что существует ненулевой вектор-столбец  $x$ , для которого  $x^T A x = 0$ .

**5.2.** Докажите, что для любой ненулевой квадратной матрицы  $A$  найдётся такая матрица  $X$ , что у матриц  $X$  и  $A + X$  нет общих собственных значений.

## 3.6. Полярное разложение

Любое комплексное число  $z$  можно представить в виде  $z = |z|e^{i\varphi}$ . Аналогом этого представления является *полярное разложение* матрицы



$A = SU$ , где  $S$  — эрмитова,  $U$  — унитарная матрица.

### 3.6.1. Существование полярного разложения

**Теорема 3.6.1.1.** Любую квадратную матрицу  $A$  над  $\mathbb{R}$  (над  $\mathbb{C}$ ) можно представить в виде  $A = SU$ , где  $S$  — симметрическая (эрмитова),  $U$  — ортогональная (унитарная) матрица. Если матрица  $A$  невырожденная, то такое представление единственно.

**Доказательство.** Если  $A = SU$ , где  $S$  — эрмитова,  $U$  — унитарная матрица, то  $AA^* = SUU^*S = S^2$ . Если матрица  $S$  к тому же неотрицательно определена, то равенство  $S^2 = AA^*$  задаёт её однозначно (теорема 4.1.4.1). Для невырожденной матрицы  $A$  остаётся положить  $U = S^{-1}A$ . Матрица  $U$  унитарна, так как  $UU^* = S^{-1}AA^*S^{-1} = I$ . Отметим, что единственной полярного разложения невырожденной матрицы можно доказать, не обращаясь к теореме 4.1.4.1. А именно, предположим, что  $S_1U_1 = S_2U_2$ , где матрицы  $S_1$  и  $S_2$  положительно определённые, а матрицы  $U_1$  и  $U_2$  унитарные. Тогда  $S_2^{-1}S_1 = U_2U_1^{-1}$ . Матрица  $U_2U_1^{-1}$  унитарная, поэтому все её собственные значения по модулю равны 1. С другой стороны, согласно теореме 4.1.6.1 собственные значения матрицы  $S_2^{-1}S_1$  положительны. Поэтому все собственные значения матрицы  $U_2U_1^{-1}$  равны 1. Эта матрица диагонализируема, поэтому она равна  $I$ . Таким образом,  $U_1 = U_2$ , а значит,  $S_1 = S_2$ .

В случае вырожденной матрицы  $A$  требуются более тонкие рассуждения. Для оператора  $AA^*$  существует ортонормированный собственный базис  $z_1, \dots, z_n$ , причём  $AA^*z_i = k_i^2z_i$ , где  $k_i \in \mathbb{R}$  (теорема 3.1.3.2). Положим  $Sz_i = k_iz_i$ , где  $k_i \in \mathbb{R}$ . Можно считать, что числа  $k_1, \dots, k_m$  положительны и  $k_{m+1} = \dots = k_n = 0$ . Пусть  $w_i = A^*z_i/k_i$  для  $i \leq m$  и  $W = \langle w_1, \dots, w_m \rangle$ . Так как  $k_ik_j(w_i, w_j) = (A^*z_i, A^*z_j) = (z_i, AA^*z_j) = k_j^2(z_i, z_j)$ , то  $w_1, \dots, w_m$  — ортонормированный базис пространства  $W$ . Дополним векторы  $w_1, \dots, w_m$  до ортонормированного базиса  $w_1, \dots, w_n$  и положим  $Uw_i = z_i$ . Докажем, что  $A = SU$ . Для  $i \leq m$   $SUw_i = Sz_i = k_iz_i$  и  $Aw_i = AA^*z_i/k_i = k_iz_i$ . Для  $i > m$  получаем  $SUw_i = 0$  и  $(Aw_i, Aw_i) = (w_i, A^*Aw_i) = 0$ , так как  $A^*Aw_i \in \text{Im } A^*A \subset \text{Im } A^* = W$ .  $\square$

**Замечание.** Матрицу  $A^*$  можно представить в виде  $A^* = SU$ , поэтому  $A = (A^*)^* = U^*S$ , где матрица  $U^*$  унитарна.

### 3.6.2. Разложение $UDW$

**Теорема 3.6.2.1.** Любую матрицу  $A$  можно представить в виде  $A = UDW$ , где  $U$  и  $W$  — унитарные матрицы,  $D$  — диагональная матрица.

**Доказательство.** Пусть  $A = SV$ , где  $S$  — эрмитова,  $V$  — унитарная матрица. Для матрицы  $S$  существует такая унитарная матрица  $U$ , что  $S = UDU^*$  (теорема 3.1.3.2), где  $D$  — диагональная матрица. Матрица  $W = U^*V$  унитарна и  $A = SV = UDW$ .  $\square$

**Теорема 3.6.2.2.** Если  $A = S_1 U_1 = U_2 S_2$  — полярные разложения невырожденной матрицы  $A$ , то  $U_1 = U_2$ .

**Доказательство.** Пусть  $A = UDW$ , где  $U$  и  $W$  — унитарные матрицы, а  $D = \text{diag}(d_1, \dots, d_n)$ . Рассмотрим матрицу  $D_+ = \text{diag}(|d_1|, \dots, |d_n|)$ . Тогда  $DD_+ = D_+D$ , а значит,

$$A = (UD_+U^*)(UD_+^{-1}DW) = (UD_+^{-1}DW)(W^*D_+W).$$

Матрицы  $UD_+U^*$  и  $W^*D_+W$  положительно определённые, а матрица  $D_+^{-1}D$  унитарная. Из единственности полярного разложения невырожденной матрицы следует, что  $S_1 = UD_+U^*$ ,  $S_2 = W^*D_+W$  и  $U_1 = UD_+^{-1}DW = U_2$ .  $\square$

### Задачи

**6.1.** Докажите, что любое линейное преобразование  $\mathbb{R}^n$  является композицией ортогонального преобразования и растяжения по перпендикулярным направлениям (с разными коэффициентами).

**6.2.** Пусть  $A$  — квадратная матрица над полем  $\mathbb{C}$ ,  $f$  — многочлен над полем  $\mathbb{C}$ . Докажите, что если  $f(AA^*) = 0$ , то  $f(A^*A) = 0$ .

**6.3.** Пусть  $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$  — сжимающий оператор, т.е.  $\|Ax\| \leq \|x\|$ . Пространство  $\mathbb{R}^n$  вложено в  $\mathbb{R}^{2n}$ . Докажите, что  $A$  является ограничением на  $\mathbb{R}^n$  композиции ортогонального преобразования  $\mathbb{R}^{2n}$  и проекции на  $\mathbb{R}^n$ .

## 3.7. Разложения матриц

### 3.7.1. Разложение Шура

**Теорема 3.7.1.1 (Шур).** Любую квадратную матрицу  $A$  над  $\mathbb{C}$  можно представить в виде  $A = UTU^*$ , где  $U$  — унитарная,  $T$  — треугольная матрица,

причём матрица  $A$  нормальна тогда и только тогда, когда  $T$  — диагональная матрица.

**Доказательство.** Применим индукцию по порядку матрицы  $A$ . Пусть  $x$  — собственный вектор матрицы  $A$ , т.е.  $Ax = \lambda x$ . Можно считать, что  $\|x\| = 1$ . Пусть  $W$  — унитарная матрица, первым столбцом которой служат координаты вектора  $x$  (для построения такой матрицы достаточно дополнить вектор  $x$  до ортонормированного базиса). Тогда

$$W^*AW = \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

По предположению индукции существует такая унитарная матрица  $V$ , что матрица  $V^*A_1V$  треугольная. Тогда матрица  $U = W \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$  искомая.

Легко проверить, что равенства  $T^*T = TT^*$  и  $A^*A = AA^*$  эквивалентны. Остаётся доказать, что треугольная нормальная матрица является диагональной. Пусть

$$T = \begin{pmatrix} t_{11} & \dots & t_{1n} \\ \mathbf{0} & \ddots & \vdots \\ & & t_{nn} \end{pmatrix}.$$

Тогда  $(TT^*)_{11} = |t_{11}|^2 + |t_{12}|^2 + \dots + |t_{1n}|^2$  и  $(T^*T)_{11} = |t_{11}|^2$ . Поэтому из равенства  $TT^* = T^*T$  следует, что  $t_{12} = \dots = t_{1n} = 0$ . Из матрицы  $T$  можно вычеркнуть первую строку и первый столбец и повторить эти рассуждения.  $\square$

### 3.7.2. Разложение Ланцоша

**Теорема 3.7.2.1.** [La] Любую вещественную матрицу  $A$  размера  $m \times n$  и ранга  $p > 0$  можно представить в виде  $A = X\Lambda Y^T$ , где  $X$  и  $Y$  — матрицы размеров  $m \times p$  и  $n \times p$  с ортонормированными столбцами,  $\Lambda$  — диагональная матрица порядка  $p$ .

**Доказательство.** [Sc5] Ранг матрицы  $A^TA$  равен рангу  $A$  (см. задачу 4.3). Пусть  $U$  — ортогональная матрица, приводящая симметрическую матрицу  $A^TA$  к диагональному виду, т.е.  $U^TA^TAU = \text{diag}(\mu_1, \dots, \mu_p, 0, \dots, 0)$ , причём  $\mu_i > 0$ . Пусть, далее,  $y_1, \dots, y_p$  — первые  $p$  столбцов матрицы  $U$ ;  $Y$  — матрица, составленная из этих столбцов. Столбцы  $x_i = \lambda_i^{-1}Ay_i$ , где  $\lambda_i = \sqrt{\mu_i}$ , образуют ортонормированную

систему, так как  $(Ay_i, Ay_j) = (y_i, A^T Ay_j) = \lambda^2(y_i, y_j)$ . Ясно также, что  $AY = (\lambda_1 x_1, \dots, \lambda_p x_p) = X\Lambda$ , где  $X$  — матрица, составленная из столбцов  $x_1, \dots, x_p$ ,  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ . Докажем теперь, что  $A = X\Lambda Y^T$ . Рассмотрим для этого снова матрицу  $U = (Y, Y_0)$ . Так как  $\text{Ker } A^T A = \text{Ker } A$  (см. решение задачи 4.3) и  $(A^T A)Y_0 = 0$ , то  $AY_0 = 0$ . Следовательно,  $AU = (X\Lambda, 0)$ , а значит,  $A = (X\Lambda, 0)U^T = X\Lambda Y^T$ .  $\square$

**Замечание.** Так как  $AU = (X\Lambda, 0)$ , то  $U^T A^T = \begin{pmatrix} \Lambda X^T \\ 0 \end{pmatrix}$ . Домножив это равенство на  $U$ , получим  $A^T = Y\Lambda X^T$ . Следовательно,  $A^T X = Y\Lambda X^T X = Y\Lambda$ , так как  $X^T X = I_p$ . Поэтому  $(X^T A)(A^T X) = (\Lambda Y^T)(Y\Lambda) = \Lambda^2$ , так как  $Y^T Y = I_p$ . Значит, столбцы матрицы  $X$  являются собственными векторами матрицы  $AA^T$ .

### 3.7.3. Произведение двух симметрических матриц

**Теорема 3.7.3.1.** а) Любая комплексная матрица является произведением двух комплексных симметрических матриц.

б) Любая вещественная матрица является произведением двух вещественных симметрических матриц.

**Доказательство.** а) Заметим сначала, что

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \lambda \\ 0 & \lambda & 1 \\ \lambda & 1 & 0 \end{pmatrix};$$

аналогичные равенства справедливы и для жордановых клеток любого размера. Следовательно, жорданову матрицу  $J$  можно представить в виде произведения двух симметрических матриц, а значит, любую матрицу  $A$  можно представить в виде  $A = PJP^{-1}$ , где  $J = ST$  — произведение двух симметрических матриц. Тогда  $A = S_1 T_1$ , где  $S_1 = PSP^*$  и  $T_1 = (P^*)^{-1} TP^{-1}$  — симметрические матрицы.

б) В вещественном случае матрицу тоже можно привести к виду, аналогичному жордановой нормальной форме (теорема 3.2.4.2). Для вещественных собственных значений предыдущее доказательство проходит без изменений. В случае собственных значений  $\alpha \pm i\beta$  заметим, что

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix},$$

поэтому соответствующую вещественную жорданову клетку можно

представить в виде

$$\begin{pmatrix} 0 & 0 & I \\ 0 & I & 0 \\ I & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \Lambda \\ 0 & \Lambda & I \\ \Lambda & I & 0 \end{pmatrix},$$

где

$$I = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{и} \quad \Lambda = \begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix}$$

(для клеток произвольного размера справедливо аналогичное равенство). Дальнейшее доказательство аналогично комплексному случаю.  $\square$

## Задачи

**7.1.** Все угловые миноры квадратной матрицы  $A$  отличны от нуля. Докажите, что её можно представить в виде  $A = LU$ , где  $L$  — нижняя треугольная,  $U$  — верхняя треугольная матрица (*разложение Гаусса*).

**7.2.** Докажите, что невырожденную матрицу  $X$  можно представить в виде  $X = UT$ , где  $U$  — ортогональная матрица,  $T$  — верхняя треугольная матрица (*разложение Грама*).

**7.3.** Докажите, что положительно определённую матрицу  $S$  можно представить в виде  $S = LL^T$ , где  $L$  — нижняя треугольная матрица (*разложение Холецкого*).

**7.4.** [Ra1] Пусть  $B = \text{diag}(1, \varepsilon, \dots, \varepsilon^{n-1})$ , где  $\varepsilon$  — первообразный корень  $n$ -й степени из 1;  $C = \|c_{ij}\|_1^n$ , где  $c_{ij} = \delta_{i,j-1}$  ( $j-1$  берётся по модулю  $n$ ). Докажите, что любая матрица  $M$  над  $\mathbb{C}$  порядка  $n$  однозначно представима в виде  $\sum_{k,l=0}^{n-1} a_{kl} B^k C^l$ .

**7.5.** Докажите, что любую кососимметрическую матрицу  $A$  можно представить в виде  $A = S_1 S_2 - S_2 S_1$ , где  $S_1$  и  $S_2$  — симметрические матрицы.

## 3.8. Нормальная форма Смита. Элементарные делители матриц

### 3.8.1. Нормальная форма Смита

Нормальную форму Смита мы определим в следующих двух основных ситуациях: 1) элементы матрицы являются целыми числами; 2) эле-

менты матрицы являются многочленами с действительными или комплексными коэффициентами.

Пусть  $f_k(A)$  — наибольший общий делитель<sup>1</sup> всех миноров порядка  $k$  матрицы  $A$ . Из формулы разложения определителя по строке видно, что  $f_k$  делится на  $f_{k-1}$ .

Пусть  $A$  — квадратная матрица, причём  $\det A \neq 0$ . Формула  $A^{-1} = (\text{adj } A) / \det A$  показывает, что если  $\det A = \pm 1$  (соответственно  $\det A$  — ненулевое число), то элементами матрицы  $A^{-1}$  будут целые числа (соответственно многочлены). Наоборот, если элементы матрицы  $A^{-1}$  являются целыми числами (соответственно многочленами), то  $\det A = \pm 1$  (соответственно  $\det A$  — ненулевое число). Действительно,  $\det A$  и  $\det A^{-1}$  — целые числа (соответственно многочлены) и  $\det A \cdot \det A^{-1} = \det(AA^{-1}) = 1$ .

Матрицу  $A$ , элементами которой являются целые числа (соответственно многочлены) называют *матрицей-единицей*, если  $\det A = \pm 1$  (соответственно  $\det A$  — ненулевое число). Ясно, что произведение двух матриц-единиц является матрицей-единицей.

**Теорема 3.8.1.1.** Если  $A' = BAC$ , где  $B$  и  $C$  — матрицы-единицы, то  $f_k(A') = f_k(A)$  для всех допустимых  $k$ .

**Доказательство.** Из формулы Бине–Коши следует, что  $f_k(A')$  делится на  $f_k(A)$ . А так как  $A = B^{-1}A'C^{-1}$ , то  $f_k(A)$  делится на  $f_k(A')$ .  $\square$

**Теорема 3.8.1.2 (Смит).** Для любой прямоугольной матрицы  $A$  можно выбрать матрицы-единицы  $B$  и  $C$  так, что  $BAC = \text{diag}(g_1, g_2, \dots, g_p, 0, \dots, 0)$ , причём  $g_{i+1}$  делится на  $g_i$  при  $i = 1, \dots, p-1$  (здесь подразумевается, что диагональная матрица дополнена до прямоугольной матрицы нулями).

**Доказательство.** Фиксируем два различных целых числа  $p$  и  $q$  ( $1 \leq p, q \leq n$ ). Матрица-единица  $\|\alpha_{ij}\|_1^n$ , где  $\alpha_{ii} = 1$  при  $i \neq p, q$  и  $\alpha_{pq} = \alpha_{qp} = 1$ , а все остальные элементы равны нулю, при умножении на неё справа (слева) осуществляет перестановку столбцов (строк) с номерами  $p$  и  $q$ . Матрица-единица  $\|\alpha_{ij}\|_1^n$ , где  $\alpha_{ii} = 1$  и  $\alpha_{pq} = f$ , а все остальные элементы равны нулю, при умножении на неё справа осуществляет прибавление к  $q$ -му столбцу  $p$ -го столбца, умноженного на  $f$ , а при умножении

<sup>1</sup>Чтобы получить однозначно определённое число (многочлен) в качестве наибольшего общего делителя мы берём неотрицательное число (многочлен со старшим коэффициентом 1).

слева — прибавление к  $p$ -й строке  $q$ -й строки, умноженной на  $f$ . Остается проверить, что такими операциями данную матрицу  $A$  можно привести к требуемому виду.

*Нормой* целого числа будем называть его абсолютную величину, а *нормой* многочлена — его степень. Выберем в данной матрице ненулевой элемент  $a$  с наименьшей нормой и переставим его на место  $(1, 1)$ . Поделим элементы первой строки на  $a$  с остатком и прибавим к столбцам  $2, 3, \dots, n$  первый столбец, умноженный на подходящие целые числа (многочлены), так, чтобы в первой строке получились остатки от деления на  $a$ . Затем сделаем аналогичную операцию, чтобы в первом столбце получились остатки от деления на  $a$ . Если при этом в первом столбце или в первой строке кроме  $a$  есть ещё один ненулевой элемент, то его норма строго меньше нормы  $a$ . Переставим этот элемент на место  $(1, 1)$  и повторим указанные операции. Норма элемента, стоящего на месте  $(1, 1)$ , строго убывает, поэтому в конце концов в первой строке и в первом столбце останется единственный ненулевой элемент  $a_{11}$ .

Предположим, что в полученной матрице есть элемент  $a_{ij}$ , не делящийся на  $a_{11}$ . Прибавим к первому столбцу  $j$ -й столбец (содержащий  $a_{ij}$ ). Затем к  $i$ -й строке, содержащей теперь два элемента  $a_{ij}$ , прибавим первую строку, умноженную на подходящее целое число (многочлен), так, чтобы один из элементов  $a_{ij}$  заменился на остаток от деления на  $a_{11}$ . В результате получим элемент, норма которого строго меньше нормы  $a_{11}$ . Переставим его на место  $(1, 1)$  и повторим все указанные операции. В конце концов получим матрицу вида  $\begin{pmatrix} g_1 & 0 \\ 0 & A' \end{pmatrix}$ , причём все элементы матрицы  $A'$  делятся на  $g_1$ . Для матрицы  $A'$  можно повторить все предыдущие рассуждения.  $\square$

**Замечание.** Ясно, что  $f_k(A) = g_1 g_2 \cdots g_k$ , поэтому  $g_k = f_k(A)/f_{k-1}(A)$ . Следовательно числа (многочлены)  $g_k$  определены однозначно.

### 3.8.2. Элементарные делители матриц

Полученные в нормальной форме Смита элементы  $g_1, \dots, g_p$  называют *инвариантными множителями* матрицы  $A$ . Они выражаются через делители миноров  $f_k(A)$  следующим образом:  $g_k = f_k/f_{k-1}$ , если  $f_{k-1} \neq 0$ .

Каждый инвариантный множитель  $g_i$  можно разложить в произведение степеней простых чисел (соответственно степеней неприводимых многочленов). Такие множители называют *элементарными делителями* матрицы  $A$ . В набор элементарных делителей каждый множитель входит с учётом кратности.

Элементарными делителями вещественной или комплексной матрицы  $A$  называют элементарные делители матрицы  $A - xI$ . Произведение всех элементарных делителей матрицы  $A$  с точностью до знака равно её характеристическому многочлену.

### 3.8.3. Целочисленные решения систем линейных уравнений

Нормальную форму Смита можно применить для решения систем линейных уравнений в целых числах. Пусть  $A$  — матрица с целочисленными элементами, и нас интересуют целочисленные решения уравнения  $Ax = b$ . Если  $A' = BAC$ , где  $B$  и  $C$  — матрицы-единицы, то данное уравнение эквивалентно уравнению  $A'x' = b'$ , где  $b' = Bb$  и  $x = Cx'$ . Действительно, уравнение  $Ax = b$  эквивалентно уравнению  $BACx' = Bb$ . Для матрицы  $A' = \text{diag}(g_1, \dots, g_p, 0, \dots, 0)$  целочисленные решения уравнения  $A'x' = b'$  легко находятся: они имеют вид  $x'_1 = b'_1/g_1, \dots, x'_p = b'_p/g_p, x'_{p+1} = t_1, x'_{p+2} = t_2, \dots$ , где  $t_1, t_2, \dots$  принимают любые целые значения (если хотя бы одно из чисел  $b'_k$  не делится на  $g_k$ , то целочисленных решений нет). Решения исходного уравнения получаются по формуле  $x = Cx'$ .

Покажем, как работает этот алгоритм, на примере системы линейных уравнений

$$\begin{cases} x_1 + 5x_2 + 3x_3 = 8, \\ 7x_1 - 13x_2 + 12x_3 = 2. \end{cases}$$

Основная трудность состоит в вычислении матриц  $B$  и  $C$  для матрицы  $A = \begin{pmatrix} 1 & 5 & 3 \\ 7 & -13 & 12 \end{pmatrix}$ . Мы начинаем с того, что вычитаем из второй строки первую строку, умноженную на 7, т.е. рассматриваем преобразование

$$\begin{pmatrix} 1 & 0 \\ -7 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 & 3 \\ 7 & -13 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 3 \\ 0 & -48 & -9 \end{pmatrix}.$$

Затем с помощью первого столбца уничтожаем второй и третий элемент первой строки, а заодно меняем знак второй строки:

$$\begin{pmatrix} 1 & 5 & 3 \\ 0 & -48 & -9 \end{pmatrix} \begin{pmatrix} 1 & 5 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 48 & 9 \end{pmatrix}.$$

Оставшиеся преобразования относятся только к строке  $(48 \ 9)$ ; они за-



ключаются в нахождении наибольшего общего делителя чисел 48 и 9:

$$\begin{aligned}(48 \ 9) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} &= (9 \ 48); \\ (9 \ 48) \begin{pmatrix} 1 & -5 \\ 0 & 1 \end{pmatrix} &= (9 \ 3); \\ (9 \ 3) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} &= (3 \ 9); \\ (3 \ 9) \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} &= (3 \ 0).\end{aligned}$$

Легко проверить, что

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -5 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ -5 & 16 \end{pmatrix}.$$

Таким образом,

$$C = \begin{pmatrix} 1 & 5 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & -5 & 16 \end{pmatrix} = \begin{pmatrix} 1 & -10 & 33 \\ 0 & -1 & 3 \\ 0 & 5 & -16 \end{pmatrix}$$

и  $B = \begin{pmatrix} 1 & 0 \\ -7 & 1 \end{pmatrix}.$

Теперь решение находится следующим образом. Прежде всего заметим, что  $b' = \begin{pmatrix} 1 & 0 \\ -7 & 1 \end{pmatrix} \begin{pmatrix} 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ -54 \end{pmatrix}$ . Поэтому  $x' = (8, -54/3, t)^T$  и

$$x = \begin{pmatrix} 1 & -10 & 33 \\ 0 & -1 & 3 \\ 0 & 5 & -16 \end{pmatrix} \begin{pmatrix} 8 \\ -18 \\ t \end{pmatrix} = \begin{pmatrix} 188 + 33t \\ 18 + 3t \\ -90 - 16t \end{pmatrix}.$$

Положив  $18 + 3t = 3s$ , т.е.  $t = s - 6$ , можно несколько упростить полученное решение:  $x_1 = -10 + 33s$ ,  $x_2 = 3s$ ,  $x_3 = 6 - 16s$ .

### 3.8.4. Классификация абелевых групп

**Теорема 3.8.4.1.** Любая конечно порождённая абелева группа имеет вид  $\mathbb{Z}^r \oplus \mathbb{Z}_{g_1} \oplus \dots \oplus \mathbb{Z}_{g_p}$ , где  $g_{i+1}$  делится на  $g_i$ .

**Доказательство.** Пусть абелева группа  $G$  порождена элементами  $a_1, \dots, a_n$ . Возьмём свободную абелеву группу  $\mathbb{Z}^n$  с базисом  $e_1, \dots, e_n$  и рассмотрим гомоморфизм  $f: \mathbb{Z}^n \rightarrow G$ , заданный формулой  $f(\sum k_i e_i) = \sum k_i a_i$ . Ясно, что  $f$  — эпиморфизм, поэтому  $G \cong \mathbb{Z}^n / \text{Ker } f$ .

Группа  $\text{Ker } f$  является свободной абелевой группой (теорема 2.3.4.1). Пусть  $\varepsilon_1, \dots, \varepsilon_m$  — её базис,  $\varepsilon$  — столбец векторов  $\varepsilon_1, \dots, \varepsilon_m$ , а  $e$  — столбец векторов  $e_1, \dots, e_n$ . Тогда  $\varepsilon = Ae$ , где  $A$  — матрица размером  $m \times n$ , элементами которой являются целые числа. Выберем матрицы-единицы  $B$  и  $C$  так, что  $BAC = A'$ , где  $A' = \text{diag}(g_1, \dots, g_p, 0, \dots, 0)$ . Тогда  $\varepsilon' = A'e'$ , где  $\varepsilon' = B\varepsilon$  и  $e' = C^{-1}e$  — базисы групп  $\text{Ker } f$  и  $\mathbb{Z}^n$ . Таким образом, в группе  $\mathbb{Z}^n$  можно выбрать базис  $e'_1, \dots, e'_n$  так, что базис группы  $\text{Ker } f$  имеет вид  $\varepsilon'_1 = g_1 e'_1, \dots, \varepsilon'_p = g_p e'_p$ . При таком выборе базисов ясно, что  $\mathbb{Z}^n / \text{Ker } f \cong \mathbb{Z}^r \oplus \mathbb{Z}_{g_1} \oplus \dots \oplus \mathbb{Z}_{g_p}$ .  $\square$

## Задачи

**8.1.** Вычислите инвариантные множители жордановой клетки и циклической клетки.

**8.2.** Пусть  $A$  — матрица порядка  $n$ ,  $f_{n-1}$  — наибольший общий делитель миноров порядка  $n-1$  матрицы  $A - xI$ . Докажите, что минимальный многочлен матрицы  $A$  равен  $\pm |A - xI| / f_{n-1}$ .

## Решения

**1.1.** Оба эти выражения равны следу оператора  $A$  (см. §3.1.1).

**1.2.** а) Нет, не существуют. Действительно,  $\text{tr}(AB - BA) = \text{tr } AB - \text{tr } BA = 0$ , а  $\text{tr } I = n$ .

б) Предположим, что матрица  $A$  невырожденная. Тогда  $A^{-1}AB - A^{-1}BA = I$ . Но  $\text{tr}(B - A^{-1}BA) = 0$ , а  $\text{tr } I = n$ .

**1.3.** Пусть  $\text{Im } A = \langle e_1 \rangle$ . В базисе  $e_1, \dots, e_n$  матрица оператора  $AB$  имеет вид  $\begin{pmatrix} x_1 & \dots & x_n \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{pmatrix}$ , причём  $x_1 = \text{tr } AB = 0$ , т.е.  $ABe_1 = 0$ .

**1.4.** Пусть  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Легко проверить, что  $D^{-1}AD = \begin{pmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{pmatrix} = B$ . Пусть  $(x_1, \dots, x_n)$  — координаты вектора  $x$ . Тогда  $Bx = y$ , где  $y$  — столбец  $(a, \dots, a)$ , причём  $a = x_1 + \dots + x_n$ . Собственные векторы, соответствующие нулевому собственному значению, образуют  $(n-1)$ -мерное подпространство, заданное уравнением  $x_1 + \dots + x_n = 0$ ; собственный вектор  $(1, \dots, 1)$  соответствует собственному значению  $n$ . Ясно также, что если  $e$  — собственный вектор матрицы  $B$ , то  $De$  — собственный вектор матрицы  $A$ .

**1.5.** Пусть  $p(\lambda) = |\lambda + A|$  и  $q(\lambda) = |\lambda - A|$  — многочлены степени  $n$ . Существует число  $\lambda$ , не являющееся корнем ни одного из этих многочленов. Тогда матрицы  $B_1 = \lambda I + A$  и  $B_2 = -\lambda I + A$  невырожденные и  $A = (B_1 + B_2)/2$ .

**1.6.** Очевидно, что коэффициенты характеристического многочлена непрерывно зависят от элементов матрицы. Остается доказать, что корни многочлена  $p(x) = x^n + a_1x^{n-1} + \dots + a_n$  непрерывно зависят от  $a_1, \dots, a_n$ . Доказательство достаточно провести для нулевого корня (для ненулевого корня  $x_1$  можно рассмотреть замену переменных  $y = x - x_1$ ). Если  $p(0) = 0$ , то  $a_n = 0$ . Рассмотрим многочлен  $q(x) = x^n + b_1x^{n-1} + \dots + b_n$ , где  $|b_i - a_i| < \delta$ . Если  $x_1, \dots, x_n$  — корни многочлена  $q$ , то  $|x_1 \dots x_n| = |b_n| < \delta$ , а значит, модуль одного из корней многочлена  $q$  меньше  $\sqrt[n]{\delta}$ . Взяв  $\delta = \varepsilon^n$ , получим требуемое.

**1.7.** Если сумма элементов каждой строки матрицы  $A$  равна  $s$ , то  $Ae = se$ , где  $e$  — столбец  $(1, 1, \dots, 1)$ . Следовательно,  $A^{-1}(Ae) = A^{-1}(se)$ , а значит,  $A^{-1}e = (1/s)e$ , т.е. сумма элементов каждой строки матрицы  $A^{-1}$  равна  $1/s$ .

**1.8.** а) Пусть  $S_1, \dots, S_n$  — столбцы матрицы  $S$ . Тогда  $A(S_1 \dots) = (S_1 \dots)(\Lambda \dots)$ , где  $\Lambda$  — столбец  $(\lambda, 0, \dots, 0)$ . Приравнявая первые столбцы этих матриц, получаем  $AS_1 = \lambda S_1$ .

б) Первый столбец матрицы  $(S^{-1}AS)^T = S^T A^T (S^T)^{-1}$  имеет вид  $(\lambda, 0, 0, \dots, 0)$ , поэтому можно воспользоваться результатом задачи а).

**1.9.** Легко проверить, что  $|A - \lambda I| = \sum_{k=0}^n (-1)^k \lambda^k \Delta_{n-k}(A)$ , где  $\Delta_m(A)$  — сумма всех главных миноров порядка  $m$  матрицы  $A$ . Поэтому  $\sum_{i=1}^n |A_i - \lambda I| = \sum_{i=1}^n \sum_{k=0}^{n-1} (-1)^k \lambda^k \Delta_{n-k-1}(A_i)$ . Остается заметить, что  $\sum_{i=1}^n \Delta_p(A_i) = (n-p)\Delta_p(A)$ . В самом деле, любой главный минор порядка  $p$  матрицы  $A$  является главным минором для  $n-p$  матриц  $A_i$ .

**1.10.** Так как  $\text{adj}(P^{-1}XP) = P^{-1}(\text{adj } X)P$ , то можно считать, что матрица  $A$  жорданова. В этом случае матрица  $\text{adj } A$  верхняя треугольная (см. задачу 2.8); её диагональные элементы легко вычисляются.

**1.11.** Пусть  $S = \|\delta_{i,n-j}\|_0^n$  — матрица с единичной побочной диагональю. Тогда  $AS = \|b_{ij}\|_0^n$  и  $SA = \|c_{ij}\|_0^n$ , где  $b_{ij} = a_{i,n-j}$  и  $c_{ij} = a_{n-i,j}$ . Поэтому центральная симметричность матрицы  $A$  означает, что  $AS = SA$ . Легко проверить также, что вектор  $x$  симметричен, если  $Sx = x$ , и антисимметричен, если  $Sx = -x$ .

Пусть  $\lambda$  — собственное значение матрицы  $A$  и  $Ay = \lambda y$ , где  $y \neq 0$ . Тогда  $A(Sy) = S(Ay) = S(\lambda y) = \lambda S(y)$ . Если  $Sy = -y$ , то можно положить  $x = y$ . Если  $Sy \neq y$ , то можно положить  $x = y + Sy$ ; тогда  $Ax = \lambda x$  и  $Sx = x$ .

$$\begin{aligned} \text{1.12. а) } & \begin{vmatrix} A - \lambda I & B \\ B & A - \lambda I \end{vmatrix} \cdot \begin{vmatrix} I & I \\ I & -I \end{vmatrix} = \begin{vmatrix} A + B - \lambda I & A - B - \lambda I \\ A + B - \lambda I & -(A - B - \lambda I) \end{vmatrix} = \\ & \begin{vmatrix} A + B - \lambda I & A - B - \lambda I \\ 0 & -2(A - B - \lambda I) \end{vmatrix}. \\ \text{б) } & \begin{vmatrix} A - \lambda I & B \\ -B & A - \lambda I \end{vmatrix} \cdot \begin{vmatrix} I & iI \\ iI & I \end{vmatrix} = \begin{vmatrix} A + iB - \lambda I & i(A - iB - \lambda I) \\ i(A + iB - \lambda I) & A - iB - \lambda I \end{vmatrix} = \\ & \begin{vmatrix} A + iB - \lambda I & i(A - iB - \lambda I) \\ 0 & 2(A - iB - \lambda I) \end{vmatrix}. \end{aligned}$$

в)  $\begin{pmatrix} A & B \\ B & -A \end{pmatrix}^2 = \begin{pmatrix} A^2 + B^2 & AB - BA \\ BA - AB & A^2 + B^2 \end{pmatrix}$ , поэтому можно воспользоваться результатом задачи б).

**1.13.** Так как  $Ae_i = x_{n-i+1}e_{n-i+1}$  и  $Ae_{n-i+1} = x_ie_i$ , то подпространства  $V_i = \langle e_i, e_{n-i+1} \rangle$  инвариантны относительно оператора  $A$ . При  $i \neq n-i+1$  матрица ограничения оператора  $A$  на подпространство  $V_i$  имеет вид  $B = \begin{pmatrix} 0 & \lambda \\ \mu & 0 \end{pmatrix}$ . Собственные значения матрицы  $B$  равны  $\pm\sqrt{\lambda\mu}$ . Если  $\lambda\mu = 0$  и  $B$  диагонализуема, то  $B = 0$ . Поэтому матрица  $B$  диагонализуема тогда и только тогда, когда оба числа  $\lambda$  и  $\mu$  равны или не равны нулю одновременно.

Итак, матрица  $A$  диагонализуема тогда и только тогда, когда оба числа  $x_i$  и  $x_{n-i+1}$  одновременно равны или не равны нулю для всех  $i$ .

**1.14.** а) Предположим, что векторы-столбцы  $x_1, \dots, x_m$  соответствуют вещественным собственным значениям  $\alpha_1, \dots, \alpha_m$ . Пусть  $X = (x_1, \dots, x_m)$  и  $D = \text{diag}(\alpha_1, \dots, \alpha_m)$ . Тогда  $AX = XD$ , а так как  $D$  — вещественная матрица, то  $AXX^* = XDX^* = X(XD)^* = X(AX)^* = XX^*A^*$ . Если векторы  $x_1, \dots, x_m$  линейно независимы, то  $\text{rk } XX^* = \text{rk } X$  (см. задачу 4.3); поэтому в качестве матрицы  $S$  можно взять  $XX^*$ .

Предположим теперь, что  $AS = SA^*$  и  $S$  — неотрицательно определённая матрица ранга  $m$ . Тогда существует невырожденная матрица  $P$ , для которой  $S = PNP^*$ , где  $N = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$  (см. § 3.1.3). Домножим обе части равенства  $AS = SA^*$  слева на  $P^{-1}$ , а справа на  $(P^*)^{-1}$ ; в результате получим  $(P^{-1}AP)N = N(P^{-1}AP)^*$ . Пусть  $P^{-1}AP = B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ , где  $B_{11}$  — матрица порядка  $m$ . Так как  $BN = NB^*$ , то  $\begin{pmatrix} B_{11} & 0 \\ B_{21} & 0 \end{pmatrix} = \begin{pmatrix} B_{11}^* & B_{21}^* \\ 0 & 0 \end{pmatrix}$ , т.е.  $B = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}$ , где  $B_{11}$  — эрмитова матрица порядка  $m$ . Матрица  $B_{11}$  имеет  $m$  линейно независимых собственных векторов  $z_1, \dots, z_m$  с вещественными собственными значениями. А так как  $AP = PB$  и  $P$  — невырожденная матрица, то векторы  $P \begin{pmatrix} z_1 \\ 0 \end{pmatrix}, \dots, P \begin{pmatrix} z_m \\ 0 \end{pmatrix}$ , линейно независимы и являются собственными векторами  $A$ , соответствующими вещественным собственным значениям  $\alpha_1, \dots, \alpha_m$ .

б) Доказательство во многом аналогично задаче а); в нашем случае  $AXX^*A^* = AX(AX)^* = XD(XD)^* = XDD^*X^* = XX^*$ .

Если  $ASA^* = S$  и  $S = PNP^*$ , то  $P^{-1}APN(P^{-1}AP)^* = N$ , т.е.  $\begin{pmatrix} B_{11}B_{11}^* & B_{11}B_{21}^* \\ B_{21}B_{11}^* & B_{21}B_{21}^* \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$ . Поэтому  $B_{21} = 0$  и  $P^{-1}AP = B = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}$ , где матрица  $B_{11}$  унитарна.

**1.15.** Рассмотрим сначала случай, когда собственные значения некрратные. Из теоремы Гамильтона—Кэли следует, что  $A^{n+m} = \sum_{k=0}^{n-1} (-1)^k a_k A^{n-k-1}$  для некоторых чисел  $a_0, \dots, a_{n-1}$ . Эти числа можно найти, решив систему линей-

ных уравнений  $x_i^{n+m} = \sum_{k=0}^{n-1} (-1)^k a_k x_i^{n-k-1}$ ,  $i = 1, \dots, n$ . В результате получим

$$a_k = \frac{\begin{vmatrix} x_1^{n-1} & \dots & x_1^{n+m} & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ x_n^{n-1} & \dots & x_n^{n+m} & \dots & 1 \end{vmatrix}}{\begin{vmatrix} x_1^{n-1} & \dots & x_1^{n-k-1} & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ x_n^{n-1} & \dots & x_n^{n-k-1} & \dots & 1 \end{vmatrix}}.$$

Переставим в первом определителе столбец, состоящий из элементов  $x_i^{n+m}$ , на первое место; при этом знак определителя изменится на  $(-1)^k$ . Если этот знак опустить, то мы получаем функцию Шура, соответствующую разбиению  $\lambda_1 = m+1, \lambda_2 = 1, \dots, \lambda_{k+1} = 1$ . В обозначениях Фробениуса это разбиение записывается как  $(m|k)$ .

В случае кратных собственных значений требуемое равенство следует из того, что обе его части непрерывно зависят от элементов матрицы  $A$ .

**2.1.** Это разложение уже фактически было построено в доказательстве теоремы Жордана. Если  $V \supset \text{Im } A \supset \text{Im } A^2 \supset \dots \supset \text{Im } A^p = \text{Im } A^{p+1}$  и  $0 \subset \text{Ker } A \subset \dots \subset \text{Ker } A^p = \text{Ker } A^{p+1}$ , то должны выполняться равенства  $N = \text{Ker } A^p$  и  $S = \text{Im } A^p$ . Ясно, что  $N \cap S \subset \text{Ker } A \cap \text{Im } A^p = 0$  и  $\dim \text{Ker } A^p + \dim \text{Im } A^p = \dim V$ .

**2.2.** Пусть  $A$  — жорданова клетка порядка  $k$ . Легко проверить, что в этом случае  $S_k A = A^T S_k$ , где  $S_k = \|\delta_{i,k+1-j}\|_1^k$  — невырожденная матрица. Если  $A$  — прямая сумма жордановых клеток, то можно взять прямую сумму матриц  $S_k$ .

**2.3.** Матрица  $P^{-1}$  соответствует перестановке  $\sigma^{-1}$ , поэтому  $P^{-1} = \|q_{ij}\|_1^n$ , где  $q_{ij} = \delta_{\sigma(i),j}$ . Пусть  $P^{-1}AP = \|b_{ij}\|_1^n$ . Тогда  $b_{ij} = \sum_{s,t} \delta_{\sigma(i),s} a_{st} \delta_{t,\sigma(j)} = a_{\sigma(i)\sigma(j)}$ .

**2.4.** Оба равенства  $A^2 = A$  и  $\text{tr } A = 0$  сохраняются, если заменить матрицу  $A$  на  $P^{-1}AP$ , поэтому можно считать, что матрица  $A$  жорданова.

Из равенства  $A^2 = A$  следует, что любое собственное значение матрицы  $A$  удовлетворяет уравнению  $x^2 = x$ . Поэтому собственные значения равны 0 или 1. Тогда из равенства  $\text{tr } A = 0$  следует, что все собственные значения равны 0.

Итак, остаётся изучить ситуацию, когда  $A$  — жорданова клетка с нулевым собственным значением и выполняется равенство  $A^2 = A$ . Ясно, что это возможно только для клетки размером  $1 \times 1$ .

**2.5.** Пусть  $\lambda_i, \dots, \lambda_m$  — попарно различные собственные значения матрицы  $A$ ;  $p_i$  — кратность собственного значения  $\lambda_i$ . Тогда  $\text{tr}(A^k) = p_1 \lambda_1^k + \dots + p_m \lambda_m^k$ . Поэтому  $|b_{ij}|_0^{m-1} = p_1 \dots p_m \prod_{i>j} (\lambda_i - \lambda_j)^2$  (см. задачу 1.24). Для вычисления определителя  $|b_{ij}|_0^m$  можно, например, в выражении для  $\text{tr}(A^k)$  заменить  $p_m \lambda_m^k$  на сумму  $\lambda_m^k + (p_m - 1)\lambda_m^k$ .

**2.6.** Если  $A' = P^{-1}AP$ , то  $(A' + \lambda I)^{-1}A' = P^{-1}(A + \lambda I)^{-1}AP$ , поэтому достаточно рассмотреть случай, когда  $A$  — жорданова клетка. Если матрица  $A$  невырождена, то  $\lim_{\lambda \rightarrow 0} (A + \lambda I)^{-1} = A^{-1}$ . Пусть  $A = N$  — жорданова клетка с нулевым собственным значением. Тогда  $(N + \lambda I)^{-1}N = \lambda^{-1}(I - \lambda^{-1}N + \lambda^{-2}N^2 - \dots)N = \lambda^{-1}N - \lambda^{-2}N^2 + \dots$ ; предел при  $\lambda \rightarrow 0$  существует, только если  $N = 0$ .

Итак, чтобы указанный предел существовал, матрица  $A$  не должна иметь ненулевых жордановых клеток с нулевым собственным значением. Это условие эквивалентно тому, что  $\text{rk } A = \text{rk } A^2$ .

**3.1.** Пусть  $(\lambda_1, \dots, \lambda_n)$  — диагональ жордановой нормальной формы матрицы  $A$ ,  $\sigma_k = \sigma_k(\lambda_1, \dots, \lambda_n)$ . Тогда  $|\lambda I - A| = \sum_{k=0}^n (-1)^k \lambda^{n-k} \sigma_k$ . Поэтому достаточно доказать, что  $f_m(A) = \sum_{k=0}^m (-1)^k A^{m-k} \sigma_k$  для всех  $m$ . При  $m = 1$  это равенство совпадает с определением  $f_1$ . Предположим, что утверждение доказано для  $m$  и докажем его для  $m + 1$ . Ясно, что  $f_{m+1}(A) = \sum_{k=0}^m (-1)^k A^{m-k+1} \sigma_k - \frac{1}{m+1} \text{tr} \left( \sum_{k=0}^m (-1)^k A^{m-k+1} \sigma_k \right) I$ . Так как  $\text{tr} \left( \sum_{k=0}^m (-1)^k A^{m-k+1} \sigma_k \right) = \sum_{k=0}^m (-1)^k s_{m-k+1} \sigma_k$ , где  $s_p = \lambda_1^p + \dots + \lambda_n^p$ , то остаётся заметить, что  $\sum_{k=0}^m (-1)^k s_{m-k+1} \sigma_k + (m+1)(-1)^{m+1} \sigma_{m+1} = 0$  (см. §1.4.1).

**3.2.** Согласно решению задачи 3.1 коэффициенты характеристического многочлена матрицы  $X$  являются функциями от  $\text{tr } X, \dots, \text{tr}(X^n)$ , поэтому характеристические многочлены матриц  $A$  и  $B$  совпадают.

**3.3.** Если  $\lambda_1, \dots, \lambda_k$  — различные собственные значения матрицы  $A$ , то  $p_A(t) = (t - \lambda_1)^{n_1} \dots (t - \lambda_k)^{n_k}$  для некоторых натуральных  $n_1, \dots, n_k$ . Поэтому

$$|p_A(B)| = |B - \lambda_1 I|^{n_1} \dots |B - \lambda_k I|^{n_k} \neq 0.$$

Действительно,  $|B - \lambda_i I| \neq 0$ , поскольку  $\lambda_i$  не является собственным значением матрицы  $B$ .

**3.4.** Пусть  $f(\lambda)$  — произвольный многочлен,  $g(\lambda) = \lambda^n f(\lambda^{-1})$  и  $B = A^{-1}$ . Если  $0 = g(B) = B^n f(A)$ , то  $f(A) = 0$ . Поэтому минимальный многочлен матрицы  $B$  пропорционален  $\lambda^n p(\lambda^{-1})$ . Остаётся заметить, что старший коэффициент многочлена  $\lambda^n p(\lambda^{-1})$  равен  $\lim_{\lambda \rightarrow \infty} (\lambda^n p(\lambda^{-1})) / \lambda^n = p(0)$ .

**3.5.** Ясно, что

$$B_k = \{A \in M_n \mid \text{матрицы } I, A, A^2, \dots, A^{k-1} \text{ линейно зависимы}\}.$$

Матрицу из  $M_n$  можно записать в виде строки длиной  $n^2$ . Составим из таких строк, соответствующих матрицам  $I, A, A^2, \dots, A^{k-1}$ , матрицу  $T$ . Линейная зависимость матриц  $I, A, \dots, A^{k-1}$  эквивалентна тому, что  $\text{rk } T \leq k - 1$ , т.е. все миноры порядка  $k - 1$  матрицы  $T$  равны 0. В итоге получаем систему алгебраических соотношений для коэффициентов матрицы  $A$ .

**4.1.** Минимальный многочлен циклической клетки совпадает (с точностью до знака) с характеристическим многочленом. Минимальный многочлен оператора  $A$  аннулирует данные циклические клетки, поэтому он делится на  $p$  и  $q$ . А так как многочлены  $p$  и  $q$  взаимно просты, то минимальный многочлен  $A$  равен  $pq$ . Следовательно, в пространстве  $V$  существует вектор, минимальный многочлен которого равен  $pq$ .

**4.2.** Докажем сначала, что  $A^k e_n = e_{n-k} + \varepsilon$ , где  $\varepsilon \in \langle e_n, \dots, e_{n-k+1} \rangle$ . При  $k = 1$  получаем  $Ae_n = e_{n-1} + e_n$ , а если утверждение верно для  $k$ , то  $A^{k+1}e_n = e_{n-k-1} + \lambda e_{n-k} + A\varepsilon$  и  $e_{n-k}, A\varepsilon \in \langle e_n, \dots, e_{n-k} \rangle$ .

Поэтому, записав столбцы координат векторов  $e_n, Ae_n, \dots, A^{n-1}e_n$  относительно базиса  $e_n, e_{n-1}, \dots, e_1$  получим верхнюю треугольную матрицу с единицами на диагонали. Эта матрица невырожденная, поэтому векторы  $e_n, Ae_n, \dots, A^{n-1}e_n$  образуют базис.

**Замечание.** Можно доказать, что в качестве  $v$  можно взять любой вектор  $x_1 e_1 + \dots + x_n e_n$ , где  $x_n \neq 0$ .

**4.3.** Пусть

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_n \\ 1 & 0 & \dots & 0 & -a_{n-1} \\ 0 & 1 & \dots & 0 & -a_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -a_1 \end{pmatrix}$$

и

$$S = \begin{pmatrix} a_{n-1} & a_{n-2} & \dots & a_1 & 1 \\ a_{n-2} & a_{n-3} & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Тогда

$$AS = \begin{pmatrix} -a_n & 0 & 0 & \dots & 0 & 0 \\ 0 & a_{n-2} & a_{n-3} & \dots & a_1 & 1 \\ 0 & a_{n-3} & a_{n-4} & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_1 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \end{pmatrix}$$

— симметрическая матрица. Поэтому  $AS = (AS)^T = SA^T$ , т.е.  $A = SA^T S^{-1}$ .

**5.1.** Утверждение очевидно следует из теоремы 3.5.2.1 Приведём, однако, независимое доказательство. Так как  $x^T A x$  — число, то  $x^T A x = (x^T A x)^T = x^T A^T x$ . Поэтому  $x^T A x = x^T A_1 x$ , где  $A_1 = (A + A^T)/2$  — симметрическая матрица с нулевым следом. Соответствующую ей квадратичную форму ортогональным преобразованием можно привести к диагональному виду, поэтому она не может быть знакоопределённой.

**5.2.** Согласно теореме 3.5.3.1 существует такая матрица  $P$ , что все диагональные элементы матрицы  $B = P^{-1}AP$  ненулевые. Рассмотрим матрицу  $Z$ , диагональные элементы которой равны 1, над диагональю стоят нули, а под диагональю — такие же элементы, как у матрицы  $-B$ . Собственные значения нижней треугольной матрицы  $Z$  равны 1, а собственные значения верхней треугольной матрицы  $B + Z$  равны  $1 + b_{ii} \neq 1$ . Поэтому в качестве матрицы  $X$  можно взять  $PZP^{-1}$ .

**6.1.** Оператор  $A$  можно представить в виде  $A = SU$ , где  $U$  — ортогональный оператор,  $S$  — положительно определённый симметрический оператор. Для симметрического оператора существует ортогональный базис собственных векторов, т.е. он является растяжением по перпендикулярным направлениям.

**6.2.** Запишем матрицу  $A$  в виде  $A = SU$ , где матрица  $S$  эрмитова, а матрица  $U$  унитарная. Тогда  $A^*A = U^*S^2U = U^*(AA^*)U$ , поэтому  $f(A^*A) = U^*f(AA^*)U = 0$ .

**6.3.** Если  $A = SU$  — полярное разложение оператора  $A$ , то для  $S$  существует ортонормированный собственный базис  $e_1, \dots, e_n$ , причём все собственные значения не превосходят 1. Следовательно,  $Se_i = (\cos \varphi_i)e_i$ . Дополним базис  $e_1, \dots, e_n$  до базиса  $e_1, \dots, e_n, \varepsilon_1, \dots, \varepsilon_n$  пространства  $\mathbb{R}^{2n}$  и рассмотрим ортогональный оператор  $S_1$ , действующий в каждой плоскости  $\langle e_i, \varepsilon_i \rangle$  как поворот на угол  $\varphi_i$ . Матрица оператора  $S_1$  имеет вид  $\begin{pmatrix} S & * \\ * & * \end{pmatrix}$ . А так как

$$\begin{pmatrix} S & * \\ * & * \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} SU & * \\ * & * \end{pmatrix},$$

то  $S_1 \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix}$  — искомое ортогональное преобразование  $\mathbb{R}^{2n}$ .

**7.1.** Пусть матрица  $X_{pq}(\lambda)$  имеет единственный ненулевой внедиагональный элемент  $a_{pq} = \lambda$ , а на диагонали этой матрицы стоят единицы. Тогда матрица  $X_{pq}(\lambda)A$  получается из матрицы  $A$  прибавлением к  $p$ -й строке  $q$ -й строки, умноженной на  $\lambda$ . По условию  $a_{11} \neq 0$ , поэтому, вычтя из  $k$ -й строки первую, умноженную на  $a_{k1}/a_{11}$ , получим матрицу с  $a_{21} = \dots = a_{n1} = 0$ . Из условия следует, что  $a_{22} \neq 0$ . Поэтому снова можно вычесть из  $k$ -й строки ( $k \geq 3$ ) вторую, умноженную на  $a_{k2}/a_{22}$ , и получим матрицу с  $a_{32} = \dots = a_{n2} = 0$  и т.д. Следовательно, домножая матрицу  $A$  справа на матрицы  $X_{pq}(\lambda)$ , где  $p > q$ , можно получить верхнюю треугольную матрицу  $U$ . Так как  $p > q$ , то все матрицы  $X_{pq}$  нижние треугольные, а значит, их произведение  $T$  тоже является нижней треугольной матрицей. Из равенства  $TA = U$  следует, что  $A = T^{-1}U$ . Остаётся заметить, что матрица  $L = T^{-1}$  нижняя треугольная (см. задачу 2.8); на диагонали матрицы  $L$  стоят единицы.

**7.2.** Пусть  $x_1, \dots, x_n$  — столбцы матрицы  $X$ . Согласно теореме 2.5.2.1 существует такая ортонормированная система векторов  $y_1, \dots, y_n$ , что  $y_i \in \langle x_1, \dots, x_i \rangle$  при  $i = 1, \dots, n$ . Тогда матрица  $U$ , состоящая из столбцов  $y_1, \dots, y_n$ , ортогональна и  $U = XT_1$ , где  $T_1$  — верхняя треугольная матрица. Следовательно,  $X = UT$ , где  $T = T_1^{-1}$  — верхняя треугольная матрица.

**7.3.** Применим индукцию по порядку  $n$  матрицы  $S_n = S$ . При  $n = 1$  утверждение очевидно. Пусть  $S_n = \begin{pmatrix} S_{n-1} & b \\ b^T & a_{nn} \end{pmatrix}$ , где  $S_{n-1}$  — положительно определённая матрица порядка  $n - 1$ . Согласно предположению индукции  $S_{n-1} = L_{n-1}L_{n-1}^T$ . Матрица  $L_{n-1}$  невырожденная, поэтому существует вектор  $c$ , для которого  $L_{n-1}c = b$ . Пусть  $L_n = \begin{pmatrix} L_{n-1} & 0 \\ c^T & x \end{pmatrix}$ . Тогда

$$L_{n-1}L_{n-1}^T = \begin{pmatrix} L_{n-1} & 0 \\ c^T & x \end{pmatrix} \begin{pmatrix} L_{n-1} & c \\ 0 & x \end{pmatrix} = \begin{pmatrix} S_{n-1} & b \\ b^T & c^Tc + x^2 \end{pmatrix}.$$



Поэтому если мы положим  $x = \sqrt{a_{nn} - c^T c}$ , то получим требуемую матрицу  $L = L_n$ .

Отметим, что число  $x$  вещественно, поскольку  $(\det L_{n-1})^2 x^2 = \det S_n > 0$ .

**7.4.** На каждом пересечении фиксированной строки и фиксированного столбца единицу имеет ровно одна из матриц  $I, C, C^2, \dots, C^{n-1}$ , поэтому  $M$  однозначно представима в виде  $M = D_0 + D_1 C + \dots + D_{n-1} C^{n-1}$ , где  $D_l$  — диагональная матрица.

Диагональные матрицы  $I, B, B^2, \dots, B^{n-1}$  линейно независимы, так как из их диагоналей можно составить определитель Вандермонда. Поэтому любая матрица  $D_l$  однозначно представима в виде их линейной комбинации:  $D_l = \sum_{k=0}^{n-1} a_{kl} B^k$ .

**7.5.** Матрицу  $A/2$  можно представить в виде  $A/2 = S_1 S_2$ , где  $S_1$  и  $S_2$  — симметрические матрицы (теорема 3.7.3.1). Поэтому  $A = (A - A^T)/2 = S_1 S_2 - S_2 S_1$ .

**8.1.** Пусть  $A$  — жорданова или циклическая клетка порядка  $n$ . В обоих случаях матрица  $A - xI$  имеет треугольную подматрицу порядка  $n - 1$  с единицами на главной диагонали. Поэтому  $f_1 = \dots = f_{n-1} = 1$  и  $f_n = p_A(x)$  — характеристический многочлен матрицы  $A$ . Следовательно,  $g_1 = \dots = g_{n-1} = 1$  и  $g_n = p_A(x)$ .

**8.2.** Циклическая нормальная форма матрицы  $A$  имеет блочно диагональный вид с циклическими клетками, соответствующими многочленам  $p_1, p_2, \dots, p_k$ , где  $p_1$  — минимальный многочлен матрицы  $A$  и  $p_i$  делится на  $p_{i+1}$ . Инвариантными множителями этих циклических клеток являются  $p_1, \dots, p_k$  (задача 8.1), поэтому их нормальные формы Смита имеют вид  $\text{diag}(1, \dots, 1, p_i)$ . Следовательно, нормальная форма Смита матрицы  $A$  имеет вид  $\text{diag}(1, \dots, 1, p_k, \dots, p_2, p_1)$ . Поэтому  $f_{n-1} = p_2 p_3 \dots p_k$ .

## Глава 4.

# Матрицы специального вида

### 4.1. Симметрические и эрмитовы матрицы

Напомним сначала доказанные ранее свойства вещественных симметрических и комплексных эрмитовых матриц.

Собственные значения вещественной симметрической (комплексной эрмитовой) матрицы вещественны, причём собственные векторы, соответствующие разным собственным значениям, ортогональны (см. §3.1.3).

Симметрическая (эрмитова) матрица  $S$  ортогонально (унитарно) диагонализируема, т.е. существует такая ортогональная (унитарная) матрица  $U$ , что матрица  $U^{-1}SU$  диагональна (см. §3.1.3). Матрица  $A$  эрмитова тогда и только тогда, когда  $(Ax, x) \in \mathbb{R}$  для любого вектора  $x$  (см. §3.1.3 и теорему 2.7.4.2).

#### 4.1.1. Квадратичные и эрмитовы формы

Между квадратичными формами и симметрическими билинейными функциями существует взаимно однозначное соответствие: *квадратичной форме*  $q(x) = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j$  соответствует *симметрическая билинейная функция*  $B(x, y) = \sum_{i=1}^n a_{ii}x_iy_i + \sum_{i<j} a_{ij}(x_iy_j + x_jy_i) = \sum_{i,j=1}^n a_{ij}x_iy_j$ ,

где  $a_{ij} = a_{ji}$ . В инвариантном виде это соответствие можно записать так:  $q(x) = B(x, x)$  и  $B(x, y) = (q(x+y) - q(x) - q(y))/2$ .

Симметрическую матрицу  $A = \|a_{ij}\|_1^n$  называют *матрицей квадратич-*

ной формы  $q$  или матрицей симметрической билинейной функции  $B$  (относительно некоторого базиса). Если  $x$  и  $y$  — столбцы координат, то  $q(x) = x^T A x = (x, Ax)$  и  $B(x, y) = x^T A y = (x, Ay)$ .

В вещественном случае квадратичную форму  $q$  называют *положительно определённой*, если  $q(x) > 0$  для любого  $x \neq 0$ . Симметрическую матрицу  $A$ , соответствующую квадратичной форме  $q$ , называют в этом случае *положительно определённой*. Аналогично вводится понятие *отрицательно определённой* симметрической матрицы и квадратичной формы ( $q(x) < 0$ ); *неотрицательно определённой* ( $q(x) \geq 0$ ) и *знакоопределённой*, т.е. положительно или отрицательно определённой. Положительная определённость матрицы  $A$  обозначается так:  $A > 0$ ; используются также аналогичные обозначения  $A \geq 0$ ,  $A < 0$  и  $A \leq 0$ .

**Теорема 4.1.1.1.** Пусть  $X$  — невырожденная матрица. Тогда матрица  $A = X^T X$  положительно определённая.

**Доказательство.** Матрице  $A$  соответствует квадратичная форма  $q(x) = x^T A x = x^T X^T X x = (Xx, Xx)$ . Эта форма положительно определённая, поскольку  $Xx \neq 0$  при  $x \neq 0$ .  $\square$

**Теорема 4.1.1.2.** Матрица Грама системы линейно независимых векторов положительно определённая.

**Доказательство.** Набору линейно независимых векторов  $e_1, \dots, e_n$  можно сопоставить положительно определённую квадратичную форму  $q(x_1, \dots, x_n) = \|x_1 e_1 + \dots + x_n e_n\|^2$ . Легко проверить, что этой квадратичной форме соответствует матрица  $\begin{pmatrix} g_{11} & \dots & g_{1k} \\ \dots & \dots & \dots \\ g_{k1} & \dots & g_{kk} \end{pmatrix}$ , где  $g_{ij} = (e_i, e_j)$ .

Но это и есть матрица Грама.  $\square$

**Замечание.** Ещё одно доказательство того, что матрица Грама положительно определённая, можно получить с помощью теоремы 4.1.1.1. Действительно, матрицу Грама можно представить в виде  $X^T X$ .

В комплексном случае квадратичная форма не может быть знакоопределённой. Понятие знакоопределённых симметрических билинейных функций можно ввести для эрмитовых форм  $B(x, y) = \sum_{i,j=1}^n a_{ij} x_j \bar{y}_i$  (где  $A = \|a_{ij}\|_1^n$  — эрмитова матрица). Если  $x$  и  $y$  — столбцы координат, то  $B(x, y) = y^* A x = \overline{x^* A y}$ , т.е.  $B(x, y) = (Ax, y) = \overline{(Ay, x)}$ , где  $(Ax, y)$  —

стандартное эрмитово произведение. Ясно, что  $B(y, x) = \overline{B(x, y)}$ . Эрмитову форму называют *положительно определённой*, если  $B(x, x) > 0$  при  $x \neq 0$ .

Для вещественной симметрической или эрмитовой формы существует ортонормированный базис, в котором матрица  $A$  диагональна, т.е.  $B(x, y) = \sum_{i=1}^n \lambda_i x_i \bar{y}_i$ , причём числа  $\lambda_i$  вещественны (см. §3.1.3). Эрмитова форма  $B$  положительно определена тогда и только тогда, когда все числа  $\lambda_i$  положительны. Таким образом, матрица  $A$  положительно определённая тогда и только тогда, когда  $A = UDU^{-1}$ , где  $D$  — диагональная матрица с положительными элементами на диагонали, а  $U$  — ортогональная (эрмитова) матрица.

Пусть  $e = (e_1, \dots, e_n)$  и  $e' = (e'_1, \dots, e'_n)$  — два базиса и  $e' = eP$ , где  $P$  — матрица перехода от  $e$  к  $e'$ . Если  $x$  и  $x'$  — столбцы координат одного вектора относительно базисов  $e$  и  $e'$ , то  $ex = e'x' = ePx'$ , т.е.  $x = Px'$ . Поэтому  $y^*Ax = (y')^*P^*APx'$ , т.е. при переходе от базиса  $e$  к  $e'$  матрица  $A$  эрмитовой формы  $B(x, y) = y^*Ax$  заменяется на  $P^*AP$ .

**Замечание.** Матрицы эрмитовой формы и квадратичной формы при замене базиса в комплексном случае преобразуются по-разному. В первом случае эрмитова матрица  $A$  заменяется на  $P^*AP$ , а во втором случае комплексная симметрическая матрица  $A$  заменяется на  $P^TAP$ .

#### 4.1.2. Сигнатура квадратичной формы

Пусть  $e_1, \dots, e_n$  — произвольный базис,  $B(x, y)$  — эрмитова форма,  $\varepsilon_1, \dots, \varepsilon_n$  — ортогональный базис, относительно которого матрица эрмитовой формы имеет вид  $\text{diag}(a_1, \dots, a_n)$ , где  $a_i = B(\varepsilon_i, \varepsilon_i)$ .

Рассмотрим эрмитову матрицу  $G = \|B(e_i, e_j)\|_1^n$ ; пусть  $G_i$  — угловой минор порядка  $i$  матрицы  $G$ .

**Теорема 4.1.2.1 (критерий Сильвестра).** Матрица  $G$  положительно определена тогда и только тогда, когда  $|G_i| > 0$  при  $i = 1, \dots, n$ .

**Доказательство.** Пусть  $A_i$  — матрица перехода от базиса  $e_1, \dots, e_i$  к базису  $\varepsilon_1, \dots, \varepsilon_i$ . Тогда  $\|B(\varepsilon_k, \varepsilon_j)\|_1^i = A_i^* G_i A_i$ . Поэтому

$$a_1 \dots a_i = |G_i| \cdot |\det A_i|^2. \quad (4.1)$$

Следовательно, все элементы  $a_1, \dots, a_n$  положительны тогда и только тогда, когда  $|G_i| > 0$  при  $i = 1, \dots, n$ .  $\square$

**Теорема 4.1.2.2 (Якоби).** Пусть  $|G_i| \neq 0$  при  $i = 1, \dots, n-1$ . Тогда эрмитову форму  $B(x, y)$  можно привести к виду  $\sum_{i=1}^n a_i x_i \bar{y}_i$ , где  $a_i = |G_i|/|G_{i-1}|$ ,  $|G_0| = 1$ .

**Доказательство.** В базисе  $\varepsilon_i, \dots, \varepsilon_n$  эрмитова форма  $B$  имеет вид  $\sum_{i=1}^n a_i x_i \bar{y}_i$  где  $a_1 \dots a_i = |G_i| |\det A_i|^2$  согласно (4.1). Замена  $z_i \mapsto z_i \det A_i$  приводит форму  $B$  к требуемому виду.  $\square$

Напомним, что для любой эрмитовой матрицы  $A$  существует такая унитарная матрица  $U$ , что  $U^* A U = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Так как  $AU = U\Lambda$ , столбцы матрицы  $U$  являются собственными векторами матрицы  $A$ .

Эрмитова форма унитарным преобразованием приводится к виду  $\lambda_1 x_1 \bar{y}_1 + \dots + \lambda_n x_n \bar{y}_n$ , где  $\lambda_i \in \mathbb{R}$ ; квадратичная форма над  $\mathbb{R}$  ортогональным преобразованием приводится к виду  $\lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ . Дальнейшей заменой  $x'_i = \sqrt{|\lambda_i|} x_i$  эрмитову форму можно привести к виду

$$x_1 \bar{y}_1 + \dots + x_p \bar{y}_p - x_{p+1} \bar{y}_{p+1} - \dots - x_{p+q} \bar{y}_{p+q}, \quad (4.2)$$

а квадратичную форму можно привести к виду

$$x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2. \quad (4.3)$$

Пару чисел обычно  $(p, q)$  называют *сигнатурой* квадратичной формы, но иногда то же самое название используют для разности  $p - q$ .

**Теорема 4.1.2.3 (закон инерции квадратичных форм).** Числа  $p$  и  $q$  в (4.2) и (4.3) для каждой формы определены однозначно.

**Доказательство.** Выражения (4.2) и (4.3) задают разложения пространства  $V$  в прямую сумму подпространств  $V = V_+ \oplus V_- \oplus V_0$ , причём на подпространствах  $V_+$ ,  $V_-$ ,  $V_0$  форма соответственно положительно определена, отрицательно определена, тождественно равна нулю. Пусть  $V = W_+ \oplus W_- \oplus W_0$  — другое такое разложение. Тогда  $V_+ \cap (W_- \oplus W_0) = 0$ , поэтому  $\dim V_+ + \dim(W_- \oplus W_0) \leq n$ , т.е.  $\dim V_+ \leq \dim W_+$ . Аналогично  $\dim W_+ \leq \dim V_+$ .  $\square$

**Теорема 4.1.2.4.** Пусть невырожденная квадратичная форма  $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2$  тождественно обращается в нуль на линейном подпространстве размерности  $m$ . Тогда  $m \leq \min(p, q)$ .

**Доказательство.** Предположим, что данная квадратичная форма обращается в нуль на подпространстве размерности  $p + 1$ . Пусть  $e_1, \dots, e_{p+1}$  — базис этого подпространства. Согласно задаче 3.1 среди линейных комбинаций этих векторов есть ненулевой вектор  $(0, \dots, 0, x_{p+1}, \dots, x_{p+q})$ . Приходим к противоречию, поскольку значение квадратичной формы на этом векторе равно  $-x_{p+1}^2 - \dots - x_{p+q}^2 < 0$ . Для подпространства размерности  $q + 1$  рассуждения аналогичны: нужно рассмотреть последние  $q$  координат.  $\square$

#### 4.1.3. Приведение к диагональному виду

Приведение к диагональному виду эрмитовых форм унитарными преобразованиями, которое мы обсудили выше, алгоритмически реализуется весьма сложно. Но если приводить к диагональному виду квадратичную форму произвольными преобразованиями, то алгоритм получается весьма простой.

**Теорема 4.1.3.1 (Лагранж).** *Квадратичную форму можно привести к виду  $q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ .*

**Доказательство.** Пусть  $A = \|a_{ij}\|_1^n$  — матрица квадратичной формы  $q$ . Доказательство проведём индукцией по  $n$ . При  $n = 1$  утверждение очевидно. Рассмотрим два случая.

а) Существует ненулевой диагональный элемент, например,  $a_{11} \neq 0$ . Тогда  $q(x_1, \dots, x_n) = a_{11}y_1^2 + q'(y_2, \dots, y_n)$ , где  $y_1 = x_1 + (a_{12}x_2 + \dots + a_{1n}x_n)/a_{11}$  и  $y_i = x_i$  при  $i \geq 2$ . К форме  $q'$  применимо предположение индукции.

б) Все диагональные элементы равны нулю. Интересен лишь случай, когда есть хотя бы один ненулевой элемент, например  $a_{12} \neq 0$ . Пусть  $x_1 = y_1 + y_2$ ,  $x_2 = y_1 - y_2$  и  $x_i = y_i$  при  $i \geq 3$ . Тогда  $q(x_1, \dots, x_n) = 2a_{12}(y_1^2 - y_2^2) + q'(y_1, \dots, y_n)$ , причём  $q'$  не содержит членов с  $y_1^2$  и  $y_2^2$ . К полученной форме от переменных  $y_1, \dots, y_n$  можно применить замену а).  $\square$

**Замечание.** В комплексном случае дальнейшей заменой  $y_i = \sqrt{\lambda_i}x_i$  квадратичную форму  $q$  можно привести к виду  $y_1^2 + \dots + y_p^2$ , где  $p = \text{rk } A$

#### 4.1.4. Квадратный корень из неотрицательно определённой матрицы

**Теорема 4.1.4.1.** *Если эрмитова матрица  $A$  неотрицательно определена, то  $A = S^2$ , где  $S$  — некоторая неотрицательно определённая эрмитова матрица, причём такая матрица  $S$  единственна.*

**Доказательство.** Существует такая унитарная матрица  $U$ , что  $A = U^{-1}DU$ , где  $D = \text{diag}(d_1, \dots, d_n)$ , причём  $d_i \geq 0$ , т.е.  $d_i = \lambda_i^2$ ; можно считать, что  $\lambda_i \geq 0$ . Пусть  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  и  $S = U^{-1}\Lambda U$ . Тогда  $S^2 = U^{-1}\Lambda^2 U = A$ .

Докажем теперь, что если матрица  $S$  неотрицательно определена и  $S^2 = A$ , то матрица  $S$  определена однозначно. Пусть  $e_1, \dots, e_n$  — собственный ортонормированный базис для  $S$ . Тогда  $Ae_i = S^2e_i = \lambda_i^2 e_i$ . Подпространство, порождённое собственными векторами  $S$ , соответствующими собственному значению  $\lambda$ , совпадает с подпространством, порождённым собственными векторами  $A$ , соответствующими собственному значению  $\lambda^2$ . Следовательно, это подпространство однозначно определено оператором  $A$  и действие оператора  $S$  на нём тоже определено однозначно.  $\square$

**Теорема 4.1.4.2.** Если эрмитова матрица  $A$  неотрицательно определена и  $(Ax, x) = 0$ , то  $Ax = 0$ .

**Доказательство.** Согласно теореме 4.1.4.1  $A = S^*S$ . Поэтому  $0 = (Ax, x) = (S^*Sx, x) = (Sx, Sx)$ . Следовательно,  $Sx = 0$  и  $Ax = S^*Sx = 0$ .  $\square$

#### 4.1.5. Теорема Куранта–Фишера

Пусть собственные значения эрмитовой матрицы  $A$  расположены в порядке убывания:  $\lambda_1 \geq \dots \geq \lambda_n$ . Числа  $\lambda_1, \dots, \lambda_n$  обладают следующим минимаксным свойством.

**Теорема 4.1.5.1 (Курант–Фишер).** Пусть  $x$  пробегает все (допустимые) единичные векторы. Тогда

$$\lambda_1 = \max_x (x, Ax), \quad \lambda_2 = \min_{y_1} \max_{x \perp y_1} (x, Ax), \dots$$

$$\dots, \lambda_n = \min_{y_1, \dots, y_{n-1}} \max_{x \perp y_1, \dots, x \perp y_{n-1}} (x, Ax).$$

**Доказательство.** Выберем ортонормированный базис, в котором форма  $(x, Ax)$  диагональна, т.е.  $(x, Ax) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ . Рассмотрим подпространства  $W_1 = \{x \mid x_{k+1} = \dots = x_n = 0\}$  и  $W_2 = \{x \mid x \perp y_1, \dots, x \perp y_{k-1}\}$ . Так как  $\dim W_1 = k$  и  $\dim W_2 \geq n - k + 1$ , то  $W = W_1 \cap W_2 \neq \emptyset$ . Если  $x \in W$  и  $\|x\| = 1$ , то  $x \in W_1$  и  $(Ax, x) = \lambda_1 x_1^2 + \dots + \lambda_k x_k^2 \geq \lambda_k (x_1^2 + \dots + x_k^2) = \lambda_k$ . Поэтому  $\lambda_k \leq \max_{x \in W_1 \cap W_2} (x, Ax) \leq \max_{x \in W_2} (x, Ax)$ , а значит,  $\lambda_k \leq \min_{y_1, \dots, y_{k-1}} \max_{x \in W_2} (x, Ax)$ .

Рассмотрим теперь векторы  $y_i = (0, \dots, 0, 1, 0, \dots, 0)$  (единица стоит на  $i$ -м месте). Тогда  $W_2 = \{x \mid x \perp y_1, \dots, x \perp y_{k-1}\} = \{x \mid x_1 = \dots = x_{k-1} = 0\}$ . Если  $x \in W_2$  и  $\|x\| = 1$ , то  $(x, Ax) = \lambda_k x_k^2 + \dots + \lambda_n x_n^2 \leq \lambda_k (x_k^2 + \dots + x_n^2) = \lambda_k$ . Поэтому

$$\lambda_k \geq \max_{x \in W_2} (x, Ax) \geq \min_{y_1, \dots, y_{k-1}} \max_{x \perp y_1, \dots, x \perp y_{k-1}} (x, Ax).$$

□

#### 4.1.6. Собственные значения произведения двух эрмитовых матриц

Произведение двух эрмитовых матриц  $A$  и  $B$  является эрмитовой матрицей тогда и только тогда, когда  $AB = (AB)^* = B^* A^* = BA$ . Тем не менее, произведение двух положительно определённых матриц до некоторой степени схоже с положительно определённой матрицей: оно является диагонализуемой матрицей с положительными собственными значениями.

**Теорема 4.1.6.1.** Пусть  $A$  — положительно определённая матрица,  $B$  — эрмитова матрица. Тогда  $AB$  — диагонализуемая матрица, причём у неё столько же положительных, отрицательных и нулевых собственных значений, сколько и у матрицы  $B$ .

**Доказательство.** Пусть  $A = S^2$ , где  $S$  — эрмитова матрица. Тогда матрица  $AB$  подобна матрице  $S^{-1}ABS = SBS$ . Если  $S$  — обратимая эрмитова матрица, то матрицы  $B$  и  $SBS$  соответствуют одной и той же эрмитовой форме, записанной в разных базисах. Но размерность максимального подпространства, на котором эрмитова форма положительна (отрицательна, тождественно равна нулю), зависит только от самой формы. Поэтому матрица  $AB$  подобна эрмитовой матрице  $SBS$ , у которой столько же положительных, отрицательных и нулевых собственных значений, сколько и у матрицы  $B$ . □

Верна также теорема, которая до некоторой степени обратна теореме 4.1.6.1.

**Теорема 4.1.6.2.** Любую диагонализируемую матрицу с вещественными собственными значениями можно представить в виде произведения положительно определённой матрицы и эрмитовой матрицы.

**Доказательство.** Пусть  $C = PDP^{-1}$ , где  $D$  — вещественная диагональная матрица. Тогда  $C = AB$ , где  $A = PP^*$  — положительно определённая матрица, а  $B = P^{*-1}DP^{-1}$  — эрмитова матрица. □



**Задачи**

**1.1.** Пусть  $S$  — комплексная симметрическая матрица. Докажите, что матрица  $S\bar{S}$  эрмитова.

**1.2.** Пусть  $S$  — вещественная неотрицательно определённая матрица ранга  $k$ . Докажите, что хотя бы один главный минор порядка  $k$  матрицы  $S$  положителен.

**1.3.** Докажите, что вещественная симметрическая матрица ранга  $r$  является суммой  $r$  симметрических матриц ранга 1.

**1.4.** а) Докажите, что если  $A > 0$ , то  $A^{-1} > 0$ .

б) Докажите, что если  $A > 0$ , то  $\text{adj } A > 0$ .

**1.5.** Докажите, что квадратичная форма, заданная матрицей  $\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$ , где  $A$  — невырожденная матрица, имеет нулевую сигнатуру.

**1.6.** Пусть  $A$  — эрмитова матрица порядка  $n$ ,  $v_1, \dots, v_n$  — ортонормированный базис собственных векторов, соответствующих собственным значениям  $\lambda_1, \dots, \lambda_n$ . Докажите, что  $A = \sum_{k=1}^n \lambda_k v_k v_k^*$ .

**1.7.** Докажите, что если  $-1 < \lambda < 1$ , то матрица  $A = \|a_{ij}\|_1^n$ , где  $a_{ii} = 1$  и  $a_{ij} = \lambda$  при  $i \neq j$ , положительно определена.

**1.8.** Докажите, что если  $A \geq 0$ , то  $\begin{pmatrix} A & A \\ A & A \end{pmatrix} \geq 0$ .

**1.9.** Докажите, что матрица  $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \dots & \frac{1}{2n-1} \end{pmatrix}$  положительно определённая.

**1.10.** Матрица  $A$  положительно определена. Докажите, что

$$\int_{-\infty}^{+\infty} e^{-(x, Ax)} dx = (\sqrt{\pi})^n |A|^{-1/2}.$$

**1.11.** Матрицы  $A$  и  $B$  вещественные симметрические, причём матрица  $\Lambda = A^{-1}B$  диагональна. Пусть  $\Lambda e_i = \lambda_i e_i$ . Докажите, что если  $\lambda_i \neq \lambda_j$ , то  $(Ae_i, e_j) = (Be_i, e_j) = 0$ .

**1.12.** Пусть  $S$  — симметрическая невырожденная матрица порядка  $n$ , все элементы которой положительны. Каково наибольшее возможное число нулевых элементов матрицы  $S^{-1}$ ?

**1.13.** Пусть  $A \geq 0$  и  $B \geq 0$ . Докажите, что  $\operatorname{tr}(AB) \geq 0$ .

**1.14.** Пусть  $A \geq 0$ ,  $B \geq 0$  и  $A \geq B$ . Докажите, что  $\operatorname{tr}(A^2) \geq \operatorname{tr}(B^2)$ .

**1.15.** Докажите, что для любого положительного  $\lambda$  эрмитовы формы, заданные матрицами  $\begin{pmatrix} A & B \\ B^* & 0 \end{pmatrix}$  и  $\begin{pmatrix} \lambda A & B \\ B^* & 0 \end{pmatrix}$ , имеют одинаковую сигнатуру.

**1.16.** а) Докажите, что квадратичные формы с матрицами  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  и  $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$  эквивалентны над  $\mathbb{Q}$ , но не эквивалентны над  $\mathbb{Z}$ .

б) Докажите, что квадратичные формы с матрицами  $\begin{pmatrix} 0 & 1 \\ 1 & m \end{pmatrix}$  и  $\begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix}$ , где  $m$  и  $n$  — целые числа, эквивалентны над  $\mathbb{Z}$  тогда и только тогда, когда числа  $m$  и  $n$  одной чётности.

## 4.2. Одновременное приведение пары эрмитовых форм к диагональному виду

### 4.2.1. Случай положительно определённой матрицы

**Теорема 4.2.1.1.** Пусть  $A$  и  $B$  — эрмитовы матрицы, причём  $A > 0$ . Тогда существует такая матрица  $T$ , что  $T^*AT = I$ , а матрица  $T^*BT$  диагональна.

**Доказательство.** Для матрицы  $A$  существует такая матрица  $Y$ , что  $Y^*AY = I$  (см. §4.1.2). Матрица  $C = Y^*BY$  эрмитова, поэтому существует такая унитарная матрица  $U$ , что матрица  $U^*CU$  диагональна. А так как  $U^*IU = I$ , то  $T = YU$  — искомая матрица.  $\square$

Две эрмитовы формы не всегда можно одновременно привести к диагональному виду заменой базиса. Рассмотрим, например, эрмитовы формы, соответствующие матрицам  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  и  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Пусть  $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  — произвольная невырожденная матрица. Тогда

$$P^* \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} P = \begin{pmatrix} a\bar{a} & \bar{a}b \\ a\bar{b} & b\bar{b} \end{pmatrix}$$

и

$$P^* \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} P = \begin{pmatrix} a\bar{c} + \bar{a}c & \bar{a}d + b\bar{c} \\ a\bar{d} + \bar{b}c & b\bar{d} + \bar{b}d \end{pmatrix}.$$

Остаётся проверить, что равенства  $\bar{a}b = 0$  и  $\bar{a}d + b\bar{c} = 0$  не могут выполняться одновременно. Если  $\bar{a}b = 0$  и матрица  $P$  невырождена, то либо  $a = 0$  и  $b \neq 0$ , либо  $b = 0$  и  $a \neq 0$ . В первом случае  $0 = \bar{a}d + b\bar{c} = b\bar{c}$ , поэтому  $c = 0$ , а во втором случае  $\bar{a}d = 0$ , поэтому  $d = 0$ . В обоих случаях получается вырожденная матрица  $P$ .

#### 4.2.2. Случай знакоопределённой матрицы

Разумеется, пару эрмитовых форм можно привести одновременно к диагональному виду не только в том случае, когда одна из них положительно (или отрицательно) определена. Известны ещё и другие условия, при которых пару эрмитовых форм можно одновременно привести к диагональному виду.

**Теорема 4.2.2.1.** Пусть  $A$  и  $B$  — эрмитовы матрицы, причём для некоторых вещественных чисел  $\alpha$  и  $\beta$  матрица  $\alpha A + \beta B$  положительно определённая. Тогда существует невырожденная матрица  $T$ , для которой матрицы  $T^*AT$  и  $T^*BT$  диагональные.

**Доказательство.** [Au1] Ясно, что одно из чисел  $\alpha$  и  $\beta$  отлично от нуля. Пусть для определённости  $\alpha \neq 0$ . Согласно теореме 3.1.3.2 существует унитарная матрица  $U$ , для которой  $U^*(\alpha A + \beta B)U = \text{diag}(d_1, \dots, d_n)$ , где числа  $d_1, \dots, d_n$  положительны. Пусть  $S = \text{diag}(s_1, \dots, s_n)$ , где  $s_k = 1/\sqrt{d_k}$ . Тогда  $S^*U^*(\alpha A + \beta B)US = I$ .

Матрица  $S^*U^*BUS$  эрмитова, поэтому существует унитарная матрица  $W$ , для которой матрица  $T^*BT$ , где  $T = USW$ , диагональна. Покажем, что матрица  $T^*AT$  тоже диагональная. Ясно, что  $T^*(\alpha A + \beta B)T = W^*W = I$ , поскольку матрица  $W$  унитарная. Поэтому  $T^*AT = \alpha^{-1}(I - \beta T^*BT)$  — диагональная матрица.  $\square$

**Теорема 4.2.2.2.** Если эрмитовы матрицы  $A$  и  $B$  знакоопределены (т.е. неположительно или неотрицательно определены), то существует такая невырожденная матрица  $T$ , что матрицы  $T^*AT$  и  $T^*BT$  диагональны.

**Доказательство.** [Ne1] Пусть  $\text{rk } A = a$ ,  $\text{rk } B = b$  и  $a \leq b$ . Существует такая невырожденная матрица  $T_1$ , что  $T_1^*AT_1 = \begin{pmatrix} I_a & 0 \\ 0 & 0 \end{pmatrix} = A_0$ . Рассмотрим последние  $n - a$  диагональных элементов матрицы  $B_1 = T_1^*BT_1$ . Матрица  $B_1$  знакоопределённая, поэтому если её диагональный элемент равен нулю, то строка и столбец, в которых он стоит, нулевые (см. задачу 2.1). Пусть теперь какой-либо из рассматриваемых диагональных элементов не равен нулю. Легко проверить, что

$$\begin{pmatrix} I & x^* \\ 0 & \bar{\alpha} \end{pmatrix} \begin{pmatrix} C & c^* \\ c & \gamma \end{pmatrix} \begin{pmatrix} I & 0 \\ x & \alpha \end{pmatrix} = \begin{pmatrix} * & \alpha c^* + \alpha \gamma x^* \\ \bar{\alpha} c + \bar{\alpha} \gamma x & |\alpha|^2 \gamma \end{pmatrix}.$$

Если  $\gamma \neq 0$ , то, положив  $\alpha = 1/\sqrt{\gamma}$  и  $x = -(1/\gamma)c$ , получим матрицу, все внедиагональные элементы последней строки и последнего столбца которой равны нулю. Эти преобразования сохраняют матрицу  $A_0$ ; докажем теперь, что с их помощью матрицу  $B_1$  можно привести к виду

$$B_0 = \begin{pmatrix} B_a & 0 & 0 \\ 0 & I_k & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

где  $B_a$  — матрица порядка  $a$  и  $k = b - \text{rk } B_a$ . Для этого перестановочной матрицей, затрагивающей лишь последние  $n - a$  элементов матрицы  $B_1$ , поставим первыми ненулевые диагональные элементы, а последними нулевые. Затем с помощью указанных выше преобразований, начиная с последнего ненулевого элемента и постепенно сокращая размеры рассматриваемой матрицы, доходим до матрицы порядка  $a$ .

Пусть  $T_2$  — такая невырожденная матрица, что  $T_2^* B T_2 = B_0$  и  $T_2^* A T_2 = A_0$ . Существует унитарная матрица  $U$  порядка  $a$ , для которой матрица  $U^* B_a U$  диагональна. А так как  $U^* I_a U = I_a$ , то  $T = T_2 U_1$ , где  $U_1 = \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix}$  — искомая матрица.  $\square$

### 4.2.3. Ещё один случай

**Теорема 4.2.3.1.** Пусть  $A$  и  $B$  — эрмитовы матрицы, причём не существует такого ненулевого столбца  $x$ , что  $x^* A x = x^* B x = 0$ . Тогда существует такая невырожденная матрица  $T$ , что  $T^* A T$  и  $T^* B T$  — диагональные матрицы.

Так как треугольная эрмитова матрица диагональна, теорема 4.2.3.1 является частным случаем следующего утверждения.

**Теорема 4.2.3.2.** Пусть  $A$  и  $B$  — произвольные комплексные квадратные матрицы, причём не существует такого столбца  $x$ , что  $x^* A x = x^* B x = 0$ . Тогда существует такая невырожденная матрица  $T$ , что  $T^* A T$  и  $T^* B T$  — треугольные матрицы.

**Доказательство.** [Ma1] Если одна из матриц  $A$  и  $B$ , например  $B$ , невырождена, то  $p(\lambda) = |A - \lambda B| = \lambda^n |A - B| + \dots$  — многочлен ненулевой степени. Если обе матрицы  $A$  и  $B$  вырождены, то  $|A - \lambda B| = 0$  при  $\lambda = 0$ . В обоих случаях уравнение  $|A - \lambda B| = 0$  имеет корень  $\lambda$ , поэтому существует такой столбец  $x_1$ , что  $Ax_1 = \lambda Bx_1$ . Если  $\lambda \neq 0$  (соответственно  $\lambda = 0$ ), выберем такие линейно независимые столбцы  $x_2, \dots, x_n$ , что  $x_i^* A x_i = 0$  (соответственно  $x_i^* B x_i = 0$ ) при  $i = 2, \dots, n$ ; в обоих случаях

$x_i^* A x_1 = x_i^* B x_1 = 0$  при  $i = 2, \dots, n$ . Поэтому если матрица  $D$  составлена из столбцов  $x_1, \dots, x_n$ , то

$$D^* A D = \begin{pmatrix} x_1^* A x_1 & \dots & x_1^* A x_n \\ 0 & & \\ \vdots & A_1 & \\ 0 & & \end{pmatrix} \quad \text{и} \quad D^* B D = \begin{pmatrix} x_1^* B x_1 & \dots & x_1^* B x_n \\ 0 & & \\ \vdots & B_1 & \\ 0 & & \end{pmatrix}.$$

Докажем, что матрица  $D$  невырождена, т.е. столбец  $x_1$  не выражается линейно через  $x_2, \dots, x_n$ . Предположим, что  $x_1 = \lambda_2 x_2 + \dots + \lambda_n x_n$ . Тогда  $x_1^* A x_1 = (\bar{\lambda}_2 x_2^* + \dots + \bar{\lambda}_n x_n^*) A x_1 = 0$ . Аналогично  $x_1^* B x_1 = 0$ . Получено противоречие, поэтому матрица  $D$  невырождена.

Докажем теперь, что матрицы  $A_1$  и  $B_1$  удовлетворяют условию теоремы. Предположим, что существует такой ненулевой столбец  $y_1 = (\alpha_2, \dots, \alpha_n)$ , что  $y_1^* A_1 y_1 = y_1^* B_1 y_1 = 0$ . Легко проверить, что  $A_1 = D_1^* A D_1$ , где  $D_1$  — матрица, составленная из столбцов  $x_2, \dots, x_n$ . Поэтому  $y^* A y = y^* B y = 0$ , где  $y = D_1 y_1 = \alpha_2 x_2 + \dots + \alpha_n x_n \neq 0$ , так как столбцы  $x_2, \dots, x_n$  линейно независимы. Получено противоречие.

Если для матриц  $A_1$  и  $B_1$  существует такая невырожденная матрица  $T_1$ , что матрицы  $T_1^* A_1 T_1$  и  $T_1^* B_1 T_1$  верхние треугольные, то матрица  $T = D \begin{pmatrix} 1 & 0 \\ 0 & T_1 \end{pmatrix}$  искомая. Для матриц порядка 1 утверждение очевидно, поэтому можно воспользоваться индукцией по порядку матрицы.  $\square$

### Задачи

**2.1.** Эрмитова матрица  $A = \|a_{ij}\|_1^n$  неотрицательно определена и  $a_{ii} = 0$  для некоторого  $i$ . Докажите, что  $a_{ij} = a_{ji} = 0$  для всех  $j$ .

**2.2.** [Vä2] Пусть  $A$  и  $B$  — положительно определённые эрмитовы матрицы. Запишем матрицу  $\text{adj}(tA + B)$  в виде матричного полинома  $t^{n-1}A_0 + t^{n-2}A_1 + \dots + A_{n-1}$ . Докажите, что матрицы  $A_0, \dots, A_{n-1}$  положительно определённые.

**2.3.** [A1] Симметрические матрицы  $A_i$  и  $B_i$  ( $i = 1, 2$ ) таковы, что характеристические многочлены матриц  $x A_1 + y A_2$  и  $x B_1 + y B_2$  равны при всех  $x$  и  $y$ . Обязательно ли найдётся такая ортогональная матрица  $U$ , что  $U A_i U^T = B_i$  при  $i = 1, 2$ ?

### 4.3. Кососимметрические матрицы

Матрицу  $A$  называют *кососимметрической*, если  $A^T = -A$ . В этом параграфе рассматриваются вещественные кососимметрические матрицы.

Напомним, что определитель кососимметрической матрицы нечётного порядка равен нулю (см. задачу 1.1).

### 4.3.1. Простейшие свойства

**Теорема 4.3.1.1.** Если  $A$  — кососимметрическая матрица, то  $A^2$  — симметрическая неположительно определённая матрица.

**Доказательство.** Ясно, что  $(A^2)^T = (A^T)^2 = (-A)^2 = A^2$  и  $(A^2x, x) = (Ax, A^Tx) = -(A, Ax) \leq 0$ .  $\square$

**Следствие.** Ненулевые собственные значения кососимметрической матрицы являются чисто мнимыми.

**Доказательство.** Если  $Ax = \lambda x$ , то  $A^2x = \lambda^2x$ , причём  $\lambda^2 \leq 0$ .  $\square$

**Теорема 4.3.1.2.** Равенство  $(x, Ax) = 0$  выполняется для всех  $x$  тогда и только тогда, когда матрица  $A$  кососимметрическая.

**Доказательство.** В ортонормированном базисе  $(x, Ax) = \sum_{i,j} a_{ij}x_ix_j = \sum_{i \leq j} (a_{ij} + a_{ji})x_ix_j$ . Квадратичная форма равна нулю при всех  $x$  тогда и только тогда, когда её коэффициенты равны нулю, т.е.  $a_{ij} + a_{ji} = 0$ .  $\square$

### 4.3.2. Канонический вид кососимметрической формы

Билинейную форму  $B(x, y) = \sum_{i,j} a_{ij}x_iy_j$  называют кососимметрической, если  $B(x, y) = -B(y, x)$ . В этом случае  $\sum_{i \leq j} (a_{ij} + a_{ji})x_ix_j = B(x, y) + B(y, x) \equiv 0$ , т.е.  $a_{ij} = -a_{ji}$ .

**Теорема 4.3.2.1.** Кососимметрическую билинейную форму заменой базиса можно привести к виду

$$\sum_{k=1}^r (x_{2k-1}y_{2k} - x_{2k}y_{2k-1}).$$

**Доказательство.** Пусть, например,  $a_{12} \neq 0$ . Введём вместо  $x_2$  и  $y_2$  переменные  $x'_2 = a_{12}x_2 + \dots + a_{1n}x_n$  и  $y'_2 = a_{12}y_2 + \dots + a_{1n}y_n$ . Тогда  $B(x, y) = x_1y'_2 - x'_2y_1 + (c_3x_3 + \dots + c_ny_n)y'_2 - (c_3y_3 + \dots + c_ny_n)x'_2 + \dots$ . Введём вместо  $x_1$  и  $y_1$  переменные  $x'_1 = x_1 + c_3x_3 + \dots + c_nx_n$  и  $y'_1 = y_1 + c_3y_3 + \dots + c_ny_n$ . Тогда  $B(x, y) = x'_1y'_2 - x'_2y'_1 + \dots$  (многоточием обозначены слагаемые, в которые входят переменные  $x_i$  и  $y_i$  с  $i \geq 3$ ). Для переменных  $x_3, x_4, \dots, y_3, y_4, \dots$  можно повторить такую же процедуру.  $\square$

**Следствие.** Ранг кососимметрической матрицы чётен.

Элементы  $a_{ij}$ , где  $i < j$ , можно считать независимыми переменными. Тогда доказательство теоремы 4.3.2.1 показывает, что  $A = P^T J P$ , где  $J = \text{diag} \left( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$ , а коэффициенты матрицы  $P$  являются рациональными функциями от  $a_{ij}$ . Учитывая, что  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ , матрицу  $J$  можно представить в виде произведения матриц  $J_1$  и  $J_2$ , имеющих равные определители. Поэтому  $A = (P^T J_1)(J_2 P) = FG$ , где элементы матриц  $F$  и  $G$  являются рациональными функциями от элементов матрицы  $A$  и  $\det F = \det G$ .

**Замечание.** Полученное разложение  $A = FG$  нужно понимать как равенство соответствующих рациональных функций. Для конкретных матриц  $A$  эти рациональные функции могут не иметь смысла. Но для любой конкретной матрицы  $A$  имеет место аналогичное разложение (см. задачу 3.5).

### 4.3.3. Канонический вид кососимметрического оператора

Линейный оператор  $A$  в евклидовом пространстве называют *кососимметрическим*, если в некотором ортонормированном базисе его матрица кососимметрична.

**Теорема 4.3.3.1.** Для кососимметрического оператора  $A$  существует ортонормированный базис, в котором его матрица имеет вид  $\text{diag}(\Lambda_1, \dots, \Lambda_k, 0, \dots, 0)$ , где  $\Lambda_i = \begin{pmatrix} 0 & -\lambda_i \\ \lambda_i & 0 \end{pmatrix}$ .

**Доказательство.** Оператор  $A^2$  симметрический неположительно определённый. Пусть  $V_\lambda = \{v \in V \mid A^2 v = -\lambda^2 v\}$ . Тогда  $V = \bigoplus V_\lambda$ , причём  $A V_\lambda \subset V_\lambda$ . Если  $A^2 v = 0$ , то  $(Av, Av) = -(A^2 v, v) = 0$ , т.е.  $Av = 0$ . Поэтому в пространстве  $V_0$  достаточно выбрать ортонормированный базис.

При  $\lambda \neq 0$  оператор  $A|_{V_\lambda}$  не имеет вещественных собственных значений и его квадрат равен  $-\lambda^2 I$ . Пусть  $x \in V_\lambda$  — единичный вектор,  $y = \lambda^{-1} Ax$ . Тогда  $(x, y) = (x, \lambda^{-1} Ax) = 0$ ,  $Ay = -\lambda x$  и  $(y, y) = (\lambda^{-1} Ax, y) = \lambda^{-1}(x, -Ay) = (x, x) = 1$ . Для построения ортонормированного базиса в  $V_\lambda$  возьмём, далее, единичный вектор  $u \in V_\lambda$ , ортогональный  $x$  и  $y$ . Тогда  $(Au, x) = (u, -Ax) = 0$  и  $(Au, y) = (u, -Ay) = 0$ . Дальнейший ход построения ортонормированного базиса в  $V_\lambda$  очевиден.  $\square$

### Задачи

**3.1.** Докажите, что если матрица  $A$  вещественная кососимметрическая, то матрица  $I + A$  невырожденная.

**3.2.** Невырожденная матрица  $A$  кососимметрична. Докажите, что матрица  $A^{-1}$  тоже кососимметрична.

**3.3.** а) Докажите, что относительно скалярного произведения  $\text{tr}(XY^T)$  пространства симметрических и кососимметрических матриц являются ортогональными дополнениями друг друга.

б) Докажите, что относительно скалярного произведения  $\text{Re tr}(XY^*)$  пространства эрмитовых и косоэрмитовых матриц являются ортогональными дополнениями друг друга.

**3.4.** Матрица  $A$  кососимметрическая. Докажите, что у многочлена  $f(\lambda) = \det(I + \lambda A)$  все коэффициенты при нечётных степенях  $\lambda$  равны нулю.

**3.5.** Докажите, что для любой кососимметрической матрицы  $A$  порядка  $2n$  существуют такие матрицы  $F$  и  $G$ , что  $A = FG$ ,  $\det F = \det G$  и элементы матриц  $F$  и  $G$  являются рациональными функциями от элементов матрицы  $A$ .

**3.6.** Докажите, что все корни характеристического многочлена матрицы  $AB$ , где  $A$  и  $B$  — кососимметрические матрицы порядка  $2n$ , кратные.

## 4.4. Ортогональные матрицы и преобразование Кэли

Вещественную квадратную матрицу  $A$  называют *ортогональной*, если  $AA^T = I$ . Это равенство означает, что строки матрицы  $A$  образуют ортонормированную систему. А так как  $A^T A = A^{-1}(AA^T)A = I$ , то ортонормированную систему образуют и столбцы матрицы  $A$ .

Матрица  $A$  ортогональна тогда и только тогда, когда  $(Ax, Ay) = (x, A^T Ay) = (x, y)$  для всех  $x, y$ . Достаточно даже выполнения равенства  $(Ax, Ax) = (x, x)$ , так как в этом случае  $2(Ax, Ay) = (A(x+y), A(x+y)) - (Ax, Ax) - (Ay, Ay) = 2(x, y)$ .

Если первые  $k$  строк матрицы порядка  $n$  образуют ортонормированную систему, то эту ортонормированную систему можно дополнить до ортонормированного базиса и получить ортогональную матрицу. Ортогональная матрица является унитарной, поэтому её собственные значения по модулю равны 1 (см. §3.1.3).



#### 4.4.1. Преобразование Кэли

Собственные значения ортогональной матрицы лежат на единичной окружности, а собственные значения кососимметрической матрицы лежат на мнимой оси. Дробно линейное отображение  $f(z) = \frac{1-z}{1+z}$  переводит единичную окружность в мнимую ось и  $f(f(z)) = z$ . Поэтому можно ожидать, что отображение  $f(A) = (I-A)(I+A)^{-1}$  переводит ортогональные матрицы в кососимметрические, а кососимметрические в ортогональные; это отображение называют *преобразованием Кэли*. Введём обозначение  $A^\# = (I-A)(I+A)^{-1}$ . Аналогично равенству  $f(f(z)) = z$  можно проверить равенство  $(A^\#)^\# = A$ ; при доказательстве следует учесть, что все возникающие в процессе преобразования матрицы попарно коммутируют.

**Теорема 4.4.1.1.** *Преобразование Кэли переводит любую кососимметрическую матрицу в ортогональную, а любую ортогональную матрицу  $A$ , для которой  $|A+I| \neq 0$ , в кососимметрическую.*

**Доказательство.** Так как матрицы  $I-A$  и  $I+A$  коммутируют, преобразование Кэли можно записать в виде  $A^\# = \frac{I-A}{I+A}$ . Если  $AA^T = I$  и  $|I+A| \neq 0$ , то

$$(A^\#)^T = \frac{I-A^T}{I+A^T} = \frac{I-A^{-1}}{I+A^{-1}} = \frac{A-I}{A+I} = -A^\#.$$

Если  $A^T = -A$ , то

$$(A^\#)^T = \frac{I-A^T}{I+A^T} = \frac{I+A}{I-A} = (A^\#)^{-1}.$$

□

**Замечание.** Преобразование Кэли можно записать в виде  $A^\# = (2I - (I+A))(I+A)^{-1} = 2(I+A)^{-1} - I$ .

#### 4.4.2. Обобщённое преобразование Кэли

Если  $U$  — ортогональная матрица и  $|U+I| \neq 0$ , то  $U = (I-X)(I+X)^{-1} = 2(I+X)^{-1} - I$ , где  $X = U^\#$  — кососимметрическая матрица.

Если  $S$  — симметрическая матрица, то  $S = U\Lambda U^T$ , где  $\Lambda$  — диагональная матрица,  $U$  — ортогональная матрица. При  $|U+I| \neq 0$  получаем следующее равенство:  $S = (2(I+X)^{-1} - I)\Lambda(2(I+X)^{-1} - I)^T$ , где  $X = U^\#$ .

Докажем теперь, что аналогичные формулы справедливы и в том случае, когда  $|U + I| = 0$ .

**Теорема 4.4.2.1.** *Если  $A$  — произвольная квадратная матрица, то существует такая матрица  $J = \text{diag}(\pm 1, \dots, \pm 1)$ , что  $|A + J| \neq 0$ .*

**Доказательство.** [Hs] Пусть  $n$  — порядок матрицы. При  $n = 1$  утверждение очевидно. Предположим, что утверждение верно для любой матрицы порядка  $n - 1$ , и рассмотрим матрицу  $A$  порядка  $n$ . Запишем матрицу  $A$  в блочном виде:  $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & a \end{pmatrix}$ , где  $A_1$  — матрица порядка  $n - 1$ . По предположению индукции существует такая матрица  $J = \text{diag}(\pm 1, \dots, \pm 1)$ , что  $|A_1 + J_1| \neq 0$ . Тогда

$$\left| \begin{array}{cc} A_1 + J_1 & A_2 \\ A_3 & a + 1 \end{array} \right| - \left| \begin{array}{cc} A_1 + J_1 & A_2 \\ A_3 & a - 1 \end{array} \right| = 2|A_1 + J_1| \neq 0,$$

поэтому хотя бы один из определителей, стоящих в левой части, отличен от нуля.  $\square$

**Следствие.** *Если  $U$  — ортогональная матрица, то существуют такая кососимметрическая матрица  $X$  и диагональная матрица  $J = \text{diag}(\pm 1, \dots, \pm 1)$ , что  $U = J(I - X)(I + X)^{-1}$ .*

**Доказательство.** Существует такая матрица  $J = \text{diag}(\pm 1, \dots, \pm 1)$ , что  $|U + J| \neq 0$ . Ясно, что  $J^2 = I$ . Поэтому  $|JU + I| \neq 0$ , а значит,  $JU = (I - X)(I + X)^{-1}$ , где  $X = (JU)^\#$ .  $\square$

**Теорема 4.4.2.2.** *Любую симметрическую матрицу  $S$  можно привести к диагональному виду с помощью такой ортогональной матрицы  $U$ , что  $|U + I| \neq 1$ .*

**Доказательство.** [Hs] Пусть  $S = U_1 \Lambda U_1^T$ . Согласно теореме 4.4.2.1 существует такая матрица  $J = \text{diag}(\pm 1, \dots, \pm 1)$ , что  $|U_1 + J| \neq 1$ . Тогда  $|U_1 J + I| \neq 0$ . Пусть  $U = U_1 J$ ; ясно, что  $U \Lambda U^T = U_1 \Lambda U_1^T = S$ .  $\square$

**Следствие.** *Для любой симметрической матрицы  $S$  можно выбрать кососимметрическую матрицу  $X$  и диагональную матрицу  $\Lambda$  так, что*

$$S = (2(I + X)^{-1} - I) \Lambda (2(I + X)^{-1} - I)^T.$$

### Задачи

**4.1.** Докажите, что если  $p(\lambda)$  — характеристический многочлен ортогональной матрицы порядка  $n$ , то  $\lambda^n p(1/\lambda) = \pm p(\lambda)$ .

**4.2.** Докажите, что любая унитарная матрица порядка 2 с определителем 1 имеет вид  $\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$ , где  $|u|^2 + |v|^2 = 1$ .

**4.3.** Пусть  $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  — ортогональная матрица с определителем 1, причём  $A_{11}$  и  $A_{22}$  — квадратные матрицы. Докажите, что  $\det A_{11} = \det A_{22}$ .

**4.4.** Определитель ортогональной матрицы  $A$  порядка 3 равен 1. Докажите, что:

а)  $(\operatorname{tr} A)^2 - \operatorname{tr}(A)^2 = 2 \operatorname{tr} A$ ;

б)  $\left( \sum_i a_{ii} - 1 \right)^2 + \sum_{i < j} (a_{ij} - a_{ji})^2 = 4$ .

**4.5.** Пусть  $J$  — невырожденная матрица. Матрицу  $A$  называют *J-ортогональной*, если  $A^T J A = J$ , т.е.  $A^T = J A^{-1} J^{-1}$ ; матрицу  $A$  называют *J-кососимметрической*, если  $A^T J = -J A$ , т.е.  $A^T = -J A J^{-1}$ . Докажите, что преобразование Кэли переводит *J-ортогональные* матрицы в *J-кососимметрические* и наоборот.

**4.6.** Для любого натурального  $n \geq 3$  приведите пример ортогональной матрицы  $A = \|a_{ij}\|_1^n$ , у которой  $a_{ij} = 0$  при  $j > i > 1$ , а все остальные элементы ненулевые.

**4.7.** [Dj3] Все собственные значения оператора  $A$  по модулю равны 1 и  $\|Ax\| \leq \|x\|$  для всех  $x$ . Докажите, что  $A$  — унитарный оператор.

**4.8.** [Za] Унитарный оператор  $U$  переводит некоторый ненулевой вектор  $x$  в ортогональный ему вектор  $Ux$ . Докажите, что любая дуга единичной окружности, содержащая все собственные значения оператора  $U$ , имеет длину не меньше  $\pi$ .

**4.9.** Комплексная матрица  $A$  порядка  $n$  такова, что  $\|Av\| \leq \|v\|$  и  $\|A^*v\| \leq \|v\|$  для всех  $v$ . Докажите, что существует унитарная матрица порядка  $2n$ , имеющая вид  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ .

## 4.5. Нормальные матрицы

Линейный оператор  $A$  над  $\mathbb{C}$  называют *нормальным*, если  $A^* A = A A^*$ . Следующие условия эквивалентны нормальности оператора  $A$ :

- 1)  $A = B + iC$ , где  $B$  и  $C$  — коммутирующие эрмитовы операторы, т.е.  $BC = CB$  (теорема 2.7.5.1);
- 2)  $A = V\Lambda U^*$ , где  $U$  — унитарная,  $\Lambda$  — диагональная матрица, т.е. оператор  $A$  имеет ортонормированный собственный базис (см. §3.7.1);
- 3)  $\sum_{i=1}^n |\lambda_i|^2 = \sum_{i,j=1}^n |a_{ij}|^2$ , где  $\lambda_1, \dots, \lambda_n$  — собственные значения оператора  $A$  (теорема 6.2.1.1).

#### 4.5.1. Ядро и образ нормального оператора

**Теорема 4.5.1.1.** Если  $A$  — нормальный оператор, то  $\text{Ker } A^* = \text{Ker } A$  и  $\text{Im } A^* = \text{Im } A$ .

**Доказательство.** Равенства  $A^*x = 0$  и  $Ax = 0$  эквивалентны, так как  $(A^*x, A^*x) = (x, AA^*x) = (x, A^*Ax) = (Ax, Ax)$ .

Равенство  $A^*x = 0$  означает, что  $(x, Ay) = (A^*x, y) = 0$  для всех  $y$ , т.е.  $x \in (\text{Im } A)^\perp$ . Поэтому  $\text{Im } A = (\text{Ker } A^*)^\perp$  и  $\text{Im } A^* = (\text{Ker } A)^\perp$ . А так как  $\text{Ker } A = \text{Ker } A^*$ , то  $\text{Im } A = \text{Im } A^*$ .  $\square$

**Следствие.** Если  $A$  — нормальный оператор, то  $V = \text{Ker } A \oplus (\text{Ker } A)^\perp = \text{Ker } A \oplus \text{Im } A$ .

#### 4.5.2. Собственные векторы нормального оператора

**Теорема 4.5.2.1.** Оператор  $A$  нормален тогда и только тогда, когда любой собственный вектор оператора  $A$  является собственным вектором оператора  $A^*$ .

**Доказательство.** Легко проверить, что если оператор  $A$  нормален, то оператор  $A - \lambda I$  тоже нормален, а значит,  $\text{Ker}(A - \lambda I) = \text{Ker}(A^* - \bar{\lambda}I)$ , т.е. любой собственный вектор оператора  $A$  является собственным вектором оператора  $A^*$ .

Предположим теперь, что любой собственный вектор оператора  $A$  является собственным вектором оператора  $A^*$ . Докажем, что если  $Ax = \lambda x$ , то  $A(\langle x \rangle^\perp) \subset \langle x \rangle^\perp$ . В самом деле, если  $y \in \langle x \rangle^\perp$ , то  $(x, Ay) = (A^*x, y) = (\mu x, y) = \mu(x, y) = 0$ . Возьмём произвольный собственный вектор  $e_1$  оператора  $A$ . Оператор  $A$  можно ограничить на пространство  $\langle e_1 \rangle^\perp$ . В этом пространстве возьмём произвольный собственный вектор  $e_2$  оператора  $A$  и т.д. В итоге получим ортонормированный собственный базис оператора  $A$ , поэтому  $A$  — нормальный оператор.  $\square$

### 4.5.3. Полиномиальное выражение $A^*$ через $A$

**Теорема 4.5.3.1.** Если  $A$  — нормальная матрица, то  $A^*$  полиномиально выражается через  $A$ .

**Доказательство.** Пусть  $A = U\Lambda U^*$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  и  $U$  — унитарная матрица. Тогда  $A^* = U\Lambda^* U^*$ , где  $\Lambda^* = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n)$ . Построим интерполяционный многочлен  $p$ , для которого  $p(\lambda_i) = \bar{\lambda}_i$  при  $i = 1, \dots, n$ . Тогда  $p(\Lambda) = \text{diag}(p(\lambda_1), \dots, p(\lambda_n)) = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) = \Lambda^*$ . Поэтому  $p(A) = Up(\Lambda)U^* = U\Lambda^*U^* = A^*$ .  $\square$

**Следствие.** Если  $A$  и  $B$  — нормальные матрицы и  $AB = BA$ , то  $A^*B = BA^*$  и  $AB^* = B^*A$ ; в частности,  $AB$  — нормальная матрица.

### Задачи

**5.1.** Докажите, что матрица  $A$  нормальна тогда и только тогда, когда  $A^* = AU$ , где  $U$  — унитарная матрица.

**5.2.** Пусть  $A$  — нормальная матрица. Докажите, что существует такая нормальная матрица  $B$ , что  $A = B^2$ .

**5.3.** Пусть  $A$  и  $B$  — нормальные операторы, причём  $\text{Im } A \perp \text{Im } B$ . Докажите, что  $A + B$  — нормальный оператор.

**5.4.** Докажите, что если  $A$  — нормальный оператор и  $A = SU$  — его полярное разложение, то  $SU = US$ .

**5.5.** Матрицы  $A$ ,  $B$  и  $AB$  нормальны. Докажите, что матрица  $BA$  тоже нормальна.

**5.6.** Докажите, что для любой матрицы  $A$  порядка  $n$  существует нормальная матрица порядка  $2n$ , имеющая вид  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ .

## 4.6. Нильпотентные матрицы

Квадратную матрицу  $A$  называют *нильпотентной*, если  $A^p = 0$  для некоторого  $p > 0$ .

### 4.6.1. Характеристический и минимальный многочлены

**Теорема 4.6.1.1.** Если порядок нильпотентной матрицы  $A$  равен  $n$ , то  $A^n = 0$ .

**Доказательство.** Выберем наибольшее натуральное  $p$ , для которого  $A^p \neq 0$ . Тогда  $A^p x \neq 0$  для некоторого  $x$  и  $A^{p+1} = 0$ . Докажем, что векторы  $x, Ax, \dots, A^p x$  линейно независимы. Предположим, что  $A^k x = \sum_{j=k+1}^p \lambda_j A^j x$ . Тогда  $A^{p-k}(A^k x) = A^p x \neq 0$ , но  $A^{p-k}(\lambda_i A^i x) = 0$ , так как  $i > k$ . Получено противоречие. Следовательно,  $p < n$ .  $\square$

**Теорема 4.6.1.2.** Пусть  $A$  — нильпотентная матрица,  $k$  — максимальный размер жордановых клеток матрицы  $A$ . Тогда  $A^k = 0$  и  $A^{k-1} \neq 0$ .

**Доказательство.** Все собственные значения нильпотентной матрицы нулевые. Пусть  $N$  — жорданова клетка порядка  $m$  матрицы  $A$ . Тогда существует такой базис  $e_1, \dots, e_m$ , что  $Ne_i = e_{i-1}$ , а значит,  $N^p e_i = e_{i-p}$  (предполагается, что  $e_{i-p} = 0$  при  $i - p < 0$ ). Поэтому  $N^m = 0$  и  $N^{m-1} e_m = e_1$ , т.е.  $N^{m-1} \neq 0$ .  $\square$

**Теорема 4.6.1.3.** Матрица  $A$  порядка  $n$  нильпотентна тогда и только тогда, когда её характеристический многочлен равен  $(-\lambda)^n$ .

**Доказательство.** Если жорданова нормальная форма матрицы  $A$  состоит из клеток порядка  $n_1, \dots, n_k$  с собственными значениями  $\lambda_1, \dots, \lambda_k$ , то характеристический многочлен равен  $\prod_{i=1}^k (\lambda_i - \lambda)^{n_i}$ . Матрица  $A$  нильпотентна тогда и только тогда, когда все её собственные значения равны 0, т.е. тогда и только тогда, когда характеристический многочлен равен  $(-\lambda)^n$ .  $\square$

#### 4.6.2. Критерии нильпотентности

**Теорема 4.6.2.1.** Матрица  $A$  порядка  $n$  нильпотентна тогда и только тогда, когда  $\text{tr}(A^p) = 0$  для  $p = 1, \dots, n$ .

**Доказательство.** Приведём матрицу  $A$  к нормальной жордановой форме. Матрица  $A$  нильпотентна тогда и только тогда, когда все её собственные значения равны нулю. Предположим, что матрица  $A$  имеет ненулевые собственные значения  $\lambda_1, \dots, \lambda_n$ ; пусть  $n_i$  — сумма порядков жордановых клеток, соответствующих собственному значению  $\lambda_i$ . Тогда  $\text{tr}(A^p) = n_1 \lambda_1^p + \dots + n_k \lambda_k^p$ . Так как  $k \leq n$ , достаточно доказать, что равенства  $n_1 \lambda_1^p + \dots + n_k \lambda_k^p = 0$  ( $p = 1, \dots, k$ ) не могут выполняться. Эти равенства можно рассмотреть как систему уравнений относительно  $n_1, \dots, n_k$ . Определитель этой системы равен произведению  $\lambda_1 \dots \lambda_k$  на определитель Вандермонда и, значит, отличен от нуля; поэтому  $n_1 = \dots = n_k = 0$ .  $\square$

**Теорема 4.6.2.2.** Пусть  $A: V \rightarrow V$  — линейный оператор и  $W$  — инвариантное подпространство, т.е.  $AW \subset W$ . Оператор  $A$  индуцирует операторы  $A_1: W \rightarrow W$  и  $A_2: V/W \rightarrow V/W$ . Если операторы  $A_1$  и  $A_2$  нильпотентны, то оператор  $A$  тоже нильпотентен.

**Доказательство.** Пусть  $A_1^p = 0$  и  $A_2^q = 0$ . Равенство  $A_2^q = 0$  означает, что  $A^q V \subset W$ , а равенство  $A_1^p = 0$  означает, что  $A^p W = 0$ . Поэтому  $A^{p+q} V \subset A^p W = 0$ .  $\square$

### 4.6.3. Диаграмма Юнга нильпотентной матрицы

Жорданова нормальная форма нильпотентной матрицы  $A$  имеет блочно диагональный вид с жордановыми клетками  $J_{n_1}(0), \dots, J_{n_k}(0)$ , причём  $n_1 + \dots + n_k = n$ , где  $n$  — порядок матрицы  $A$ . Можно считать, что  $n_1 \geq \dots \geq n_k$ . Тогда набору  $(n_1, n_2, \dots, n_k)$  можно сопоставить диаграмму Юнга, состоящую из  $n$  клеток, расположенных таким образом, что в  $i$ -й клетке стоит  $n_i$  клеток, причём первые клетки всех строк находятся в первом столбце (как на рис. 1.1 на с. 73).

Нильпотентные матрицы подобны тогда и только тогда, когда им соответствует одна и та же диаграмма Юнга. Размерность пространства  $\text{Ker } A^m$  выражается через  $(n_1, \dots, n_k)$  по формулам:

$$\dim \text{Ker } A = k = \text{Card}\{j \mid n_j \geq 1\},$$

$$\dim \text{Ker } A^m = \dim \text{Ker } A^{m-1} + \text{Card}\{j \mid n_j \geq m\};$$

(здесь  $\text{Card } M$  — мощность  $M$ ).

Набор  $(n'_1, \dots, n'_l)$ , где  $n'_i = \text{Card}\{j \mid n_j \geq i\}$  называют *двойственным* к  $(n_1, \dots, n_k)$ .

Диаграммы Юнга двойственных разбиений числа  $n$  получаются друг из друга транспонированием. Если нильпотентной матрице  $A$  соответствует набор  $(n_1, \dots, n_k)$ , то  $\dim \text{Ker } A^m = n'_1 + \dots + n'_m$ .

### Задачи

**6.1.** Пусть  $A$  и  $B$  — две матрицы порядка  $n$ . Докажите, что если матрица  $A + \lambda B$  нильпотентна для  $n + 1$  различных значений  $\lambda$ , то матрицы  $A$  и  $B$  нильпотентны.

**6.2.** Матрица  $A^2 - A$  нильпотентна. Докажите, что либо  $A$  нильпотентна, либо существует такой многочлен  $p$  с целыми коэффициентами и без свободного члена, что матрица  $p(A) \neq 0$  идемпотентна.

**6.3.** Укажите такие матрицы  $A$  и  $B$ , что матрица  $\lambda A + \mu B$  нильпотентна для любых  $\lambda$  и  $\mu$ , но матрицы  $A$  и  $B$  нельзя одновременно привести к треугольному виду.

## 4.7. Проекторы

Оператор  $P: V \rightarrow V$  называют *проектором* (или *идемпотентом*), если  $P^2 = P$ .

### 4.7.1. Канонический вид матрицы проектора

**Теорема 4.7.1.1.** Матрица проектора  $P$  в некотором базисе имеет вид  $\text{diag}(1, \dots, 1, 0, \dots, 0)$ .

**Доказательство.** Любой вектор  $v \in V$  можно представить в виде  $v = Pv + (v - Pv)$ , где  $Pv \in \text{Im } P$  и  $v - Pv \in \text{Ker } P$ . Кроме того, если  $x \in \text{Im } P \cap \text{Ker } P$ , то  $x = 0$ . В самом деле, тогда  $x = Py$  и  $Px = 0$ , поэтому  $0 = Px = P^2y = Py = x$ . Следовательно,  $V = \text{Im } P \oplus \text{Ker } P$ . Выберем в качестве базиса  $V$  объединение базисов  $\text{Im } P$  и  $\text{Ker } P$ . В этом базисе матрица оператора  $P$  имеет требуемый вид.  $\square$

**Следствие 1.** Существует взаимно однозначное соответствие между проекторами и разложениями  $V = W_1 \oplus W_2$ . Каждому такому разложению соответствует проектор  $P(w_1 + w_2) = w_1$  для  $w_1 \in W_1$  и  $w_2 \in W_2$ , а каждому проектору соответствует разложение  $V = \text{Im } P \oplus \text{Ker } P$ .

Оператор  $P$  можно назвать *проекцией* на  $W_1$  параллельно  $W_2$ .

**Следствие 2.** Если  $P$  — проектор, то  $\text{rk } P = \text{tr } P$ .

**Теорема 4.7.1.2.** Если  $P$  — проектор, то  $I - P$  — тоже проектор, причём  $\text{Ker}(I - P) = \text{Im } P$  и  $\text{Im}(I - P) = \text{Ker } P$ .

**Доказательство.** Если  $P^2 = P$ , то  $(I - P)^2 = I - 2P + P^2 = I - P$ . Согласно доказательству теоремы 4.7.1.1  $\text{Ker } P$  состоит из векторов  $v - Pv$ , т.е.  $\text{Ker } P = \text{Im}(I - P)$ . Аналогично  $\text{Ker}(I - P) = \text{Im } P$ .  $\square$

**Следствие.** Если  $P$  — проекция на  $W_1$  параллельно  $W_2$ , то  $I - P$  — проекция на  $W_2$  параллельно  $W_1$ .



### 4.7.2. Эрмитовы проекторы

Пусть  $P$  — проектор и  $V = \text{Im } P \oplus \text{Ker } P$ . Если  $\text{Im } P \perp \text{Ker } P$ , то  $Pv$  — ортогональная проекция вектора  $v$  на  $\text{Im } P$  (см. §2.5.3).

**Теорема 4.7.2.1.** *Проектор  $P$  эрмитов тогда и только тогда, когда  $\text{Im } P \perp \text{Ker } P$ .*

**Доказательство.** Если оператор  $P$  эрмитов, то  $\text{Ker } P = (\text{Im } P^*)^\perp = (\text{Im } P)^\perp$ . Предположим теперь, что  $P$  — проектор и  $\text{Im } P \perp \text{Ker } P$ . Векторы  $x - Px$  и  $y - Py$  лежат в  $\text{Ker } P$ , поэтому  $(Px, y - Py) = 0$  и  $(Py, x - Px) = 0$ , т.е.  $(Px, y) = (Px, Py) = (x, Py)$ .  $\square$

**Замечание.** Если проектор  $P$  эрмитов, то  $(Px, y) = (Px, Py)$ ; в частности,  $(Px, x) = \|Px\|^2$ .

**Теорема 4.7.2.2.** *Проектор  $P$  эрмитов тогда и только тогда, когда  $\|Px\| \leq \|x\|$  для всех  $x$ .*

**Доказательство.** Если оператор  $P$  эрмитов, то  $x - Px \perp Px$ , поэтому  $\|x\|^2 = \|P\|^2 + \|Px - x\|^2 \geq \|Px\|^2$ .

Докажем теперь, что если  $\|Px\| \leq \|x\|$ , то  $\text{Ker } P \perp \text{Im } P$ . Предположим, что вектор  $v \in \text{Im } P$  не перпендикулярен  $\text{Ker } P$  и  $v_1$  — проекция вектора  $v$  на  $\text{Ker } P$ . Тогда  $\|v - v_1\| < \|v\|$  и  $v = P(v - v_1)$ , поэтому  $\|v - v_1\| < \|P(v - v_1)\|$ . Получено противоречие.  $\square$

Эрмитовы проекторы  $P$  и  $Q$  называют *ортогональными*, если  $\text{Im } P \perp \text{Im } Q$ , т.е.  $PQ = QP = 0$ .

**Теорема 4.7.2.3.** *Пусть  $P_1, \dots, P_n$  — эрмитовы проекторы. Оператор  $P = P_1 + \dots + P_n$  является проектором тогда и только тогда, когда  $P_i P_j = 0$  при  $i \neq j$ .*

**Доказательство.** Если  $P_i P_j = 0$  при  $i \neq j$ , то  $P^2 = (P_1 + \dots + P_n)^2 = P_1^2 + \dots + P_n^2 = P_1 + \dots + P_n = P$ .

Предположим теперь, что  $P = P_1 + \dots + P_n$  — проектор. Этот проектор эрмитов, значит, если  $x \in \text{Im } P$ , то  $x = P_i x$  и  $\|x\|^2 = \|P_i x\|^2 \leq \|P_1 x\|^2 + \dots + \|P_n x\|^2 = (P_1 x, x) + \dots + (P_n x, x) = (Px, x) = \|Px\|^2 \leq \|x\|^2$ . Поэтому  $P_j x = 0$  при  $i \neq j$ , т.е.  $P_j P_i = 0$ .  $\square$

### 4.7.3. Матрица проектора на подпространство

Пусть  $W \subset V$ ;  $a_1, \dots, a_k$  — базис  $W$ . Рассмотрим матрицу  $A$  размера  $n \times k$ , столбцами которой являются координаты векторов  $a_1, \dots, a_k$  относительно ортонормированного базиса пространства  $V$ . Тогда  $\text{rk } A^*A = \text{rk } A = k$ , поэтому матрица  $A^*A$  невырождена.

Ортогональную проекцию  $Pv$  вектора  $v$  на подпространство  $W$  можно выразить с помощью матрицы  $A$ . В самом деле, с одной стороны,  $Pv = x_1a_1 + \dots + x_ka_k$ , т.е.  $Pv = Ax$ , где  $x$  — столбец  $(x_1, \dots, x_k)^T$ . С другой стороны,  $Pv - v \perp W$ , т.е.  $A^*(v - Ax) = 0$ . Поэтому  $x = (A^*A)^{-1}A^*v$ , а значит,  $Pv = Ax = A(A^*A)^{-1}A^*x$ , т.е.  $P = A(A^*A)^{-1}A^*$ .

Если базис  $a_1, \dots, a_k$  ортонормирован, то  $A^*A = I$ , а значит,  $P = AA^*$ .

### 4.7.4. Определитель суммы проекторов

**Теорема 4.7.4.1.** [Dj3] Пусть  $V = V_1 \oplus \dots \oplus V_k$ , причём  $V_i \neq 0$  при  $i = 1, \dots, k$ ;  $P_i: V \rightarrow V_i$  — ортогональный проектор и  $A = P_1 + \dots + P_k$ . Тогда  $0 < |A| \leq 1$ , причём  $|A| = 1$  тогда и только тогда, когда  $V_i \perp V_j$ .

**Доказательство.** Докажем предварительно две леммы; в дальнейшем  $P_i$  обозначает ортогональный проектор на  $V_i$ ,  $P_{ij}: V_i \rightarrow V_j$  — ограничение  $P_j$  на  $V_i$ .

**Лемма 1.** Пусть  $V = V_1 \oplus V_2$ , причём  $V_i \neq 0$ . Тогда  $|I - P_{12}P_{21}| \leq 1$ , причём равенство достигается тогда и только тогда, когда  $V_1 \perp V_2$ .

**Доказательство.** Операторы  $P_1$  и  $P_2$  эрмитовы неотрицательно определённые, поэтому оператор  $A = P_1 + P_2$  тоже неотрицательно определён. Кроме того, если  $Ax = P_1x + P_2x = 0$ , то  $P_1x = P_2x = 0$ , так как  $P_1x \in V_1$  и  $P_2x \in V_2$ . Поэтому  $x \perp V_1$  и  $x \perp V_2$ , а значит,  $x = 0$ . Следовательно, оператор  $A$  положительно определён и  $|A| > 0$ .

Выберем в качестве базиса пространства  $V$  объединение базисов  $V_1$  и  $V_2$ . В этом базисе матрица  $A$  имеет вид  $\begin{pmatrix} I & P_{21} \\ P_{12} & I \end{pmatrix}$ ; положим  $B = \begin{pmatrix} I & 0 \\ P_{12} & I - P_{12}P_{21} \end{pmatrix}$ . Легко проверить, что  $|I - P_{12}P_{21}| = |B| = |A| > 0$ . Докажем теперь, что все собственные значения оператора  $I - P_{12}P_{21}$  (т.е. ограничения  $B$  на  $V_2$ ) по модулю не превосходят 1. В самом деле, если  $x \in V_2$ , то  $\|Bx\|^2 = (Bx, Bx) = (x - P_2P_1x, x - P_2P_1x) = \|x\|^2 - (P_2P_1x, x) - (x, P_2P_1x) + \|P_2P_1x\|^2$ . А так как  $(P_2P_1x, x) = (P_1x, P_2x) = (P_1x, x) = \|P_1x\|^2$ ,  $(x, P_2P_1x) = \|P_1x\|^2$  и  $\|P_1x\|^2 \geq \|P_2P_1x\|^2$ , то

$$\|Bx\|^2 \leq \|x\|^2 - \|P_1x\|^2. \quad (4.1)$$

Все собственные значения оператора  $I - P_{12}P_{21}$  по модулю не превосходят 1 и его определитель положителен, поэтому  $0 < |I - P_{12}P_{21}| \leq 1$ .

Если  $|I - P_{12}P_{21}| = 1$ , то все собственные значения оператора  $I - P_{12}P_{21}$  равны 1, поэтому, учитывая (4.1), получаем, что этот оператор унитарен (см. задачу 4.7). Следовательно,  $\|Bx\| = \|x\|$  для любого  $x \in V_2$ . Снова учитывая (4.1), получаем  $\|P_1x\| = 0$ , т.е.  $V_2 \perp V_1$ .  $\square$

**Лемма 2.** Пусть  $V = V_1 \oplus V_2$ , причём  $V_i \neq 0$ ;  $H$  — такой эрмитов оператор, что  $\text{Im } H = V_1$  и оператор  $H_1 = H|_{V_1}$  положительно определён. Тогда  $0 < |H + P_2| \leq |H_1|$ , причём равенство достигается тогда и только тогда, когда  $V_1 \perp V_2$ .

**Доказательство.** Возьмём в качестве базиса пространства  $V$  объединение базисов пространств  $V_1$  и  $V_2$ . В этом базисе матрица оператора  $H + P_2$  имеет вид  $\begin{pmatrix} H_1 & H_1P_{21} \\ P_{12} & I \end{pmatrix}$ . В самом деле, так как  $\text{Ker } H = (\text{Im } H)^\perp = (V_1)^\perp$ , то  $H = HP_1$ ; поэтому  $H|_{V_2} = H_1P_{21}$ . Легко проверить, что

$$\begin{vmatrix} H_1 & H_1P_{21} \\ P_{12} & I \end{vmatrix} = \begin{vmatrix} H_1 & 0 \\ P_{12} & I - P_{12}P_{21} \end{vmatrix} = |H_1| \cdot |I - P_{12}P_{21}|.$$

Остаётся воспользоваться леммой 1.  $\square$

Займёмся теперь непосредственно доказательством теоремы 4.7.4.1. Аналогично доказательству леммы 1 можно доказать, что  $|A| > 0$ . Доказательство второго неравенства проведём индукцией по  $k$ . При  $k = 1$  утверждение очевидно. При  $k > 1$  рассмотрим пространство  $W = V_1 \oplus \dots \oplus V_{k-1}$ . Пусть  $Q_i = P_i|_W$  ( $i = 1, \dots, k-1$ ),  $H_1 = Q_1 + \dots + Q_{k-1}$ . По предположению индукции  $|H_1| \leq 1$ ; кроме того,  $|H_1| > 0$ . Применив к оператору  $H = P_1 + \dots + P_{k-1}$  лемму 2, получим  $0 < |A| = |H + P_k| \leq |H_1| \leq 1$ .

Если  $|A| = 1$ , то согласно лемме 2  $V_k \perp W$ ; кроме того,  $|H_1| = 1$ , поэтому  $V_i \perp V_j$  при  $i, j \leq k-1$ .  $\square$

**Теорема 4.7.4.2.** Пусть  $N$  — нормальный оператор в  $V$ , ненулевые собственные значения которого равны  $\lambda_{r_1i}, \dots, \lambda_{r_ki}$ . Пусть, далее,  $r_1 + \dots + r_k \leq \dim V$ . Если ненулевые собственные значения оператора  $N = N_1 + \dots + N_k$  равны  $\lambda_{ji}$ , где  $j = 1, \dots, r_i$ , то оператор  $N$  нормален,  $\text{Im } N_i \perp \text{Im } N_j$  и  $N_iN_j = 0$  при  $i \neq j$ .

**Доказательство.** [Dj3] Пусть  $V_i = \text{Im } N_i$ . Так как  $\text{rk } N = \text{rk } N_1 + \dots + \text{rk } N_k$ , то  $W = V_1 + \dots + V_k$  — прямая сумма. Для нормального оператора  $\text{Ker } N_i = (\text{Im } N_i)^\perp$ , поэтому  $\text{Ker } N_i \subset W^\perp$ , а значит,  $\text{Ker } N \subset W^\perp$ . Ясно также, что  $\dim \text{Ker } N = \dim W^\perp$ . Поэтому без потери общности можно

ограничиться пространством  $W$  и считать, что  $r_1 + \dots + r_k = \dim V$ , т.е.  $\det N \neq 0$ .

Пусть  $M_i = N_i|_{V_i}$ . Возьмём в качестве базиса пространства  $V$  объединение базисов пространств  $V_i$ . Так как  $N = \sum N_i = \sum N_i P_i = \sum M_i P_i$ , в этом базисе матрица оператора  $N$  имеет вид

$$\begin{pmatrix} M_1 P_{11} & \dots & M_1 P_{k1} \\ \dots & \dots & \dots \\ M_k P_{1k} & \dots & M_k P_{kk} \end{pmatrix} = \begin{pmatrix} M_1 & & 0 \\ & \ddots & \\ 0 & & M_k \end{pmatrix} \begin{pmatrix} P_{11} & \dots & P_{k1} \\ \dots & \dots & \dots \\ P_{1k} & \dots & P_{kk} \end{pmatrix}.$$

Условие на собственные значения операторов  $N_i$  и  $N$  влечёт равенство  $|N - \lambda I| = \prod_{i=1}^k |M_i - \lambda I|$ . В частности, при  $\lambda = 0$  получаем  $|N| = \prod_{i=1}^k |M_i|$ .

Следовательно,  $\begin{vmatrix} P_{11} & \dots & P_{k1} \\ \dots & \dots & \dots \\ P_{1k} & \dots & P_{kk} \end{vmatrix} = 1$ , т.е.  $|P_1 + \dots + P_k| = 1$ . Применяя теорему 4.7.4.1, получаем, что  $V = V_1 \oplus \dots \oplus V_k$  — прямая сумма ортогональных подпространств. Следовательно, оператор  $N$  нормален (см. §3.7.1) и  $N_i N_j = 0$ , так как  $\text{Im } N_j \subset (\text{Im } N_i)^\perp = \text{Ker } N_i$ .  $\square$

## Задачи

**7.1.** Пусть  $P_1$  и  $P_2$  — проекторы. Докажите, что:

- а)  $P_1 + P_2$  — проектор тогда и только тогда, когда  $P_1 P_2 = P_2 P_1 = 0$ ;
- б)  $P_1 - P_2$  — проектор тогда и только тогда, когда  $P_1 P_2 = P_2 P_1 = P_2$ .

**7.2.** Найдите все матрицы порядка 2, являющиеся проекторами.

**7.3.** Пусть  $A$  — унитарный оператор. Тогда  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} A^i x = Px$ , где  $P$  — эрмитов проектор на  $\text{Ker}(A - I)$  (*эргодическая теорема*).

**7.4.** Операторы  $A_1, \dots, A_k$  в пространстве  $V$  размерности  $n$  таковы, что  $A_1 + \dots + A_k = I$ . Докажите, что следующие условия эквивалентны:

- (а) все операторы  $A_i$  — проекторы;
- (б)  $A_i A_j = 0$  при  $i \neq j$ ;
- (в)  $\text{rk } A_1 + \dots + \text{rk } A_k = n$ .

**7.5.** Пусть  $A$  — неотрицательно определённая матрица порядка  $n$ , причём матрица  $I - A$  тоже неотрицательно определённая. Докажите, что существует эрмитов проектор порядка  $2n$ , имеющий вид  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ .

## 4.8. Инволюции

Линейный оператор  $A$  называют *инволюцией*, если  $A^2 = I$ . Легко проверить, что оператор  $P$  является проектором тогда и только тогда, когда оператор  $2P - I$  является инволюцией. В самом деле, равенство  $I = (2P - I)^2 = 4P^2 - 4P + I$  эквивалентно равенству  $P^2 = P$ .

### 4.8.1. Канонический вид матрицы инволюции

**Теорема 4.8.1.1.** *Любая инволюция в некотором базисе имеет вид  $\text{diag}(\pm 1, \dots, \pm 1)$ .*

**Доказательство.** Если  $A$  — инволюция, то  $P = (A + I)/2$  — проектор; этот проектор в некотором базисе имеет вид  $\text{diag}(1, \dots, 1, 0, \dots, 0)$  (теорема 4.7.1.1). Оператор  $A = 2P - I$  в этом базисе имеет вид  $\text{diag}(1, \dots, 1, -1, \dots, -1)$ .  $\square$

**Замечание.** Воспользовавшись разложением  $x = \frac{1}{2}(x - Ax) + \frac{1}{2}(x + Ax)$ , можно доказать, что  $V = \text{Ker}(A + I) \oplus \text{Ker}(A - I)$ .

### 4.8.2. Произведение двух инволюций

**Теорема 4.8.2.1.** *Матрицу  $A$  можно представить в виде произведения двух инволюций тогда и только тогда, когда матрицы  $A$  и  $A^{-1}$  подобны.*

**Доказательство.** [Dj2] Если  $A = ST$ , где  $S$  и  $T$  — инволюции, то  $A^{-1} = TS = S(ST)S = SAS^{-1}$ .

Предположим теперь, что матрицы  $A$  и  $A^{-1}$  подобны. Жорданова нормальная форма матрицы  $A$  имеет вид  $\text{diag}(J_1, \dots, J_k)$ , поэтому матрицы  $\text{diag}(J_1, \dots, J_k)$  и  $\text{diag}(J_1^{-1}, \dots, J_k^{-1})$  подобны. Если  $J$  — жорданова клетка, то матрица  $J^{-1}$  подобна жордановой клетке. Следовательно, матрицы  $J_1, \dots, J_k$  можно разбить на два класса: для матриц первого класса  $J_\alpha^{-1} \sim J_\alpha$ , а для матриц второго класса  $J_\alpha^{-1} \sim J_\beta$  и  $J_\beta^{-1} \sim J_\alpha$ . Достаточно доказать, что в виде произведения двух инволюций можно представить матрицу  $J_\alpha$  в первом случае и матрицу  $\text{diag}(J_\alpha, J_\beta)$  во втором случае.

Характеристический многочлен жордановой клетки совпадает с минимальным многочленом, поэтому если  $p$  и  $q$  — минимальные многочлены матриц  $J_\alpha$  и  $J_\alpha^{-1}$ , то  $q(\lambda) = p(0)^{-1}\lambda^n p(\lambda^{-1})$ , где  $n$  — порядок матрицы  $J_\alpha$  (см. задачу 3.4).

Пусть  $J_\alpha \sim J_\alpha^{-1}$ . Тогда  $p(\lambda) = p(0)^{-1}\lambda^n p(\lambda^{-1})$ , т.е.  $p(\lambda) = \sum \alpha_i \lambda^{n-i}$ , где  $\alpha_0 = 1$  и  $\alpha_n \alpha_{n-i} = \alpha_i$ . Матрица  $J_\alpha$  подобна циклической клетке, поэтому существует такой базис  $e_1, \dots, e_n$ , что  $J_\alpha e_k = e_{k+1}$  при  $k \leq n-1$  и  $J_\alpha e_n = -\alpha_n e_1 - \alpha_{n-1} e_2 - \dots - \alpha_1 e_n$ . Пусть  $Te_k = e_{n+1-k}$ ; очевидно, что  $T$  — инволюция. Если  $STe_k = J_\alpha e_k$ , то  $Se_{n+1-k} = e_{k+1}$  при  $k \neq n$  и  $Se_1 = -\alpha_n e_1 - \dots - \alpha_1 e_n$ . Проверим, что  $S$  — инволюция:  $S^2 e_1 = \alpha_n(\alpha_n e_1 + \dots + \alpha_1 e_n) - \alpha_{n-1} e_n - \dots - \alpha_1 e_2 = e_1 + (\alpha_n \alpha_{n-1} - \alpha_1) e_2 + \dots + (\alpha_n \alpha_1 - \alpha_{n-1}) e_n = e_1$ ; при  $i \neq 1$  равенство  $S^2 e_i = e_i$  очевидно.

Рассмотрим теперь случай  $J_\alpha^{-1} \sim J_\beta$ ; пусть  $\sum \alpha_i \lambda^{n-i}$  и  $\sum \beta_i \lambda^{n-i}$  — минимальные многочлены матриц  $J_\alpha$  и  $J_\beta$ . Тогда  $\sum \alpha_i \lambda^{n-i} = \beta_n^{-1} \lambda^n \sum \beta_i \lambda^{i-n} = \beta_n^{-1} \sum \beta_i \lambda^i$ . Поэтому  $\alpha_{n-i} \beta_n = \beta_i$  и  $\alpha_n \beta_n = \beta_0 = 1$ . Существуют такие базисы  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_n$ , что  $J_\alpha e_k = e_{k+1}$ ,  $J_\alpha e_n = -\alpha_n e_1 - \dots - \alpha_1 e_n$  и  $J_\beta \varepsilon_{k+1} = \varepsilon_k$ ,  $J_\beta \varepsilon_1 = -\beta_1 \varepsilon_1 - \dots - \beta_n \varepsilon_n$ . Пусть  $Te_k = \varepsilon_k$  и  $T\varepsilon_k = e_k$ . Если  $\text{diag}(J_\alpha, J_\beta) = ST$ , то  $Se_{k+1} = ST\varepsilon_{k+1} = J_\beta \varepsilon_{k+1} = \varepsilon_k$ ,  $S\varepsilon_k = e_{k+1}$ ,  $Se_1 = ST\varepsilon_1 = J_\beta \varepsilon_1 = -\beta_1 \varepsilon_1 - \dots - \beta_n \varepsilon_n$  и  $S\varepsilon_n = -\alpha_n e_1 - \dots - \alpha_1 e_n$ . Проверим, что  $S$  — инволюция. Равенства  $S^2 e_i = e_i$  и  $S^2 \varepsilon_j = \varepsilon_j$  очевидны при  $i \neq 1$  и  $j \neq n$ .  $S^2 e_1 = S(-\beta_1 \varepsilon_1 - \dots - \beta_n \varepsilon_n) = -\beta_1 e_2 - \dots - \beta_{n-1} e_n + \beta_n(\alpha_n e_1 + \dots + \alpha_1 e_n) = e_1 + (\beta_n \alpha_{n-1} - \beta_1) e_2 + \dots + (\beta_n \alpha_1 - \beta_{n-1}) e_n = e_1$ . Аналогично  $S^2 \varepsilon_n = \varepsilon_n$ .  $\square$

**Следствие.** Если  $B$  — невырожденная матрица и  $X^T B X = B$ , то матрицу  $X$  можно представить в виде произведения двух инволюций.

**Доказательство.** Если  $X^T B X = B$ , то  $X^T = B X^{-1} B^{-1}$ , т.е. матрицы  $X^T$  и  $X^{-1}$  подобны. Кроме того, матрицы  $X$  и  $X^T$  подобны для любой матрицы  $X$  (см. задачу 2.2).  $\square$

## Задачи

**8.1.** Найдите все матрицы порядка 2, являющиеся инволюциями.

**8.2.** Пусть  $A$  — симметрическая матрица,  $H = (\bar{A}A + I)^{-1}(\bar{A}A - I)$  и  $S = 2A(\bar{A}A + I)^{-1}$ . Докажите, что  $\begin{pmatrix} -H & \bar{S} \\ S & \bar{H} \end{pmatrix}$  — инволюция.

## Решения

**1.1.** Ясно, что  $(S\bar{S})^T = (\bar{S})^T S^T = \bar{S}S = \overline{(S\bar{S})}$ .

**1.2.** Для матрицы  $S$  существует такая матрица  $A$ , что  $S = A^T A$ . Главные миноры порядка  $k$  матрицы  $S$  неотрицательны, поэтому достаточно доказать, что

сумма главных миноров порядка  $k$  матрицы  $S$  положительна. Но эта сумма равна сумме квадратов всех миноров порядка  $k$  матрицы  $A$  (задача 2.4). Так как  $\text{rk } A = k$ , то по крайней мере один минор порядка  $k$  матрицы  $A$  отличен от нуля.

**Замечание.** Аналогичное утверждение верно и для эрмитовой неотрицательно определённой матрицы  $S$ .

**1.3.** Пусть  $S = U\Lambda U^T$ , где  $U$  — ортогональная матрица,  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ . Тогда  $S = S_1 + \dots + S_r$ , где  $S_i = U \text{diag}(0, \dots, 0, \lambda_i, 0, \dots, 0) U^T$ .

**1.4.** а) Пусть  $A = UDU^{-1}$ , где  $U$  — унитарная матрица,  $D$  — диагональная матрица с положительными элементами. Тогда  $A^{-1} = UD^{-1}U^{-1}$ . Здесь  $D^{-1}$  — тоже диагональная матрица с положительными элементами.

б) Следует из а), поскольку  $\text{adj } A = (\det A)A^{-1}$  и  $\det A > 0$ .

**1.5.** Равенство  $(x^T 0) \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} x \\ 0 \end{pmatrix} = 0$  показывает, что рассматриваемая квадратичная форма тождественно равна нулю на пространстве, состоящем из векторов, у которых последние  $n$  координат (где  $n$  — порядок матрицы  $A$ ) нулевые. Пусть рассматриваемая квадратичная форма приводится к виду  $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2$ . Тогда согласно теореме 4.1.2.4  $m \leq \min(p, q)$ . С другой стороны,  $p + q = 2n$ . Поэтому  $p = q = n$ .

**1.6.** Рассмотрим оператор  $B = \sum_{k=1}^n \lambda_k v_k v_k^*$ . Достаточно доказать, что  $Bv_i = \lambda_i v_i$ , т.е.  $\sum_{k=1}^n \lambda_k v_k v_k^* v_i = \lambda_i v_i$ . Но из ортонормированности базиса следует, что  $v_k^* v_i = \delta_{ki}$ .

**1.7.** Матрице  $A$  соответствует квадратичная форма  $f(x, x) = \sum_i x_i^2 + 2\lambda \sum_{i < j} x_i x_j$ . Ясно, что  $2 \sum_{i < j} x_i x_j \leq \sum_i x_i^2$ . Поэтому если  $-1 < \lambda < 1$  и  $\sum_i x_i^2 \neq 0$ , то  $|2\lambda \sum_{i < j} x_i x_j| < \sum_i x_i^2$ . Следовательно,  $f(x, x) = \sum_i x_i^2 + 2\lambda \sum_{i < j} x_i x_j > 0$ .

**1.8.** Легко проверить, что

$$\begin{pmatrix} A & A \\ A & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (A(x+y), (x+y)) \geq 0.$$

**1.9.** Рассмотрим в пространстве многочленов степени не выше  $n-1$  скалярное произведение  $(f, g) = \int_0^1 f(x)g(x)dx$ . Данная матрица является матрицей Грама системы линейно независимых векторов  $1, x, x^2, \dots, x^{n-1}$ , поэтому согласно теореме 4.1.1.2 эта матрица положительно определённая.

**1.10.** Пусть  $U$  — такая ортогональная матрица, что  $U^{-1}AU = \Lambda$  и  $|U| = 1$ . Сделаем замену  $x = Uy$ . Тогда  $(x, Ax) = (y, \Lambda y)$  и  $dx_1 \dots dx_n = dy_1 \dots dy_n$ , так

как якобиан замены переменных равен  $|U|$ . Поэтому

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-(x, Ax)} dx &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{-\lambda_1 y_1^2 - \dots - \lambda_n y_n^2} dy = \\ &= \prod_{i=1}^n \int_{-\infty}^{+\infty} e^{-\lambda_i y_i^2} dy_i = \prod_{i=1}^n \sqrt{\frac{\pi}{\lambda_i}} = (\sqrt{\pi})^n |A|^{-1/2}. \end{aligned}$$

**1.11.** Ясно, что  $(Be_i, e_j) = (AA^{-1}Be_i, e_j) = \lambda_i(Ae_i, e_j)$ . Аналогично  $(Be_j, e_i) = \lambda_j(Ae_j, e_i)$ . Остаётся заметить, что  $(Be_i, e_j) = (Be_j, e_i)$  и  $(Ae_i, e_j) = (Ae_j, e_i)$ , так как матрицы  $A$  и  $B$  симметричны.

**1.12.** Скалярное произведение  $i$ -й строки  $S$  на  $j$ -й столбец  $S^{-1}$  равно нулю при  $i \neq j$ . Поэтому в каждом столбце  $S^{-1}$  есть положительный и отрицательный элемент, а значит, число её ненулевых элементов не менее  $2n$ , а число нулевых элементов не более  $n^2 - 2n$ .

Пример матрицы  $S^{-1}$ , имеющей  $2n$  ненулевых элементов, выглядит следующим образом:

$$S^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & \dots \\ 1 & 2 & 2 & 2 & 2 & \dots \\ 1 & 2 & 1 & 1 & 1 & \dots \\ 1 & 2 & 1 & 2 & 2 & \dots \\ 1 & 2 & 1 & 2 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 0 & 1 & & & \\ & 1 & 0 & -1 & & \\ & & -1 & 0 & & \\ & & & & \ddots & \\ & & & & & 0 & -s \\ 0 & & & & & -s & -s \end{pmatrix},$$

где  $s = (-1)^n$ .

**1.13.** Для матриц  $A$  и  $B$  выберем неотрицательно определённые матрицы  $S$  и  $T$  так, что  $A = S^2$  и  $B = T^2$  (теорема 4.1.4.1). Тогда  $\text{tr}(AB) = \text{tr}(SSTT) = \text{tr}(TSST) = \text{tr}((ST)^*ST) \geq 0$ .

**1.14.** Ясно, что  $\text{tr}(A^2) - \text{tr}(B^2) = \text{tr}((A - B)A) + \text{tr}(B(A - B))$ . Остаётся воспользоваться результатом задачи 1.13.

**1.15.** Достаточно воспользоваться тождеством

$$\begin{pmatrix} \lambda I_n & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} A & B \\ B^* & 0 \end{pmatrix} \begin{pmatrix} \lambda I_n & 0 \\ 0 & I_m \end{pmatrix} = \lambda \begin{pmatrix} \lambda A & B \\ B^* & 0 \end{pmatrix}.$$

**1.16.** а) Эквивалентность этих форм над  $\mathbb{Q}$  следует из того, что они имеют одинаковую сигнатуру. Предположим, что эти формы эквивалентны над  $\mathbb{Z}$ . Тогда существует целочисленная матрица  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , для которой

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2ac & ad + bc \\ ad + bc & 2bd \end{pmatrix}.$$



Но если  $b$  и  $d$  — целые числа, то  $2bd \neq 1$ .

б) Равенства

$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2k \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \\ k & k+1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & k \\ 1 & k+1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2k+1 \end{pmatrix}$$

показывают, что квадратичная форма с матрицей  $\begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix}$  эквивалентна квадратичной форме с матрицей  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  при  $n$  чётном и квадратичной форме с матрицей  $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$  при  $n$  нечётном.

**2.1.** Пусть  $a_{ii} = 0$  и  $a_{ij} \neq 0$ . Рассмотрим вектор  $u$ ,  $i$ -я и  $j$ -я координаты которого равны  $x_i$  и  $x_j$ , а все остальные координаты равны нулю. Тогда  $B(u, u) = a_{ij}x_j\bar{x}_i + \bar{a}_{ij}\bar{x}_jx_i + a_{jj}|x_j|^2$ . Фиксируем  $x_j \neq 0$ , и пусть  $x_i = \lambda a_{ij}x_j$ , где  $\lambda \in \mathbb{R}$ . Тогда  $B(u, u) = 2\lambda|a|^2 + b$ , где  $a = a_{ij}x_j \neq 0$  и  $b = a_{jj}|x_j|^2$ . Величина  $2\lambda|a|^2 + b$  принимает как положительные, так и отрицательные значения, поэтому форма  $B$  не является знакоопределённой.

**2.2.** Выберем матрицу  $T$  так, что  $A = T^*T$  и  $B = T^*DT$ , где  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  и числа  $\lambda_1, \dots, \lambda_n$  положительны. Тогда  $\text{adj}(tA+B) = \text{adj } T \text{ adj}(tI+D) \text{ adj } T^*$ , причём  $\text{adj } T^* = (\text{adj } T)^*$ . Матрица  $\text{adj}(tI+D)$  диагональна, причём на диагонали на  $i$ -м месте стоит элемент  $\prod_{j \neq i} (t + \lambda_j)$ . Таким образом, эта матрица имеет вид  $t^{n-1}I + t^{n-2}\hat{A}_1 + \dots + \hat{A}_{n-1}$ , где матрицы  $\hat{A}_1, \dots, \hat{A}_{n-1}$  диагональные и на  $i$ -м месте матрицы  $\hat{A}_k$  стоит число  $\sigma_k(\lambda_1, \dots, \hat{\lambda}_i, \dots, \lambda_n)$ . Это число положительно, поэтому матрицы  $\hat{A}_0 = I, \hat{A}_1, \dots, \hat{A}_{n-1}$  положительно определённые. Остаётся заметить, что если  $X$  — положительно определённая матрица, а  $S$  — невырожденная матрица, то матрица  $S^*XS$  положительно определённая.

**2.3.** Нет, не обязательно. Пусть  $A_1 = B_1 = \text{diag}(0, 1, -1)$ ,  $A_2 = \begin{pmatrix} 0 & \sqrt{2} & 2 \\ \sqrt{2} & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$  и  $B_2 = \begin{pmatrix} 0 & 0 & \sqrt{2} \\ 0 & 0 & 2 \\ \sqrt{2} & 2 & 0 \end{pmatrix}$ . Легко проверить, что  $|xA_1 + yA_2 + \lambda I| = \lambda^3 - \lambda(x^2 + 6y^2) - 2y^2x = |xB_1 + yB_2 + \lambda I|$ . Предположим теперь, что существует такая ортогональная матрица  $U$ , что  $UA_1U^T = B_1 = A_1$  и  $UA_2U^T = B_2$ . Тогда  $UA_1 = A_1U$ , а так как матрица  $A_1$  диагональная и на диагонали стоят попарно различные числа, то  $U$  — диагональная ортогональная матрица (см. задачу 1.1 а), т.е.  $U = \text{diag}(\lambda, \mu, \nu)$ , где  $\lambda, \mu, \nu = \pm 1$ . Следовательно,

$$\begin{pmatrix} 0 & 0 & \sqrt{2} \\ 0 & 0 & 2 \\ \sqrt{2} & 2 & 0 \end{pmatrix} = B_2 = UA_2U^T = \begin{pmatrix} 0 & \sqrt{2}\lambda\mu & 2\lambda\mu \\ \sqrt{2}\lambda\mu & 0 & 0 \\ 2\lambda\mu & 0 & 0 \end{pmatrix}.$$

Получено противоречие.

**3.1.** Ненулевые собственные значения матрицы  $A$  чисто мнимые, поэтому  $-1$  не может быть её собственным значением.

**3.2.** Так как  $(-A)^{-1} = -A^{-1}$ , то  $(A^{-1})^T = (A^T)^{-1} = (-A)^{-1} = -A^{-1}$ .

**3.3.** Размерности этих пространств дополнительные, поэтому нужно лишь проверить, что если  $X^* = X$  и  $Y^* = -Y$ , то  $\operatorname{Re} \operatorname{tr}(XY) = 0$ , т.е.  $\operatorname{tr}(XY + \overline{XY}) = 0$ . Но  $\operatorname{tr}(XY) = \operatorname{tr}(XY)^T = \operatorname{tr}(-\bar{Y} \cdot \bar{X}) = -\operatorname{tr}(\bar{Y}\bar{X}) = -\operatorname{tr}(\overline{XY})$ .

**3.4.** Ясно, что  $|I + \lambda A| = |(I + \lambda A)^T| = |I - \lambda A|$ . Поэтому  $f(\lambda) = f(-\lambda)$ , т.е.  $f$  — чётная функция.

**3.5.** Доказательство проведём индукцией по  $n$ . При  $n = 1$  имеем  $\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} = \begin{pmatrix} 0 & x \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -x & 0 \\ 0 & 1 \end{pmatrix}$ . Для доказательства шага индукции разберём два случая.

1) Предположим сначала, что  $A = \begin{pmatrix} A_1 & A_2 \\ -A_2^T & A_3 \end{pmatrix}$ , где  $A_1 = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}$  и  $x \neq 0$ . Пусть  $F = \begin{pmatrix} F_1 & 0 \\ X & F_2 \end{pmatrix}$  и  $G = \begin{pmatrix} G_1 & Y \\ 0 & G_2 \end{pmatrix}$ . Тогда  $FG = \begin{pmatrix} F_1 G_1 & F_1 Y \\ X G_1 & X Y + F_2 G_2 \end{pmatrix}$ . Матрицы  $F_1$  и  $G_1$  можно выбрать так, что  $F_1 G_1 = A_1$ . Возьмём, далее,  $Y = F_1^{-1} A_2$  и  $X = -A_2^T G_1^{-1}$ . Тогда  $XY + F_2 G_2 = -A_2^T G_1^{-1} F_1^{-1} A_2 + F_2 G_2 = -A_2^T A_1^{-1} A_2 + F_2 G_2$ . Матрица  $A_3 + A_2^T A_1^{-1} A_2$  кососимметрична, поэтому по предположению индукции матрицы  $F_2$  и  $G_2$  можно подобрать так, что  $F_2 G_2 = A_3 + A_2^T A_1^{-1} A_2$  и  $\det F = \det G$ .

2) Рассмотрим теперь случай, когда  $a_{12} = 0$ , но  $a_{ij} \neq 0$  для некоторых  $i$  и  $j$ . Пусть  $P$  — ортогональная матрица, соответствующая перестановке  $\begin{pmatrix} 1 & 2 \\ i & j \end{pmatrix} \dots$  (см. задачу 2.3). Матрица  $PAP^T$  относится к первому случаю, поэтому существуют такие матрицы  $F_1$  и  $G_1$ , что  $PAP^T = F_1 G_1$ . Тогда  $A = (P^{-1} F_1)(G_1 P) = FG$ , где  $F = P^{-1} F_1$  и  $G = G_1 P$ .

**3.6.** Нам неоднократно придётся пользоваться тем, что для кососимметрической матрицы  $A$  чётного порядка число  $\dim \operatorname{Ker} A$  чётное, так как ранг кососимметрической матрицы чётен (см. § 3.2).

Рассмотрим сначала случай нулевого собственного значения, т.е. докажем, что если  $\dim \operatorname{Ker} AB \geq 1$ , то  $\dim \operatorname{Ker} AB \geq 2$ . Если  $|B| = 0$ , то  $\dim \operatorname{Ker} AB \geq \dim \operatorname{Ker} B \geq 2$ . Если же  $|B| \neq 0$ , то  $\operatorname{Ker} AB = B^{-1} \operatorname{Ker} A$ , а значит,  $\dim \operatorname{Ker} AB \geq 2$ .

Предположим теперь, что  $\dim \operatorname{Ker}(AB - \lambda I) \geq 1$  при  $\lambda \neq 1$ , и докажем, что  $\dim \operatorname{Ker}(AB - \lambda I) \geq 2$ . Если  $(ABA - \lambda A)u = 0$ , то  $(AB - \lambda I)Au = 0$ , т.е.  $AU \subset \operatorname{Ker}(AB - \lambda I)$ , где  $U = \operatorname{Ker}(ABA - \lambda A)$ . Поэтому достаточно доказать, что  $\dim AU \geq 2$ . Так как  $\operatorname{Ker} A \subset U$ , то  $\dim AU = \dim U - \dim \operatorname{Ker} A$ . Матрица  $ABA$  кососимметрическая, поэтому числа  $\dim U$  и  $\dim \operatorname{Ker} A$  чётные, а значит, число  $\dim AU$  тоже чётное. Остаётся проверить, что  $\operatorname{Ker} A \neq U$ . Предположим, что из равенства  $(AB - \lambda I)Ax = 0$  следует, что  $Ax = 0$ . Тогда  $\operatorname{Im} A \cap \operatorname{Ker}(AB - \lambda I) = 0$ . С другой стороны, если  $(AB - \lambda I)x = 0$ , то  $x = A(\lambda^{-1} Bx) \in \operatorname{Im} A$ , т.е.  $\operatorname{Ker}(AB - \lambda I) \subset \operatorname{Im} A$ , причём  $\dim \operatorname{Ker}(AB - \lambda I) \geq 1$ . Приходим к противоречию.

**4.1.** Корни многочлена  $p(\lambda)$  таковы, что если  $z$  — корень, то и  $1/z = \bar{z}/z\bar{z} = \bar{z}$  — тоже корень. Поэтому многочлен  $q(\lambda) = \lambda^n p(1/\lambda)$  имеет те же корни, что и  $p$  (и с теми же кратностями). Кроме того, свободный член многочлена  $p(\lambda)$  равен  $\pm 1$ , поэтому старшие коэффициенты многочленов  $p(\lambda)$  и  $q(\lambda)$  могут отличаться лишь знаком.

**4.2.** Пусть  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  — унитарная матрица с определителем 1. Тогда  $\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$ , т.е.  $\bar{a} = d$  и  $\bar{b} = -c$ . Кроме того,  $ad - bc = 1$ , т.е.  $|a|^2 + |b|^2 = 1$ .

**4.3.** Рассмотрим матрицу  $A^{-1} = \begin{pmatrix} A_{11}^T & A_{12}^T \\ A_{21}^T & A_{22}^T \end{pmatrix}$ . Согласно задаче 3.3  $\det A \det A_{22}^T = \det A_{11}$ , поэтому  $\det A_{22} = \det A_{11}$ .

**4.4.** а) Оператор  $A$  является поворотом на некоторый угол  $\varphi$ , поэтому  $\operatorname{tr} A = 1 + 2 \cos \varphi$  и  $\operatorname{tr}(A^2) = 1 + 2 \cos 2\varphi = 4 \cos^2 \varphi - 1$ .

б) Ясно, что  $\sum_{i < j} (a_{ij} - a_{ji})^2 = \sum_{i \neq j} a_{ij}^2 - 2 \sum_{i < j} a_{ij} a_{ji}$  и  $\operatorname{tr}(A^2) = \sum_i a_{ii}^2 + 2 \sum_{i < j} a_{ij} a_{ji}$ . С другой стороны, согласно задаче а)  $\operatorname{tr}(A^2) = \operatorname{tr}(A)^2 - 2 \operatorname{tr} A = (\operatorname{tr} A - 1)^2 - 1 = \left( \sum_i a_{ii} - 1 \right)^2 - 1$ . Поэтому  $\sum_{i < j} (a_{ij} - a_{ji})^2 + \left( \sum_i a_{ii} - 1 \right)^2 - 1 = \sum_{i \neq j} a_{ij}^2 + \sum_i a_{ii}^2 = 3$ .

**4.5.** Введём обозначение  $\frac{A}{B} = AB^{-1}$ ; тогда правило сокращения выглядит следующим образом:  $\frac{AB}{CB} = \frac{A}{C}$ . Если  $A^T = JA^{-1}J^{-1}$ , то

$$\begin{aligned} (A^\#)^T &= \frac{I - A^T}{I + A^T} = \frac{I - JA^{-1}J^{-1}}{I + JA^{-1}J^{-1}} = \frac{J(A - I)A^{-1}J^{-1}}{J(A + I)A^{-1}J^{-1}} = \\ &= \frac{J(A - I)}{J(A + I)} = -JA^\#J^{-1}. \end{aligned}$$

Если  $A^T = -JAJ^{-1}$ , то

$$(A^\#)^T = \frac{I - A^T}{I + A^T} = \frac{I + JAJ^{-1}}{I - JAJ^{-1}} = \frac{J(I + A)J^{-1}}{J(I - A)J^{-1}} = J(A^\#)^{-1}J^{-1}.$$

**4.6.** Рассмотрим матрицу

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & -1 & 0 & 0 & \dots & 0 \\ 1 & 1 & -2 & 0 & \dots & 0 \\ 1 & 1 & 1 & -3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & 1 & \dots & -(n-1) \end{pmatrix}.$$

Строки этой матрицы попарно ортогональны и квадраты их длин равны  $n, 1 \cdot 2, 2 \cdot 3, 3 \cdot 4, \dots$ . Пусть  $D = \operatorname{diag}(d_1, d_2, \dots, d_n)$ , где  $d_1 = \sqrt{n}, d_2 = \sqrt{2}, d_3 = \sqrt{6}, \dots$ . Тогда  $A = DH$  — требуемая матрица.

**4.7.** Так как все собственные значения оператора  $A$  по модулю равны 1, достаточно проверить, что оператор  $A$  унитарно диагонализуем. Докажем сначала, что оператор  $A$  диагонализуем. Предположим, что жорданова нормальная форма оператора  $A$  имеет клетку порядка не менее 2. Тогда существуют такие векторы  $e_1$  и  $e_2$ , что  $Ae_1 = \lambda e_1$  и  $Ae_2 = \lambda e_2 + e_1$ ; можно считать, что  $\|e_1\| = 1$ . Рассмотрим вектор  $x = e_2 - (e_1, e_2)e_1$ . Легко проверить, что  $x \perp e_1$  и  $Ax = \lambda x + e_1$ .

Поэтому  $\|Ax\|^2 = \|\lambda x\|^2 + \|e_1\|^2 = \|x\|^2 + 1$ , а значит,  $\|Ax\| > \|x\|$ . Получено противоречие.

Остаётся доказать, что если  $Ax = \lambda x$  и  $Ay = \mu y$ , причём  $\lambda \neq \mu$ , то  $(x, y) = 0$ . Предположим, что  $(x, y) \neq 0$ . Заменяя вектор  $x$  на  $\alpha x$ , где  $\alpha$  — подходящее комплексное число, можно считать, что  $\operatorname{Re}[(\lambda\bar{\mu} - 1)(x, y)] > 0$ . Тогда

$$\begin{aligned} \|A(x+y)\|^2 - \|x+y\|^2 &= \|\lambda x + \mu y\|^2 - \|x+y\|^2 = \\ &= 2 \operatorname{Re}[(\lambda\bar{\mu} - 1)(x, y)] > 0, \end{aligned}$$

т.е.  $\|Az\| > \|z\|$ , где  $z = x+y$ . Получено противоречие.

**4.8.** Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения оператора  $U$ ,  $e_1, \dots, e_n$  — соответствующие им попарно ортогональные собственные векторы. Тогда  $x = \sum x_i e_i$  и  $Bx = \sum \lambda_i x_i e_i$ , а значит,  $0 = (Bx, x) = \sum \lambda_i |x_i|^2$ . Пусть  $t_i = |x_i|^2 / \|x\|^2$ . Так как  $t_i \geq 0$ ,  $\sum t_i = 1$  и  $\sum t_i \lambda_i = 0$ , то начало координат лежит внутри выпуклой оболочки точек  $\lambda_1, \dots, \lambda_n$ .

**4.9.** Матрица  $I - A^*A$  неотрицательно определённая, поскольку  $((I - A^*A)x, x) = (x, x) - (Ax, Ax) \geq 0$ . Матрица  $I - AA^*$  тоже неотрицательно определённая, поэтому согласно теореме 4.1.4.1 существуют неотрицательно определённые матрицы  $S$  и  $T$ , для которых  $S^2 = I - AA^*$  и  $T^2 = I - A^*A$ . Покажем, что матрица  $U = \begin{pmatrix} A & S \\ T & -A^* \end{pmatrix}$  унитарна. Ясно, что  $U^* = \begin{pmatrix} A^* & T \\ S & -A \end{pmatrix}$ , поэтому  $UU^* = \begin{pmatrix} AA^* + S^2 & AT - SA \\ TA^* - A^*S & T^2 + A^*A \end{pmatrix}$ . Остаётся доказать, что  $AT = SA$ . Ясно, что  $AT^2 = S^2A$ , поэтому по индукции получаем, что  $AT^{2n} = S^{2n}A$  для  $n = 0, 1, 2, \dots$ , а значит,  $Ap(T^2) = p(S^2)A$  для любого многочлена  $p$ . Если мы выберем многочлен  $p$  степени  $2n$  так, что  $p(\lambda_i^2) = \lambda_i$  для каждого числа  $\lambda_i$ , которое является собственным значением матрицы  $T$  или  $S$ , то получим  $p(T^2) = T$  и  $p(S^2) = S$ . Таким образом,  $AT = SA$ .

**Замечание.** Для вещественных матриц другое доказательство приведено в решении задачи 6.3.

**5.1.** Если  $A^* = AU$ , где  $U$  — унитарная матрица, то  $A = U^*A^*$ , а значит,  $UA = UU^*A^* = A^*$ . Следовательно,  $AU = UA$  и  $A^*A = AUA = AAU = AA^*$ .

Если оператор  $A$  нормален, то для него существует ортонормированный собственный базис  $e_1, \dots, e_n$ . При этом  $Ae_i = \lambda_i e_i$  и  $A^*e_i = \bar{\lambda}_i e_i$ . Пусть  $U = \operatorname{diag}(d_1, \dots, d_n)$ , где  $d_i = \bar{\lambda}_i / \lambda_i$  при  $\lambda_i \neq 0$  и  $d_i = 1$  при  $\lambda_i = 0$ . Тогда  $A^* = AU$ .

**5.2.** Пусть  $A = U\Lambda U^*$ , где  $U$  — унитарная матрица,  $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ . Тогда

$$B = UDU^*, \quad \text{где } D = \operatorname{diag}(\pm\sqrt{\lambda_1}, \dots, \pm\sqrt{\lambda_n}).$$

**5.3.** По условию  $\operatorname{Im} B \subset (\operatorname{Im} A)^\perp = \operatorname{Ker} A^*$ , т.е.  $A^*B = 0$ . Аналогично  $B^*A = 0$ . Поэтому  $(A^* + B^*)(A + B) = A^*A + B^*B$ . Учитывая, что для нормального оператора  $\operatorname{Ker} A = \operatorname{Ker} A^*$  и  $\operatorname{Im} A = \operatorname{Im} A^*$ , аналогично получаем  $(A + B)(A^* + B^*) = AA^* + BB^*$ .

**5.4.** Рассмотрим ортонормированный базис, в котором оператор  $A$  диагонален. Можно считать, что  $A = \operatorname{diag}(d_1, \dots, d_k, 0, \dots, 0)$ , причём  $d_i \neq 0$ . Тогда  $S =$

$\text{diag}(|d_1|, \dots, |d_k|, 0, \dots, 0)$ . Пусть  $D = \text{diag}(d_1, \dots, d_k)$  и  $D_+ = \text{diag}(|d_1|, \dots, |d_k|)$ . Равенство

$$\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} D_+ & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix} = \begin{pmatrix} D_+ U_1 & D_+ U_2 \\ 0 & 0 \end{pmatrix}$$

выполняется, только если  $U_1 = D_+^{-1} D = \text{diag}(e^{i\varphi_1}, \dots, e^{i\varphi_k})$ , поэтому матрица  $\begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix}$  унитарна, только если  $U_2 = 0$  и  $U_3 = 0$ . Ясно, что  $\begin{pmatrix} D_+ & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_1 & 0 \\ 0 & U_4 \end{pmatrix} = \begin{pmatrix} U_1 & 0 \\ 0 & U_4 \end{pmatrix} \begin{pmatrix} D_+ & 0 \\ 0 & 0 \end{pmatrix}$ .

**5.5.** Матрица  $X$  нормальна тогда и только тогда, когда  $\text{tr}(X^* X) = \sum |\lambda_i|^2$ , где  $\lambda_i$  — собственные значения матрицы  $X$  (теорема 6.2.1.1).

Кроме того, собственные значения матриц  $X = AB$  и  $Y = BA$  совпадают (теорема 3.1.6.1). Остаётся проверить, что  $\text{tr}(X^* X) = \text{tr}(Y^* Y)$ . Это легко сделать, если учесть, что  $A^* A = A A^*$  и  $B^* B = B B^*$ .

**5.6.** Рассмотрим матрицу  $N = \begin{pmatrix} A & A^* \\ A^* & A \end{pmatrix}$ . Легко проверить, что

$$N^* N = \begin{pmatrix} A^* A + A A^* & A^2 + A^{*2} \\ A^{*2} + A^2 & A^* A + A A^* \end{pmatrix} = N N^*.$$

**6.1.** Матрицу  $(A + \lambda B)^n$  можно представить в виде  $(A + \lambda B)^n = A^n + \lambda C_1 + \dots + \lambda^{n-1} C_{n-1} + \lambda^n B^n$ , где элементы матриц  $C_1, \dots, C_{n-1}$  не зависят от  $\lambda$ . Пусть  $a, c_1, \dots, c_{n-1}, b$  — элементы матриц  $A^n, C_1, \dots, C_{n-1}, B^n$ , стоящие на месте  $(i, j)$ . Тогда  $a + \lambda c_1 + \dots + \lambda^{n-1} c_{n-1} + \lambda^n b = 0$  для  $n+1$  различных значений  $\lambda$ . Получена система из  $n+1$  уравнения для  $n+1$  неизвестных  $a, c_1, \dots, c_{n-1}, b$ . Определитель этой системы является определителем Вандермонда, поэтому он не равен нулю. Следовательно, полученная система уравнений имеет лишь нулевое решение. В частности,  $a = b = 0$ , а значит,  $A^n = B^n = 0$ .

**6.2.** Пусть  $(A^2 - A)^k = 0$ . Раскрывая скобки, для некоторого многочлена  $q$  с целыми коэффициентами получаем  $A^k = A^{k+1} q(A) = A^k [A q(A)] = A^{k+1} q(A) [A q(A)] = A^{k+1} q(A) [A q(A)] = A^k [A q(A)]^2 = \dots = A^k [A q(A)]^k = A^{2k} q(A)^k$ . Если  $A^k \neq 0$ , то  $X = A^k q(A)^k \neq 0$  и  $X^2 = [A^{2k} q(A)^k] q(A)^k = A^k q(A)^k = X$ , т.е.  $p(t) = t^k q(t)^k$  — искомым многочлен.

**6.3.** Пусть  $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$  и  $C = \lambda A + \mu B$ . Легко проверить, что  $C^3 = 0$ .

Матрицы  $A$  и  $B$  нельзя одновременно привести к треугольному виду, так как матрица  $AB$  не нильпотентна.

**7.1.** а) Ясно, что  $(P_1 + P_2)^2 = P_1^2 + P_2^2 + (P_1 P_2 + P_2 P_1) = P_1 + P_2 + (P_1 P_2 + P_2 P_1)$ . Если  $P_1 P_2 + P_2 P_1 = 0$ , то  $2P_1 P_2 = P_1 P_2 - P_2 P_1 = P_1^2 P_2 - P_2 P_1^2 = P_1(P_1 P_2 + P_2 P_1) - (P_1 P_2 + P_2 P_1)P_1 = 0$ .

б) Так как  $I - (P_1 - P_2) = (I - P_1) + P_2$ , то оператор  $P_1 - P_2$  является проектором тогда и только тогда, когда  $(I - P_1)P_2 = P_2(I - P_1) = 0$ , т.е.  $P_1 P_2 = P_2 P_1 = P_2$ .

**7.2.** Если  $P$  — матрица порядка 2 и  $\text{rk } P = 1$ , то  $\text{tr } P = 1$  и  $\det P = 0$  (если  $\text{rk } P \neq 1$ , то  $P = I$  или  $P = 0$ ). Поэтому  $P = \frac{1}{2} \begin{pmatrix} 1+a & b \\ c & 1-a \end{pmatrix}$ , где  $a^2 + bc = 1$ .

Ясно также, что если  $\operatorname{tr} P = 1$  и  $\det P = 0$ , то согласно теореме Гамильтона–Кэли  $P^2 - P = P^2 - (\operatorname{tr} P)P + \det P = 0$ .

**7.3.** Так как  $\operatorname{Im}(I - A) = \operatorname{Ker}((I - A)^*)^\perp$ , то любой вектор  $x$  можно представить в виде  $x = x_1 + x_2$ , где  $x_1 \in \operatorname{Im}(I - A)$  и  $x_2 \in \operatorname{Im}(I - A^*)$ . Достаточно отдельно рассмотреть  $x_1$  и  $x_2$ . Вектор  $x_1$  имеет вид  $y - Ay$ , поэтому

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} A^i x_1 \right| = \left| \frac{1}{n} (y - A^n y) \right| \leq \frac{2|y|}{n} \rightarrow 0.$$

Так как  $x_2 \in \operatorname{Ker}(I - A^*)$ , то  $x_2 = A^* x_2 = A^{-1} x_2$ , т.е.  $Ax_2 = x_2$ . Поэтому

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} A^i x_2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} x_2 = x_2.$$

Остаётся заметить, что  $\operatorname{Ker}(I - A^*) = \operatorname{Ker}(A - I)$ .

**7.4.** (б)  $\Rightarrow$  (а). Достаточно домножить равенство  $A_1 + \dots + A_k = I$  на  $A_i$ .

(а)  $\Rightarrow$  (в).  $A_i$  — проектор, поэтому  $\operatorname{rk} A_i = \operatorname{tr} A_i$ . Следовательно,  $\sum \operatorname{rk} A_i = \sum \operatorname{tr} A_i = \operatorname{tr} (\sum A_i) = n$ .

(в)  $\Rightarrow$  (б). Так как  $\sum A_i = I$ , то  $\operatorname{Im} A_1 + \dots + \operatorname{Im} A_k = V$ . Но  $\operatorname{rk} A_1 + \dots + \operatorname{rk} A_k = \dim V$ , поэтому  $V = \operatorname{Im} A_1 \oplus \dots \oplus \operatorname{Im} A_k$ . Для любого  $x \in V$  справедливо равенство  $A_j x = (A_1 + \dots + A_k) A_j x = A_1 A_j x + \dots + A_k A_j x$ , причём  $A_i A_j x \in \operatorname{Im} A_i$  и  $A_j x \in \operatorname{Im} A_j$ . Следовательно,  $A_i A_j = 0$  при  $i \neq j$  и  $A_j^2 = A_j$ .

**7.5.** Матрица  $A(I - A)$  неотрицательно определённая, поэтому согласно теореме 4.1.4.1 существует единственная неотрицательно определённая матрица  $S$ , для которой  $S^2 = A(I - A)$ . Конструкция матрицы  $S$  показывает, что  $AS = SA$ . Используя это свойство, легко проверить, что эрмитова матрица  $\begin{pmatrix} A & S \\ S & I - A \end{pmatrix}$  является проектором.

**8.1.** Если инволюция  $A$  в двумерном пространстве не равна  $\pm I$ , то её матрица в некотором базисе имеет вид  $\operatorname{diag}(1, -1)$ . Поэтому  $\operatorname{tr} A = 0$  и  $\det A = -1$ , т.е.  $A = \begin{pmatrix} x & y \\ z & -x \end{pmatrix}$ , где  $x^2 + yz = 1$ . Согласно теореме Гамильтона–Кэли такая матрица  $A$  удовлетворяет уравнению  $A^2 - (\operatorname{tr} A)A + (\det A)I = 0$ , т.е.  $A^2 = I$ .

**8.2.** Требуется доказать, что  $H^2 + \bar{S}S = I$  и  $SH = \overline{HS}$ . Для доказательства этих равенств достаточно воспользоваться перестановочностью матриц  $\bar{A}A + I$  и  $\bar{A}A - I$ , а также равенствами  $A(\bar{A}A \pm I) = (A\bar{A} \pm I)A$  и  $A(\bar{A}A \pm I)^{-1} = (\bar{A}A \pm I)^{-1}A$ .

## Глава 5.

# Полилинейная алгебра

### 5.1. Полилинейные отображения. Тензорные произведения

#### 5.1.1. Определения

Пусть  $V, V_1, \dots, V_k$  — линейные пространства;  $\dim V_i = n_i$ . Отображение  $f: V_1 \times \dots \times V_k \rightarrow V$  называют *полилинейным* (или *k-линейным*), если оно линейно по каждой из  $k$  переменных при фиксированных остальных переменных.

Выберем в пространствах  $V_1, \dots, V_k$  базисы  $\{e_{1i}\}, \dots, \{e_{kj}\}$ . Если  $f$  — полилинейное отображение, то

$$f(\sum x_{1i}e_{1i}, \dots, \sum x_{kj}e_{kj}) = \sum x_{1i} \dots x_{kj} f(e_{1i}, \dots, e_{kj}).$$

Отображение  $f$  задаётся своими  $n_1 \dots n_k$  значениями  $f(e_{1i}, \dots, e_{kj}) \in V$ , причём эти значения могут быть любыми. Рассмотрим пространство  $V_1 \otimes \dots \otimes V_k$  размерности  $n_1 \dots n_k$ , базисные элементы которого обозначим  $e_{1i} \otimes \dots \otimes e_{kj}$ . Рассмотрим, далее, отображение  $p: V_1 \times \dots \times V_k \rightarrow V_1 \otimes \dots \otimes V_k$ , заданное формулой  $p(\sum x_{1i}e_{1i}, \dots, \sum x_{kj}e_{kj}) = \sum x_{1i} \dots x_{kj} e_{1i} \otimes \dots \otimes e_{kj}$ ; элемент  $p(v_1, \dots, v_k)$  обозначим  $v_1 \otimes \dots \otimes v_k$ . Каждому полилинейному отображению  $f$  соответствует линейное отображение  $\tilde{f}: V_1 \otimes \dots \otimes V_k \rightarrow V$ , где  $\tilde{f}(e_{1i} \otimes \dots \otimes e_{kj}) = f(e_{1i}, \dots, e_{kj})$ ; соответствие между полилинейными отображениями  $f$  и линейными отображениями  $\tilde{f}$  взаимно однозначное. Легко проверить также, что  $\tilde{f}(v_1 \otimes \dots \otimes v_k) = f(v_1, \dots, v_k)$  для любых векторов  $v_i \in V_i$ .

Элементу  $v_1 \otimes \dots \otimes v_k$  можно сопоставить полилинейную функцию  $f(w_1, \dots, w_k) = w_1(v_1) \dots w_k(v_k)$  на  $V_1^* \times \dots \times V_k^*$ . Продолжив это отображение по линейности, получим изоморфизм пространства  $V_1 \otimes \dots \otimes V_k$  на пространство полилинейных функций на  $V_1^* \times \dots \times V_k^*$ . Таким образом можно получить инвариантное определение пространства  $V_1 \otimes \dots \otimes V_k$ ; это пространство называют *тензорным произведением* пространств  $V_1, \dots, V_k$ .

Линейное отображение  $V_1^* \otimes \dots \otimes V_k^* \rightarrow (V_1 \otimes \dots \otimes V_k)^*$ , переводящее  $f_1 \otimes \dots \otimes f_k \in V_1^* \otimes \dots \otimes V_k^*$  в линейную функцию  $f(v_1 \otimes \dots \otimes v_k) = f_1(v_1) \dots f_k(v_k)$ , является каноническим изоморфизмом.

### 5.1.2. Билинейные формы

Билинейная функция  $f : V_1 \times V_2 \rightarrow \mathbb{R}$  задаётся матрицей  $B = (b_{ij})$  размера  $n \times m$ , где  $n = \dim V_1$  и  $m = \dim V_2$ . А именно, выберем в  $V_1$  базис  $e_1, \dots, e_n$ , в  $V_2$  — базис  $\varepsilon_1, \dots, \varepsilon_m$  и положим  $b_{ij} = f(e_i, \varepsilon_j)$ . Тогда  $f(\sum x_i e_i, \sum y_j \varepsilon_j) = \sum x_i y_j b_{ij}$ . Таким образом,  $f(x, y) = x^T B y$ , где векторы  $x$  и  $y$  представлены как столбцы координат.

В случае, когда  $V_1 = V_2$ , билинейную функцию называют *билинейной формой*. Для билинейной формы мы предполагаем, что в  $V_1$  и в  $V_2$  выбран один и тот же базис. Билинейную форму  $f$  называют *симметрической*, если  $f(x, y) = f(y, x)$  для всех  $x$  и  $y$ ; *кососимметрической* — если  $f(x, y) = -f(y, x)$  для всех  $x$  и  $y$ . Симметрическая форма задаётся симметрической матрицей, а кососимметрическая — кососимметрической.

Билинейную форму называют *невырожденной*, если она задаётся невырожденной матрицей. Это свойство эквивалентно следующему: для любого  $y \neq 0$  можно выбрать  $x$  так, что  $f(x, y) \neq 0$ . Действительно, оба эти условия эквивалентны тому, что если  $y \neq 0$ , то  $B y \neq 0$ .

Если задана невырожденная билинейная форма  $g(x, y) = x^T G y$ , то оператору  $A$  в пространстве  $V_1 = V_2$  можно сопоставить оператор  $A^*$ , для которого  $g(Ax, y) = g(x, A^*y)$  для всех  $x$  и  $y$ . Это равенство полностью определяет оператор  $A^*$ . Действительно, если  $(Ax)^T G y \equiv x^T G A^* y$ , то  $A^T G = G A^*$  и  $A^* = G^{-1} A^T G$ .

**Теорема 5.1.2.1.** *Равенство  $(A^*)^* = A$  выполняется для всех операторов  $A$  тогда и только тогда, когда форма симметрическая или кососимметрическая.*

**Доказательство.** Ясно, что  $(A^*)^* = G^{-1}(G^{-1}A^T G)^T G = G^{-1}G^T A(G^{-1})^T G = G^{-1}G^T A(G^{-1}G^T)^{-1}$ . Поэтому равенство  $(A^*)^* = A$  эквивалентно тому, что матрица  $G^{-1}G^T$  коммутирует с  $A$ .



Матрица  $G^{-1}G^T$  коммутирует со всеми матрицами тогда и только тогда, когда  $G^{-1}G^T = \lambda I$  (задача 1.2), т.е.  $G^T = \lambda G$ . В таком случае  $G = (G^T)^T = \lambda G^T = \lambda^2 G$ , поэтому  $\lambda = \pm 1$ .  $\square$

**Следствие.** Пусть  $g(x, y)$  — невырожденная билинейная форма, причём из равенства  $g(x, y) = g(u, v)$  следует равенство  $g(y, x) = g(v, u)$ . Тогда форма  $g$  симметрическая или кососимметрическая.

**Доказательство.** По определению  $g(Ay, x) = g(y, A^*x)$ , поэтому  $g(x, Ay) = g(A^*x, y) = g(x, (A^*)^*y)$ . Таким образом,  $(A^*)^* = A$  для любого оператора  $A$ .  $\square$

### 5.1.3. Изоморфизм $V^* \otimes W \rightarrow \text{Hom}(V, W)$

**Теорема 5.1.3.1.** Пусть  $\text{Hom}(V, W)$  — пространство линейных отображений  $V$  в  $W$ . Тогда существует канонический изоморфизм  $\alpha: V^* \otimes W \rightarrow \text{Hom}(V, W)$ .

**Доказательство.** Пусть  $\{e_i\}$  и  $\{\varepsilon_j\}$  — базисы  $V$  и  $W$ . Положим  $\alpha(e_i^* \otimes \varepsilon_j)v = e_i^*(v)\varepsilon_j = v_i\varepsilon_j$  и продолжим отображение  $\alpha$  по линейности. Если  $v \in V$ ,  $f \in V^*$  и  $w \in W$ , то  $\alpha(f \otimes w)v = f(v)w$ , поэтому  $\alpha$  можно определить инвариантно.

Пусть  $Ae_p = \sum_q a_{qp}\varepsilon_q$ , тогда  $A(\sum_p v_p e_p) = \sum_{p,q} a_{qp}v_p\varepsilon_q$ . Поэтому отображению  $\alpha(e_i^* \otimes \varepsilon_j)$  соответствует матрица  $(a_{qp})$ , где  $a_{qp} = \delta_{qj}\delta_{pi}$ . Такие матрицы образуют базис пространства  $\text{Hom}(V, W)$ . Ясно также, что размерности пространств  $V^* \otimes W$  и  $\text{Hom}(V, W)$  равны.  $\square$

**Теорема 5.1.3.2.** Пусть  $V$  — линейное пространство над полем  $K$ . Рассмотрим отображение свёртки  $\varepsilon: V^* \otimes V \rightarrow K$ , заданное формулой  $\varepsilon(x^* \otimes y) = x^*(y)$  и продолженное по линейности. Тогда  $\text{tr } A = \varepsilon\alpha^{-1}(A)$  для любого линейного оператора  $A$  в пространстве  $V$ .

**Доказательство.** Выберем в пространстве  $V$  базис. Доказательство достаточно провести для матриц  $I_{ij} = (a_{qp})$ , где  $a_{qp} = \delta_{qj}\delta_{pi}$ . Ясно, что  $\text{tr } I_{ij} = \delta_{ij}$  и  $\varepsilon\alpha^{-1}(I_{ij}) = \varepsilon(e_i^* \otimes e_j) = e_i^*(e_j) = \delta_{ij}$ .  $\square$

**Замечание.** Пространство  $V^* \otimes V$  и отображения  $\alpha$  и  $\varepsilon$  определены инвариантно, поэтому теорема 5.1.3.2 даёт инвариантное определение следа оператора.

**Теорема 5.1.3.3.** Имеет место канонический изоморфизм

$$\text{Hom}(V \otimes W, U) \cong \text{Hom}(V, \text{Hom}(W, U)).$$

**Доказательство.** Согласно теореме 5.1.3.1

$$\begin{aligned}\operatorname{Hom}(V \otimes W, U) &\cong (V \otimes W)^* \otimes U, \\ \operatorname{Hom}(V, \operatorname{Hom}(W, U)) &\cong V^* \otimes (W^* \otimes U).\end{aligned}$$

Остаётся заметить, что тензорное произведение ассоциативно и имеет место канонический изоморфизм  $V^* \otimes W^* \cong (V \otimes W)^*$ .  $\square$

#### 5.1.4. Валентность тензора

*Тензором типа  $(p, q)$*  на пространстве  $V$  называют элемент пространства  $T_p^q(V) = V^* \otimes \dots \otimes V^* \otimes V \otimes \dots \otimes V$  ( $p$  экземпляров пространства  $V^*$  и  $q$  экземпляров пространства  $V$ ), изоморфного пространству линейных функций на  $V \times \dots \times V \times V^* \times \dots \times V^*$ . Число  $p$  называют *ковариантной валентностью* тензора,  $q$  — *контравариантной валентностью*,  $p + q$  — *общей валентностью*. Векторы являются тензорами типа  $(0, 1)$ , а ко-векторы — тензорами типа  $(1, 0)$ .

Пусть в пространстве  $V$  выбран базис  $e_1, \dots, e_n$  и  $e_1^*, \dots, e_n^*$  — двойственный ему базис пространства  $V^*$ . Каждый тензор типа  $(p, q)$  имеет вид

$$\sum T_{i_1 \dots i_p}^{j_1 \dots j_q} e_{i_1}^* \otimes \dots \otimes e_{i_p}^* \otimes e_{j_1} \otimes \dots \otimes e_{j_q}; \quad (5.1)$$

числа  $T_{i_1 \dots i_p}^{j_1 \dots j_q}$  называют *координатами* (или *компонентами*) тензора относительно базиса  $e_1, \dots, e_n$ .

Выясним, как изменяются координаты тензора при переходе к другому базису. Пусть  $\varepsilon_j = Ae_j = \sum a_{ij}e_i$  и  $\varepsilon_j^* = \sum b_{ij}e_i^*$ . Легко проверить, что  $B = (A^T)^{-1}$  (см. §2.1.4). Введём обозначения  $a_j^i = a_{ij}$  и  $b_i^j = b_{ij}$ ; тензор (5.1) для краткости обозначим  $\sum T_\alpha^\beta e_\alpha^* \otimes e_\beta$ . Тогда  $\sum T_\alpha^\beta e_\alpha^* \otimes e_\beta = \sum S_\mu^\nu \varepsilon_\mu^* \otimes \varepsilon_\nu = \sum S_\mu^\nu b_\alpha^\mu a_\nu^\beta e_\alpha^* \otimes e_\beta$ , т.е.

$$T_{i_1 \dots i_p}^{j_1 \dots j_q} = b_{i_1}^{l_1} \dots b_{i_p}^{l_p} a_{k_1}^{j_1} \dots a_{k_q}^{j_q} S_{l_1 \dots l_p}^{k_1 \dots k_q} \quad (5.2)$$

(по одинаковым индексам производится суммирование). Формула (5.2) связывает координаты  $S$  тензора в базисе  $\{\varepsilon_i\}$  и координаты  $T$  в базисе  $\{e_i\}$ . Несколько иную интерпретацию закона преобразования координат тензоров типа  $(0, 1)$  и  $(1, 0)$  см. в §2.3.1.

Для тензоров типа  $(1, 1)$ , которые можно отождествить с линейными операторами, определено отображение свёртки, переводящее  $v^* \otimes w$  в  $v^*(w)$ . При этом отображении оператору сопоставляется его след (см. теорему 5.1.3.2).

Общее отображение свёртки определяется следующим образом. Пусть  $1 \leq i \leq p$  и  $1 \leq j \leq q$ . Рассмотрим линейное отображение  $T_p^q(V) \rightarrow T_{p-1}^{q-1}(V)$ , переводящее  $f_1 \otimes \dots \otimes f_p \otimes v_1 \otimes \dots \otimes v_q$  в  $f_i(v_j)f_i \otimes v_j$ , где  $f_i$  и  $v_j$  — тензорные произведения элементов  $f_1, \dots, f_p$  и  $v_1, \dots, v_q$  без  $f_i$  и  $v_j$  соответственно. Это отображение называют *свёрткой тензора* по  $i$ -му нижнему индексу и  $j$ -му верхнему индексу.

### 5.1.5. Тензорное произведение отображений

Линейные отображения  $A_i: V_i \rightarrow W_i$  ( $i = 1, \dots, k$ ) индуцируют линейное отображение  $A_1 \otimes \dots \otimes A_k: V_1 \otimes \dots \otimes V_k \rightarrow W_1 \otimes \dots \otimes W_k$ , переводящее элемент  $e_{1i} \otimes \dots \otimes e_{kj}$  в  $A_1 e_{1i} \otimes \dots \otimes A_k e_{kj}$ . Легко проверить, что при этом элемент  $v_1 \otimes \dots \otimes v_k$  переходит в  $A_1 v_1 \otimes \dots \otimes A_k v_k$ . Отображение  $A_1 \otimes \dots \otimes A_k$  называют *тензорным произведением отображений*  $A_1, \dots, A_k$ , или *произведением Кронекера*.

Если  $Ae_j = \sum a_{ij}\varepsilon_i$  и  $Be'_q = \sum b_{pq}\varepsilon'_p$ , то  $A \otimes B(e_j \otimes e'_q) = \sum a_{ij}b_{pq}\varepsilon_i \otimes \varepsilon'_p$ . Поэтому, упорядочив соответствующим образом базисы  $e_j \otimes e'_q$  и  $\varepsilon_i \otimes \varepsilon'_p$ , матрицу  $A \otimes B$  можно записать в виде

$$\begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \dots & \dots & \dots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}$$

или в виде

$$\begin{pmatrix} b_{11}A & \dots & a_{1l}A \\ \dots & \dots & \dots \\ b_{k1}B & \dots & b_{kl}B \end{pmatrix};$$

матрицу  $A \otimes B$  называют *кронекеровским произведением* матриц  $A$  и  $B$ .

**Теорема 5.1.5.1.** Пусть собственные значения квадратных матриц  $A$  и  $B$  равны  $\alpha_1, \dots, \alpha_m$  и  $\beta_1, \dots, \beta_n$ . Тогда собственные значения матрицы  $A \otimes B$  равны  $\alpha_i\beta_j$ , а собственные значения матрицы  $A \otimes I_n + I_m \otimes B$  равны  $\alpha_i + \beta_j$ .

**Доказательство.** Приведём матрицы  $A$  и  $B$  к жордановой нормальной форме (достаточно привести к треугольному виду) и в тензорном произведении рассмотрим базис, являющийся произведением этих базисов. Остаётся заметить, что  $J_p(\alpha) \otimes J_q(\beta)$  — верхняя треугольная матрица с диагональю  $(\alpha\beta, \dots, \alpha\beta)$ , а  $J_p(\alpha) \otimes I_q$  и  $I_p \otimes J_q(\beta)$  — верхние треугольные матрицы с диагоналями  $(\alpha, \dots, \alpha)$  и  $(\beta, \dots, \beta)$ .  $\square$

**Следствие.**  $\text{tr}(A \otimes B) = (\text{tr } A)(\text{tr } B)$  и  $\det(A \otimes B) = (\det A)^n(\det B)^m$ .

### 5.1.6. Матричные уравнения

Тензорное произведение операторов можно применить для решения матричных уравнений вида

$$A_1XB_1 + \cdots + A_rXB_r = C, \quad (1)$$

где

$$V^k \xrightarrow{B_i} V^l \xrightarrow{X} V^m \xrightarrow{A_i} V^n.$$

Покажем, что при естественных отождествлениях  $\text{Hom}(V^l, V^m) = (V^l)^* \otimes V^m$  и  $\text{Hom}(V^k, V^n) = (V^k)^* \otimes V^n$  отображение  $X \mapsto A_iXB_i$  отождествляется с  $B_i^T \otimes A_i$ , а значит, уравнение (1) можно переписать в виде

$$(B_1^T \otimes A_1 + \cdots + B_r^T \otimes A_r)X = C,$$

где  $X \in (V^l)^* \otimes V^m$  и  $C \in (V^k)^* \otimes V^n$  (т.е. матрицы  $X$  и  $C$  мы рассматриваем здесь как векторы). В самом деле, если элемент  $f \otimes v \in (V^l)^* \otimes V^m$  соответствует отображению  $Xx = (f \otimes v)x = f(x)v$ , то элемент  $B^T f \otimes Av \in (V^k)^* \otimes V^n$  соответствует отображению  $(B^T f \otimes Av)y = f(By)Av = AXBy$ .

**Теорема 5.1.6.1.** Пусть  $A$  и  $B$  — квадратные матрицы порядков  $m$  и  $n$ ,  $X$  и  $C$  — матрицы размером  $m \times n$ . (Матрицы  $A$ ,  $B$ ,  $C$  заданы, а матрица  $X$  неизвестна.) Уравнение  $AX - XB = C$  имеет единственное решение тогда и только тогда, когда матрицы  $A$  и  $B$  не имеют общих собственных значений.

**Доказательство.** Данное уравнение можно переписать в виде  $(I_n \otimes A - B^T \otimes I_m)X = C$ . Собственные значения оператора  $I_n \otimes A - B^T \otimes I_m$  равны  $\alpha_i - \beta_j$ , где  $\alpha_i$  — собственные значения оператора  $A$ ,  $\beta_j$  — собственные значения оператора  $B^T$  (т.е. собственные значения оператора  $B$ ). Ясно, что оператор  $I_n \otimes A - B^T \otimes I_m$  невырожден тогда и только тогда, когда  $\alpha_i - \beta_j \neq 0$  для всех  $i$  и  $j$ .  $\square$

В том случае когда матрицы  $A$  и  $B$  не имеют общих собственных значений, решение уравнения  $AX - XB = C$  можно найти в явном виде. Введём для этого обозначения  $C_r = \sum_{k=1}^r A^{r-k}CB^{k-1}$ ,  $C_0 = 0$ .

**Теорема 5.1.6.2.** [Ш] Предположим, что матрицы  $A$  и  $B$  не имеют общих собственных значений. Пусть  $p_A(t) = t^l + p_1t^{l-1} + \cdots + p_l$  — минимальный многочлен матрицы  $A$ . Тогда решение уравнения  $AX - XB = C$  можно найти по формуле

$$X = -(C_l + p_1C_{l-1} + \cdots + p_{l-1}C_1)(p_A(B))^{-1}.$$

**Доказательство.** Поскольку матрицы  $A$  и  $B$  не имеют общих собственных значений, матрица  $p_A(B)$  обратима (задача 3.3). Ясно также, что одна из матриц  $A$  и  $B$  обратима. Действительно, если одна матрица имеет собственное значение 0, то у другой матрицы нет такого собственного значения. Пусть для определённости матрица  $A$  обратима. Тогда  $X = A^{-1}C + A^{-1}CXB = A^{-1}C + A^{-2}CB + A^{-2}CXB^2 + \dots = \sum_{k=1}^r A^{-k}CB^k + A^{-r}XB^r$ . Следовательно,  $A^rX - XB^r = \sum_{k=1}^r A^{r-k}CB^{r-k} = \sum_{k=1}^r A^{r-k}CB^{k-1}$ . (В случае, когда обратима матрица  $B$ , это равенство доказывается аналогично).

Сложим равенства

$$\begin{aligned} p_l(A^0X - XB^0) &= p_lC_0, \\ p_{l-1}(AX - XB) &= p_{l-1}C_1, \\ &\dots \dots \\ p_1(A^{l-1}X - XB^{l-1}) &= p_1C_{l-1}, \\ A^lX - XB^l &= C_l. \end{aligned}$$

В результате получим

$$p_A(A)X - Xp_A(B) = C_l + p_1C_{l-1} + \dots + p_{l-1}C_1.$$

Учитывая, что  $p_A(A) = 0$ , а матрица  $p_A(B)$  обратима, приходим к требуемому равенству.  $\square$

**Теорема 5.1.6.3.** Пусть  $A$  и  $B$  — квадратные матрицы одного порядка. Уравнение  $AX - XB = \lambda X$  имеет ненулевое решение тогда и только тогда, когда  $\lambda = \alpha_i - \beta_j$ , где  $\alpha_i$  и  $\beta_j$  — собственные значения матриц  $A$  и  $B$ .

**Доказательство.** Уравнение  $(I \otimes A - B^T \otimes I)X = \lambda X$  имеет ненулевое решение тогда и только тогда, когда  $\lambda$  — собственное значение оператора  $I \otimes A - B^T \otimes I$ , т.е.  $\lambda = \alpha_i - \beta_j$ .  $\square$

### 5.1.7. Подпространство, соответствующее полилинейной функции

Полилинейной функции  $f \in \text{Hom}(V \times \dots \times V, K) \cong (V^*)^{\otimes p}$  можно сопоставить подпространство  $W_f \subset V^*$ , порождённое ковекторами  $\xi$  вида  $\xi(x) = f(a_1, \dots, a_{i-1}, x, a_i, \dots, a_{p-1})$ , где векторы  $a_i, \dots, a_{p-1}$  и число  $i$  фиксированы.

**Теорема 5.1.7.1.**  $f \in W_f^{\otimes p}$ .

**Доказательство.** Пусть  $\varepsilon_1, \dots, \varepsilon_r$  — базис  $W_f$ ; дополним его до базиса  $\varepsilon_1, \dots, \varepsilon_n$  пространства  $V^*$ . Требуется доказать, что  $f = \sum f_{i_1 \dots i_p} \varepsilon_{i_1} \otimes \dots \otimes \varepsilon_{i_p}$ , где  $f_{i_1 \dots i_p} = 0$ , когда один из индексов  $i_1, \dots, i_p$  больше  $r$ . Пусть  $e_1, \dots, e_n$  — базис, двойственный к  $\varepsilon_1, \dots, \varepsilon_n$ . Так как  $f(a_1, \dots, x, \dots, a_{p-1}) = \sum_{i=1}^r \alpha_i \varepsilon_i(x)$ , то  $f(e_{j_1}, \dots, e_{j_p}) = 0$ , если хотя бы одно из чисел  $j_1, \dots, j_p$  больше  $r$ ; кроме того,  $\varepsilon_{i_1} \otimes \dots \otimes \varepsilon_{i_p}(e_{j_1}, \dots, e_{j_p}) = 1$ , если  $i_1 = j_1, \dots, i_p = j_p$ , а во всех остальных случаях эта величина равна нулю.  $\square$

**Теорема 5.1.7.2.** Если ковекторы  $\varepsilon_1, \dots, \varepsilon_r$  таковы, что  $f = \sum f_{i_1 \dots i_p} \varepsilon_{i_1} \otimes \dots \otimes \varepsilon_{i_p}$ , то  $W_f \subset \langle \varepsilon_1, \dots, \varepsilon_r \rangle$ .

**Доказательство.** Ясно, что  $f(a_1, \dots, a_{k-1}, x, a_k, \dots, a_{p-1}) = \sum f_{i_1 \dots i_p} \varepsilon_{i_1}(a_1) \dots \varepsilon_{i_k}(x) \dots \varepsilon_{i_p}(a_{p-1}) = \sum c_s \varepsilon_s(x)$ .  $\square$

### Задачи

**1.1.** Докажите, что  $v \otimes w = v' \otimes w' \neq 0$  тогда и только тогда, когда  $v = \lambda v'$  и  $w' = \lambda w$ .

**1.2.** Пусть  $A_i: V_i \rightarrow W_i$  ( $i = 1, 2$ ) — линейные отображения. Докажите, что:

- а)  $\text{Im}(A_1 \otimes A_2) = (\text{Im } A_1) \otimes (\text{Im } A_2)$ ;
- б)  $\text{Im}(A_1 \otimes A_2) = (\text{Im } A_1 \otimes W_2) \cap (W_1 \otimes \text{Im } A_2)$ ;
- в)  $\text{Ker}(A_2 \otimes A_2) = \text{Ker } A_1 \otimes W_2 + W_1 \otimes \text{Ker } A_2$ .

**1.3.** Пусть  $V_1, V_2 \subset V$  и  $W_1, W_2 \subset W$ . Докажите, что

$$(V_1 \otimes W_1) \cap (V_2 \otimes W_2) = (V_1 \cap V_2) \otimes (W_1 \cap W_2).$$

**1.4.** Докажите, что  $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$ .

**1.5.** а) Пусть  $A$  и  $B$  — симметрические матрицы. Докажите, что матрица  $A \otimes B$  симметрическая.

б) Пусть  $A$  и  $B$  — кососимметрические матрицы. Докажите, что матрица  $A \otimes B$  симметрическая.

**1.6.** а) Пусть  $A$  и  $B$  — симметрические матрицы с сигнатурами  $\sigma(A)$  и  $\sigma(B)$ . Докажите, что сигнатура симметрической матрицы  $A \otimes B$  равна  $\sigma(A)\sigma(B)$ .

б) Пусть  $A$  и  $B$  — кососимметрические матрицы. Докажите, что сигнатура симметрической матрицы  $A \otimes B$  равна нулю.

1.7. Пусть  $V$  — евклидово пространство и  $V^*$  естественным образом отождествлено с  $V$ . Докажите, что оператор  $A = I - 2a \otimes a$ , где  $a$  — единичный вектор, является симметрией относительно  $a^\perp$ .

1.8. Пусть  $A(x, y)$  — билинейная функция в евклидовом пространстве, причём если  $x \perp y$ , то  $A(x, y) = 0$ . Докажите, что функция  $A(x, y)$  пропорциональна скалярному произведению  $(x, y)$ .

1.9. Докажите, что пространство  $T_{p'}^{q'}(T_p^q(V))$  канонически изоморфно  $T_{pq'+qp'}^{pp'+qq'}(V)$ .

## 5.2. Симметрические и кососимметрические тензоры

### 5.2.1. Симметризация и альтернирование

Каждой перестановке  $\sigma \in S_q$  можно сопоставить линейный оператор  $f_\sigma: T_0^q(V) \rightarrow T_0^q(V)$ , переводящий элемент  $v_1 \otimes \dots \otimes v_q$  в  $v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(q)}$ . Тензор  $T \in T_0^q(V)$  называют *симметрическим* (соответственно *кососимметрическим*), если  $f_\sigma(T) = T$  (соответственно  $f_\sigma(T) = (-1)^\sigma T$ ) для любой перестановки  $\sigma$ . Симметрические тензоры образуют в  $T_0^q(V)$  подпространство  $S^q(V)$ , а кососимметрические — подпространство  $\Lambda^q(V)$ . Ясно, что  $S^q(V) \cap \Lambda^q(V) = 0$  при  $q \geq 2$ .

Оператор  $S = \frac{1}{q!} \sum_{\sigma} f_\sigma$  называют *симметризацией*, а оператор  $A = \frac{1}{q!} \sum_{\sigma} (-1)^\sigma f_\sigma$  — *антисимметризацией* или *альтернированием*.

**Теорема 5.2.1.1.** Оператор  $S$  является проектором  $T_0^q(V)$  на  $S^q(V)$ , а оператор  $A$  — проектором на  $\Lambda^q(V)$ .

**Доказательство.** Очевидно, что симметризация любого тензора симметрична и на симметрических тензорах оператор действует тождественно.

Так как для любого  $T \in T_0^q(V)$  имеем  $f_\sigma(AT) = \frac{1}{q!} \sum_{\tau} (-1)^\tau f_\sigma f_\tau(T) = (-1)^\sigma \frac{1}{q!} \sum_{\rho=\sigma\tau} (-1)^\rho f_\rho(T) = (-1)^\sigma AT$ , то  $\text{Im } A \subset \Lambda^q(V)$ . Если тензор  $T$  кососимметричен, то  $AT = \frac{1}{q!} \sum_{\sigma} (-1)^\sigma f_\sigma(T) = \frac{1}{q!} \sum_{\sigma} (-1)^\sigma (-1)^\sigma T = T$ .  $\square$

Введём обозначения  $S(e_{i_1} \otimes \dots \otimes e_{i_q}) = e_{i_1} \dots e_{i_q}$  и  $A(e_{i_1} \otimes \dots \otimes e_{i_q}) = e_{i_1} \wedge \dots \wedge e_{i_q}$ . Например,  $e_i e_j = (e_i \otimes e_j + e_j \otimes e_i)/2$  и  $e_i \wedge e_j = (e_i \otimes e_j - e_j \otimes e_i)/2$ . Если  $e_1, \dots, e_n$  — базис  $V$ , то тензоры  $e_{i_1} \dots e_{i_q}$  порождают  $S^q(V)$ , а тензоры  $e_{i_1} \wedge \dots \wedge e_{i_q}$  порождают  $\Lambda^q(V)$ . Тензор  $e_{i_1} \dots e_{i_q}$  зависит лишь от того,

сколько раз  $e_i$  встречается в этой записи, поэтому можно ввести обозначение  $e_{i_1} \dots e_{i_q} = e_1^{k_1} \dots e_n^{k_n}$ , где вектор  $e_i$  встречается в записи  $e_{i_1} \dots e_{i_q}$  ровно  $k_i$  раз. Тензор  $e_{i_1} \wedge \dots \wedge e_{i_q}$  при перестановке любых двух векторов  $e_{i_\alpha}$  и  $e_{i_\beta}$  меняет знак, поэтому  $e_{i_1} \wedge \dots \wedge e_{i_q} = 0$ , если  $e_{i_\alpha} = e_{i_\beta}$ , и тензоры  $e_{i_1} \wedge \dots \wedge e_{i_q}$ , где  $1 \leq i_1 < \dots < i_q \leq n$ , порождают пространство  $\Lambda^q(V)$ . В частности,  $\Lambda^q(V) = 0$  при  $q > n$ .

**Теорема 5.2.1.2.** Элементы  $e_1^{k_1} \dots e_n^{k_n}$ , где  $k_1 + \dots + k_n = q$ , образуют базис пространства  $S^q(V)$ , а элементы  $e_{i_1} \wedge \dots \wedge e_{i_q}$  образуют базис пространства  $\Lambda^q(V)$ .

**Доказательство.** Достаточно проверить линейную независимость этих векторов. Если наборы  $(k_1, \dots, k_n)$  и  $(l_1, \dots, l_n)$  различны, то тензоры  $e_1^{k_1} \dots e_n^{k_n}$  и  $e_1^{l_1} \dots e_n^{l_n}$  являются линейными комбинациями двух непересекающихся подмножеств базисных элементов пространства  $T_0^q(V)$ . Для тензоров вида  $e_{i_1} \wedge \dots \wedge e_{i_q}$  доказательство аналогично.  $\square$

**Следствие.**  $\dim \Lambda^q(V) = \binom{n}{q}$  и  $\dim S^q(V) = \binom{n+q-1}{q}$ .

**Доказательство.** Ясно, что количество наборов  $1 \leq i_1 < \dots < i_q \leq n$  равно  $\binom{n}{q}$ . Для вычисления количества наборов  $k_1 + \dots + k_n = q$  поступим следующим образом. Каждому такому набору сопоставим последовательность из  $q + n - 1$  шаров, среди которых  $q$  белых и  $n - 1$  чёрных; в этой последовательности сначала идёт  $k_1$  белых шаров, за ними один чёрный, за ним  $k_2$  белых, за ними один чёрный и т.д. Из  $n + q - 1$  шаров  $q$  белых шаров можно выбрать  $\binom{n+q-1}{q}$  способами.  $\square$

### 5.2.2. Алгебра Грассмана

В пространстве  $\Lambda(V) = \bigoplus_{q=0}^n \Lambda^q(V)$  можно ввести операцию внешнего произведения  $T_1 \wedge T_2 = A(T_1 \otimes T_2)$  для  $T_1 \in \Lambda^p(V)$  и  $T_2 \in \Lambda^q(V)$ ; на  $\Lambda(V)$  эта операция продолжается по линейности. Полученную алгебру  $\Lambda(V)$  называют *внешней алгеброй* или *алгеброй Грассмана* пространства  $V$ .

**Теорема 5.2.2.1.** Алгебра  $\Lambda(V)$  ассоциативна и косокоммутативна, т.е.  $T_1 \wedge T_2 = (-1)^{pq} T_2 \wedge T_1$  для  $T_1 \in \Lambda^p(V)$  и  $T_2 \in \Lambda^q(V)$ .



**Доказательство.** Докажем сначала, что  $A(T_1 \otimes T_2) = A(A(T_1) \otimes T_2)$ . Для этого достаточно рассмотреть случай, когда  $T_1 = x_1 \otimes \cdots \otimes x_p$  и  $T_2 = x_{p+1} \otimes \cdots \otimes x_{p+q}$ . Так как

$$A(x_1 \otimes \cdots \otimes x_p) = \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^\sigma x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(p)},$$

то

$$\begin{aligned} A(A(T_1) \otimes T_2) &= \\ &= A\left(\frac{1}{p!} \sum_{\sigma \in S_p} (-1)^\sigma x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(p)} \otimes x_{p+1} \otimes \cdots \otimes x_{p+q}\right) = \\ &= \frac{1}{p!(p+q)!} \sum_{\sigma \in S_p} \sum_{\tau \in S_{p+q}} (-1)^{\sigma\tau} x_{\tau(\sigma(1))} \otimes \cdots \otimes x_{\tau(p+q)}. \end{aligned}$$

Остаётся заметить, что

$$\begin{aligned} \sum_{\sigma \in S_p} (-1)^{\sigma\tau} x_{\tau(\sigma(1))} \otimes \cdots \otimes x_{\tau(p+q)} &= \\ &= p! \sigma(-1)^{\tau_1} x_{\tau_1(1)} \otimes \cdots \otimes x_{\tau_1(p+q)}, \end{aligned}$$

где  $\tau_1 = (\tau(\sigma(1)), \dots, \tau(\sigma(p)), \tau(p+1), \dots, \tau(p+q))$ .

Аналогично доказывается, что  $A(T_1 \otimes T_2) = A(T_1 \otimes A(T_2))$ , поэтому  $(T_1 \wedge T_2) \wedge T_3 = A(A(T_1 \otimes T_2) \otimes T_3) = A(T_1 \otimes T_2 \otimes T_3) = A(T_1 \otimes A(T_2 \otimes T_3)) = T_1 \wedge (T_2 \wedge T_3)$ .

Ясно, что  $x_{p+1} \otimes \cdots \otimes x_{p+q} \otimes x_1 \otimes \cdots \otimes x_p = x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(p+q)}$ , где  $\sigma = (p+1, \dots, p+q, 1, \dots, p)$ . Чтобы в  $\sigma$  последовательно переставить элемент 1 на первое место, ..., элемент  $p$  на  $p$ -е место, требуется совершить  $pq$  транспозиций. Следовательно,  $(-1)^\sigma = (-1)^{pq}$  и  $A(T_1 \otimes T_2) = (-1)^{pq} A(T_2 \otimes T_1)$ .  $\square$

В алгебре  $\Lambda(V)$   $k$ -я степень элемента  $\omega$  обозначается  $\Lambda^k \omega$ , иными словами  $\Lambda^k \omega = \omega \wedge \cdots \wedge \omega$  ( $k$  раз).

### 5.2.3. Кососимметрические функции

Кососимметрической функцией на  $V \times \cdots \times V$  называют полилинейная функция  $f(v_1, \dots, v_q)$ , для которой  $f(v_{\sigma(1)}, \dots, v_{\sigma(q)}) = (-1)^\sigma f(v_1, \dots, v_q)$  для любой перестановки  $\sigma$ .

**Теорема 5.2.3.1.** *Пространство  $\Lambda^q V^*$  канонически изоморфно пространству кососимметрических функций на  $V \times \cdots \times V$ , а также пространству  $(\Lambda^q V)^*$ .*

**Доказательство.** Легко проверить, что функция  $(f_1 \wedge \cdots \wedge f_q)(v_1, \dots, v_q) = A(f_1 \otimes \cdots \otimes f_q)(v_1, \dots, v_q) = \frac{1}{q!} \sum_{\sigma} (-1)^\sigma f_1(v_{\sigma(1)}) \cdots f_q(v_{\sigma(q)})$  — кососимметрическая. Если  $e_1, \dots, e_n$  — базис пространства  $V$ , то кососимметрическая функция  $f$  задаётся значениями  $f(e_{i_1}, \dots, e_{i_q})$ , где  $1 \leq i_1 < \cdots < i_q \leq n$ , причём любой такой набор значений соответствует некоторой кососимметрической функции. Поэтому размерность пространства кососимметрических функций равна размерности пространства  $\Lambda^q V^*$ , а значит, эти пространства изоморфны.

Построим теперь канонический изоморфизм  $\Lambda^q V^* \rightarrow (\Lambda^q V)^*$ . Линейное отображение  $V^* \otimes \cdots \otimes V^* \rightarrow (V \otimes \cdots \otimes V)^*$ , переводящее элемент  $(f_1, \dots, f_q) \in V^* \otimes \cdots \otimes V^*$  в полилинейную функцию  $f(v_1, \dots, v_q) = f_1(v_1) \cdots f_q(v_q)$ , является каноническим изоморфизмом. Рассмотрим ограничение этого отображения на  $\Lambda^q V^*$ . Элемент  $f_1 \wedge \cdots \wedge f_q = A(f_1 \otimes \cdots \otimes f_q) \in \Lambda^q V^*$  переходит в полилинейную функцию  $f(v_1, \dots, v_q) = \frac{1}{q!} \sum_{\sigma} (-1)^\sigma f_1(v_{\sigma(1)}) \cdots f_q(v_{\sigma(q)})$ .

Функция  $f$  кососимметрична, поэтому получаем отображение  $\Lambda^q V^* \rightarrow (\Lambda^q V)^*$ . Проверим, что это отображение — изоморфизм. Ясно, что

$$\begin{aligned} \tilde{f}(A(v_1 \otimes \cdots \otimes v_q)) &= \\ &= \left( \frac{1}{q!} \right)^2 \sum_{\sigma, \tau} (-1)^{\sigma\tau} f_1(v_{\sigma\tau(1)}) \cdots f_q(v_{\sigma\tau(q)}) = \\ &= \frac{1}{q!} \sum_{\sigma} (-1)^\sigma f_1(v_{\sigma(1)}) \cdots f_q(v_{\sigma(q)}) = \frac{1}{q!} \begin{vmatrix} f_1(v_1) & \cdots & f_1(v_q) \\ \dots & \dots & \dots \\ f_q(v_1) & \cdots & f_q(v_q) \end{vmatrix}. \end{aligned}$$

Пусть  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_n$  — двойственные базисы пространств  $V$  и  $V^*$ . Элементы  $e_{i_1} \wedge \cdots \wedge e_{i_q}$  образуют базис пространства  $\Lambda^q V$ ; рассмотрим двойственный базис пространства  $(\Lambda^q V)^*$ . Из доказанного выше равенства следует, что при рассматриваемом отображении элемент  $\varepsilon_{i_1} \wedge \cdots \wedge \varepsilon_{i_q}$  переходит в базисный элемент, двойственный  $q! e_{i_1} \wedge \cdots \wedge e_{i_q}$ .  $\square$

**Замечание.** Попутно доказано равенство  $\tilde{f}(A(v_1 \otimes \cdots \otimes v_q)) = \frac{1}{q} \tilde{f}(v_1 \otimes \cdots \otimes v_q)$  для  $f \in \Lambda^q V^*$ .

### 5.2.4. Свойства внешнего произведения

**Теорема 5.2.4.1.**  $T_0^2(V) = \Lambda^2(V) \oplus S^2(V)$ .

**Доказательство.** Для доказательства достаточно заметить, что  $a \otimes b = \frac{1}{2}(a \otimes b - b \otimes a) + \frac{1}{2}(a \otimes b + b \otimes a)$ .  $\square$

**Теорема 5.2.4.2.** Имеют место следующие канонические изоморфизмы:

$$\begin{aligned} \text{а) } \Lambda^q(V \oplus W) &\cong \bigoplus_{i=0}^q (\Lambda^i V \otimes \Lambda^{q-i} W); \\ \text{б) } S^q(V \oplus W) &\cong \bigoplus_{i=0}^q (S^i V \otimes S^{q-i} W). \end{aligned}$$

**Доказательство.** Ясно, что  $\Lambda^i V \subset T_0^i(V \oplus W)$  и  $\Lambda^{q-i} W \subset T_0^{q-i}(V \oplus W)$ . Поэтому имеется каноническое вложение  $\Lambda^i V \otimes \Lambda^{q-i} W \subset T_0^q(V \oplus W)$ . Спроектируем  $T_0^q(V \oplus W)$  на  $\Lambda^q(V \oplus W)$  с помощью оператора альтернирования. В результате получим каноническое отображение  $\Lambda^i V \otimes \Lambda^{q-i} W \rightarrow \Lambda^q(V \oplus W)$ . При этом отображении элемент  $(v_1 \wedge \dots \wedge v_i) \otimes (w_1 \wedge \dots \wedge w_{q-i})$  переходит в  $v_1 \wedge \dots \wedge v_i \wedge w_1 \wedge \dots \wedge w_{q-i}$ . Выбрав в пространствах  $V$  и  $W$  базисы, легко проверить, что построенное отображение

$$\bigoplus_{i=0}^q (\Lambda^i V \otimes \Lambda^{q-i} W) \rightarrow \Lambda^q(V \oplus W)$$

является изоморфизмом.

Для пространства  $S^q(V \oplus W)$  доказательство аналогично.  $\square$

**Теорема 5.2.4.3.** Если  $\dim V = n$ , то имеется канонический изоморфизм  $\Lambda^p V \cong (\Lambda^{n-p} V)^* \otimes \Lambda^n V$ .

**Доказательство.** Внешнее произведение является отображением  $\Lambda^p V \times \Lambda^{n-p} V \rightarrow \Lambda^n V$ , поэтому каждому элементу  $\Lambda^p V$  соответствует некоторое отображение  $\Lambda^{n-p} V \rightarrow \Lambda^n V$ . В итоге получаем отображение  $\Lambda^p V \rightarrow \text{Hom}(\Lambda^{n-p} V, \Lambda^n V) \cong (\Lambda^{n-p} V)^* \otimes \Lambda^n V$ . Докажем, что это отображение — изоморфизм. Выберем в пространстве  $V$  базис  $e_1, \dots, e_n$ . Элементу  $e_{i_1} \wedge \dots \wedge e_{i_p}$  соответствует отображение, переводящее  $e_{j_1} \wedge \dots \wedge e_{j_{n-p}}$  в 0 или в  $\pm e_1 \wedge \dots \wedge e_n$  в зависимости от того, пересекаются ли множества  $\{i_1, \dots, i_p\}$  и  $\{j_1, \dots, j_{n-p}\}$  или дополняют друг друга до  $\{1, \dots, n\}$ . Такие отображения образуют базис пространства  $\text{Hom}(\Lambda^{n-p} V, \Lambda^n V)$ .  $\square$

### 5.2.5. Тензорная степень оператора

Линейный оператор  $B: V \rightarrow V$  индуцирует линейный оператор  $B_q: T_0^q(V) \rightarrow T_0^q(V)$ , отображающий  $v_1 \otimes \cdots \otimes v_q$  в  $Bv_1 \otimes \cdots \otimes Bv_q$ . Если  $T = v_1 \otimes \cdots \otimes v_q$ , то  $B_q f_\sigma(T) = f_\sigma B_q(T)$ , поэтому

$$B_q f_\sigma(T) = f_\sigma B_q(T) \quad \text{для любого } T \in T_0^q(V). \quad (5.1)$$

Следовательно,  $B_q$  переводит симметрические тензоры в симметрические, а кососимметрические — в кососимметрические. Ограничения  $B_q$  на  $S^q(V)$  и  $\Lambda^q(V)$  обозначим  $S^q B$  и  $\Lambda^q B$  соответственно. Пусть  $S$  и  $A$  — операторы симметризации и альтернирования. Из равенства (5.1) следует, что  $B_q S = S B_q$  и  $B_q A = A B_q$ . Поэтому  $B_q(e_1^{k_1} \cdots e_n^{k_n}) = (B e_1)^{k_1} \cdots (B e_n)^{k_n}$  и  $B_q(e_{i_1} \wedge \cdots \wedge e_{i_q}) = (B e_{i_1}) \wedge \cdots \wedge (B e_{i_q})$ .

**Теорема 5.2.5.1.** Пусть

$$B_q(e_{i_1} \wedge \cdots \wedge e_{i_q}) = \sum_{1 \leq i_1 < \cdots < i_q \leq n} b_{j_1 \dots j_q}^{i_1 \dots i_q} e_{i_1} \wedge \cdots \wedge e_{i_q}.$$

Тогда  $b_{j_1 \dots j_q}^{i_1 \dots i_q}$  — минор  $B \begin{pmatrix} i_1 & \dots & i_q \\ j_1 & \dots & j_q \end{pmatrix}$  матрицы оператора  $B$ .

**Доказательство.** Ясно, что

$$\begin{aligned} B e_{j_1} \wedge \cdots \wedge B e_{j_q} &= \left( \sum_{i_1} b_{i_1 j_1} e_{i_1} \right) \wedge \cdots \wedge \left( \sum_{i_q} b_{i_q j_q} e_{i_q} \right) = \\ &= \sum_{i_1, \dots, i_q} b_{i_1 j_1} \cdots b_{i_q j_q} e_{i_1} \wedge \cdots \wedge e_{i_q} = \\ &= \sum_{1 \leq i_1 < \cdots < i_q \leq n} \left( \sum_{\sigma} (-1)^\sigma b_{i_{\sigma(1)} j_1} \cdots b_{i_{\sigma(q)} j_q} \right) e_{i_1} \wedge \cdots \wedge e_{i_q}. \end{aligned}$$

□

**Следствие.**  $\Lambda^q(B) = C_q(B)$  — ассоциированная матрица (см. §1.2.8).

Введём на множестве наборов индексов  $(i_1, \dots, i_q)$  лексикографический порядок, т.е. будем считать, что  $(i_1, \dots, i_q) < (j_1, \dots, j_q)$ , если  $i_1 = j_1, \dots, i_r = j_r$  и  $i_{r+1} < j_{r+1}$ . В соответствии с лексикографическим порядком упорядочим базисные векторы  $e_1^{k_1} \cdots e_n^{k_n}$  и  $e_{i_1} \wedge \cdots \wedge e_{i_q}$ .

**Теорема 5.2.5.2.** Если матрица оператора  $B$  треугольна в базисе  $e_1, \dots, e_n$ , то матрицы  $S^q B$  и  $\Lambda^q B$  треугольны в базисах  $e_1^{k_1} \cdots e_n^{k_n}$  и  $e_{i_1} \wedge \cdots \wedge e_{i_q}$ , упорядоченных лексикографически.

**Доказательство.** Пусть, например,  $Be_i \in \langle e_1, \dots, e_i \rangle$ , т.е. в соответствии с нашим отношением порядка  $Be_i \leq e_i$ . Если  $i_1 \leq j_1, \dots, i_q \leq j_q$ , то  $e_{i_1} \wedge \dots \wedge e_{i_q} \leq e_{j_1} \wedge \dots \wedge e_{j_q}$  и  $e_{i_1} \dots e_{i_q} = e_1^{k_1} \dots e_n^{k_n} \leq e_1^{l_1} \dots e_n^{l_n} = e_{j_1} \dots e_{j_q}$ . Поэтому имеем  $\Lambda^q B(e_{i_1} \wedge \dots \wedge e_{i_q}) \leq e_{i_1} \wedge \dots \wedge e_{i_q}$  и  $S^q B(e_1^{k_1} \dots e_n^{k_n}) \leq e_1^{k_1} \dots e_n^{k_n}$ .  $\square$

**Теорема 5.2.5.3.**  $\det(\Lambda^q B) = (\det B)^p$ , где  $p = \binom{n-1}{q-1}$  и  $\det(S^q B) = (\det B)^r$ , где  $r = \frac{q}{n} \binom{n+q-1}{q}$ .

**Доказательство.** Можно считать, что  $B$  — оператор над  $\mathbb{C}$ . Пусть  $e_1, \dots, e_n$  — жорданов базис для оператора  $B$ . Согласно теореме 5.2.5.2 матрицы операторов  $\Lambda^q B$  и  $S^q B$  треугольны в базисах  $e_{i_1} \wedge \dots \wedge e_{i_q}$  и  $e_1^{k_1} \dots e_n^{k_n}$ , упорядоченных лексикографически. Если вектору  $e_i$  соответствует диагональный элемент  $\lambda_i$ , то векторам  $e_{i_1} \wedge \dots \wedge e_{i_q}$  и  $e_1^{k_1} \dots e_n^{k_n}$  соответствуют диагональные элементы  $\lambda_{i_1} \dots \lambda_{i_q}$  и  $\lambda_1^{k_1} \dots \lambda_n^{k_n}$ , причём  $k_1 + \dots + k_n = q$ . Поэтому произведение всех диагональных элементов матриц  $\Lambda^q B$  и  $S^q B$  имеет по  $\lambda$  суммарную степень  $q \dim \Lambda^q(V)$  и  $q \dim S^q(V)$  соответственно. Следовательно,  $|\Lambda^q B| = |B|^p$  и  $|S^q B| = |B|^r$ , где  $p = \frac{q}{n} \binom{n}{q} = \binom{n-1}{q-1}$  и  $r = \frac{q}{n} \binom{n+q-1}{q}$ .  $\square$

Матрице  $B$  порядка  $n$  можно сопоставить многочлен  $\Lambda_B(t) = 1 + \sum_{q=1}^n \operatorname{tr}(\Lambda^q B) t^q$  и ряд  $S_B(t) = 1 + \sum_{q=1}^{\infty} \operatorname{tr}(S^q B) t^q$ .

**Теорема 5.2.5.4.**  $S_B(t) = (\Lambda_B(-t))^{-1}$ .

**Доказательство.** Как и при доказательстве теоремы 5.2.5.3, получаем, что если матрица  $B$  треугольная с диагональю  $(\lambda_1, \dots, \lambda_n)$ , то матрицы  $\Lambda^q B$  и  $S^q B$  треугольные с диагональными элементами  $\lambda_{i_1}, \dots, \lambda_{i_q}$  и  $\lambda_1^{k_1} \dots \lambda_n^{k_n}$  ( $k_1 + \dots + k_n = q$ ). Поэтому  $\Lambda_B(-t) = (1 - t\lambda_1) \dots (1 - t\lambda_n)$  и  $S_B(t) = (1 + t\lambda_1 + t^2\lambda_1^2 + \dots) \dots (1 + t\lambda_n + t^2\lambda_n^2 + \dots)$ . Остаётся заметить, что  $1/(1 - t\lambda_i) = 1 + t\lambda_i + t^2\lambda_i^2 + \dots$ .  $\square$

### 5.2.6. Отображение Ходжа

Пусть  $V$  — вещественное векторное пространство размерности  $n$ . Если фиксировать в  $V$  базис  $e_1, \dots, e_n$ , то можно рассмотреть отображение  $*$ :  $\Lambda^p V \rightarrow \Lambda^{n-p} V$ , которое переводит вектор  $e_{i_1} \wedge \dots \wedge e_{i_p}$  в вектор  $(-1)^\sigma e_{j_1} \wedge \dots \wedge e_{j_{n-p}}$ , где  $\{j_1, \dots, j_{n-p}\}$  — дополнение множества

$\{i_1, \dots, i_p\}$  в множестве  $\{1, \dots, n\}$  и  $\sigma = (i_1, \dots, i_p, j_1, \dots, j_{n-p})$ ; при этом предполагается, что  $i_1 < \dots < i_p$  и  $j_1 < \dots < j_{n-p}$ . Отображение  $*$  переводит базис пространства  $\Lambda^p V$  в базис пространства  $\Lambda^{n-p} V$ , поэтому оно является изоморфизмом. Это отображение называют *отображением Ходжа*.

**Теорема 5.2.6.1.** *Квадрат отображения Ходжа действует на пространстве  $\Lambda^p V$  как умножение на  $(-1)^{(n-p)p}$ .*

**Доказательство.** Ясно, что  $* * (e_{i_1} \wedge \dots \wedge e_{i_p}) = (-1)^{\sigma\tau} e_{i_1} \wedge \dots \wedge e_{i_p}$ , где  $\sigma = (i_1, \dots, i_p, j_1, \dots, j_{n-p})$  и  $\tau = (j_1, \dots, j_{n-p}, i_1, \dots, i_p)$ . В перестановках  $\sigma$  и  $\tau$  элементы набора  $(i_1, \dots, i_p)$  не образуют инверсий; элементы набора  $(j_1, \dots, j_{n-p})$  тоже не образуют инверсий. Кроме того, для любой пары индексов  $\alpha$  и  $\beta$  ровно одна из пар  $(i_\alpha, j_\beta)$  и  $(j_\beta, i_\alpha)$  образует инверсию. Поэтому сумма числа инверсий в перестановке  $\sigma$  и числа инверсий в перестановке  $\tau$  равна  $p(n-p)$ .  $\square$

Отображение Ходжа зависит от выбора базиса в пространстве  $V$ . Например, если вместо базиса  $e_1, e_2$  в 2-мерном пространстве взять базис  $\varepsilon_1, \varepsilon_2$ , где  $\varepsilon_1 = e_2, \varepsilon_2 = e_1$ , то отображение Ходжа изменит знак. Но если вместо одного ортонормированного базиса взять другой ортонормированный базис, который одинаково ориентирован с первым, то отображение Ходжа не изменится. Другими словами, имеет место следующее утверждение.

**Теорема 5.2.6.2.** *Выбор скалярного произведения и ориентации в пространстве  $V$  однозначно задаёт отображение Ходжа.*

**Доказательство.** Пусть  $A$  — ортогональная матрица, причём  $\det A = 1$ . В таком случае  $\text{adj } A = A^{-1} = A^T$ , поэтому, применяя тождество Якоби (теорема 1.2.6.3), получаем, что для любой перестановки

$$\tau = \begin{pmatrix} \alpha_1 & \dots & \alpha_p & \alpha_{p+1} & \dots & \alpha_n \\ i_1 & \dots & i_p & j_1 & \dots & j_{n-p} \end{pmatrix}$$

имеет место равенство

$$\begin{vmatrix} a_{\alpha_1 i_1} & \dots & a_{\alpha_p i_p} \\ \dots & \dots & \dots \\ a_{\alpha_p i_1} & \dots & a_{\alpha_p i_p} \end{vmatrix} = (-1)^\tau \begin{vmatrix} a_{\alpha_{p+1} j_1} & \dots & a_{\alpha_{p+1} j_{n-p}} \\ \dots & \dots & \dots \\ a_{\alpha_n j_1} & \dots & a_{\alpha_n j_{n-p}} \end{vmatrix}.$$

Поэтому отображение  $*$ , которое задаётся базисом  $e_1, \dots, e_n$ , переводит вектор

$$Ae_{\alpha_1} \wedge \dots \wedge Ae_{\alpha_p} = \sum_{i_1 < \dots < i_p} \begin{vmatrix} a_{\alpha_1 i_1} & \dots & a_{\alpha_p i_p} \\ \dots & \dots & \dots \\ a_{\alpha_p i_1} & \dots & a_{\alpha_p i_p} \end{vmatrix} e_{i_1} \wedge \dots \wedge e_{i_p}$$

в вектор

$$\sum_{i_1 < \dots < i_p} (-1)^\sigma (-1)^\tau \begin{vmatrix} a_{\alpha_{p+1}, j_1} & \dots & a_{\alpha_{p+1}, j_{n-p}} \\ \dots & \dots & \dots \\ a_{\alpha_n, j_1} & \dots & a_{\alpha_n, j_{n-p}} \end{vmatrix} e_{j_1} \wedge \dots \wedge e_{j_{n-p}},$$

где набор  $j_1, \dots, j_{n-p}$  и перестановка  $\sigma$  строятся по набору  $i_1, \dots, i_p$ , как описано выше. Ясно, что  $(-1)^\sigma (-1)^\tau = (-1)^\rho$ , где  $\rho = (\alpha_1, \dots, \alpha_n)$ . Таким образом, вектор  $Ae_{\alpha_1} \wedge \dots \wedge Ae_{\alpha_p}$  переходит в вектор  $(-1)^\rho Ae_{\alpha_{p+1}} \wedge \dots \wedge Ae_{\alpha_n}$ , т.е. отображение  $*$ , которое задаётся базисом  $e_1, \dots, e_n$ , совпадает с отображением  $*$ , которое задаётся базисом  $Ae_1, \dots, Ae_n$ .  $\square$

Отображение Ходжа является изометрией относительно скалярного произведения в  $\Lambda^{n-1}(\mathbb{R}^n)$ , заданного формулой

$$(v_1 \wedge \dots \wedge v_{n-1}, w_1 \wedge \dots \wedge w_{n-1}) = \|(v_i, w_j)\|_1^{n-1},$$

т.е. если  $*(v_1 \wedge \dots \wedge v_{n-1}) = v_n$  и  $*(w_1 \wedge \dots \wedge w_{n-1}) = w_n$ , то имеет место равенство

$$(v_n, w_n) = \|(v_i, w_j)\|_1^{n-1}.$$

Действительно, это равенство очевидно в случае, когда  $v_n$  и  $w_n$  — векторы из канонического базиса. Кроме того, обе части этого равенства — полилинейные функции от  $v_1, \dots, v_{n-1}$  и от  $w_1, \dots, w_{n-1}$ .

### 5.2.7. Векторное произведение

Отображение Ходжа в  $\mathbb{R}^3$  тесно связано с векторным произведением, а именно,  $v \times w = *(v \wedge w)$ . Действительно, обе части этого равенства линейны по  $v$  и  $w$ , поэтому его достаточно проверить для базисных векторов. Остаётся заметить, что  $*(e_1 \wedge e_2) = e_3$ ,  $*(e_1 \wedge e_3) = -e_2$  и  $*(e_2 \wedge e_3) = e_1$ .

Аналогичным образом для векторов  $v_1, \dots, v_{n-1} \in \mathbb{R}^{n-1}$  можно определить их *векторное произведение* следующим образом:

$$v_1 \times v_2 \times \dots \times v_{n-1} = *(v_1 \wedge v_2 \wedge \dots \wedge v_{n-1}).$$

Эквивалентное определение векторного произведения таково. Фиксируем векторы  $v_1, \dots, v_{n-1} \in \mathbb{R}^n$  и рассмотрим функцию  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ , заданную формулой  $\varphi(u) = \det(v_1, \dots, v_{n-1}, u)$  — определитель матрицы, столбцами которой служат координаты векторов  $v_1, \dots, v_{n-1}, u$ . Ясно, что  $\varphi$  — линейная функция, поэтому существует единственный вектор  $v_n \in \mathbb{R}^n$ , для которого  $\varphi(u) = (v_n, u)$  для всех  $u \in \mathbb{R}^n$ . Мы полагаем  $v_1 \times v_2 \times \dots \times v_{n-1} = v_n$ .

Полученный вектор  $v_n$  характеризуется тем свойством, что  $(v_n, u) = \det(v_1, \dots, v_{n-1}, u)$  для любого вектора  $u \in \mathbb{R}^n$ . Это согласуется с первым определением векторного произведения, поскольку

$$(*)(v_1 \wedge v_2 \wedge \dots \wedge v_{n-1}), u) = \det(v_1, \dots, v_{n-1}, u)$$

(это равенство достаточно проверить для набора векторов, который получен перестановкой векторов канонического базиса).

**Теорема 5.2.7.1.** *Векторное произведение обладает следующими свойствами:*

- а) вектор  $v_1 \times \dots \times v_{n-1}$  ортогонален каждому из векторов  $v_1, \dots, v_{n-1}$ ;
- б) если  $v_1 \times \dots \times v_{n-1} \neq 0$ , то набор векторов  $v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1}$  образует положительно ориентированный базис;
- в)  $|v_1 \times \dots \times v_{n-1}|$  — это  $(n-1)$ -мерный объём параллелепипеда, натянутого на векторы  $v_1, \dots, v_{n-1}$ .

**Доказательство.** а) Ясно, что  $(v_1 \times \dots \times v_{n-1}, v_i) = \det(v_1, \dots, v_{n-1}, v_i) = 0$ .

б) Ориентированный  $n$ -мерный объём параллелепипеда, натянутого на векторы  $v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1}$ , равен

$$\det(v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1}) = (v_1 \times \dots \times v_{n-1}, v_1 \times \dots \times v_{n-1}) > 0.$$

в) Случай, когда вектор  $v_1 \times \dots \times v_{n-1}$  нулевой, очевиден, поэтому будем считать, что этот вектор ненулевой. Вектор  $v_1 \times \dots \times v_{n-1}$  ортогонален гиперплоскости, порождённой векторами  $v_1, \dots, v_{n-1}$ . Поэтому  $n$ -мерный объём параллелепипеда, натянутого на векторы  $v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1}$ , равен произведению высоты  $|v_1 \times \dots \times v_{n-1}|$  на  $(n-1)$ -мерный объём параллелепипеда, натянутого на векторы  $v_1, \dots, v_{n-1}$ . Учитывая, что  $n$ -мерный объём параллелепипеда, натянутого на векторы  $v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1}$ , равен  $|v_1 \times \dots \times v_{n-1}|^2$ , получаем требуемое.  $\square$

## Задачи

**2.1.** Трилинейная функция  $f$  симметрична по двум первым аргументам и кососимметрична по двум последним. Докажите, что  $f \equiv 0$ .

**2.2.** Пусть  $f: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  — симметрическое билинейное отображение, причём  $(f(x, x), f(y, y)) \leq |f(x, y)|^2$  и  $f(x, x) \neq 0$  при  $x \neq 0$ . Докажите, что  $m \leq n$ .

**2.3.** Пусть  $e_1, \dots, e_{2n}$  — базис и  $\omega = e_1 \wedge e_2 + e_3 \wedge e_4 + \dots + e_{2n-1} \wedge e_{2n}$ . Докажите, что  $\Lambda^n \omega = n! e_1 \wedge \dots \wedge e_{2n}$ .



**2.4.** Пусть  $|A - \lambda I| = \sum_{k=0}^n a_k \lambda^k$  — характеристический многочлен матрицы  $A$ . Докажите, что  $a_k = (-1)^k \operatorname{tr}(\Lambda^{n-k} A)$ .

**2.5.** Пусть в пространстве  $V$  задано скалярное произведение. Докажите, что формула

$$(x \wedge y, z \wedge t) = \det \begin{pmatrix} (x, z) & (x, t) \\ (y, z) & (y, t) \end{pmatrix}$$

позволяет определить скалярное произведение на  $\Lambda^2 V$ . Постройте аналогичным образом скалярное произведение на  $\Lambda^k V$ .

**2.6.** Пусть  $d$  — определитель системы линейных уравнений  $\left( \sum_{j=1}^n a_{ij} x_j \right) \left( \sum_{q=1}^n a_{pq} x_q \right) = 0$ , где неизвестными считаются  $(n+1)n/2$  величин  $x_i x_j$ , упорядоченных лексикографически. Докажите, что  $d = (\det \|a_{ij}\|)^{n+1}$ .

**2.7.** Пусть  $s_k = \operatorname{tr}(A^k)$ ,  $\sigma_k$  — сумма главных миноров порядка  $k$  матрицы  $A$ . Докажите, что для любого натурального  $m$  выполняется равенство

$$s_m - s_{m-1} \sigma_1 + s_{m-2} \sigma_2 - \dots + (-1)^m m \sigma_m = 0.$$

**2.8.** Докажите формулу Бине–Коши с помощью внешнего произведения.

**2.9.** [Ga] Рассмотрим в пространстве  $\Lambda^2(\mathbb{R}^4)$  квадратичную форму  $Q$ , которая задаётся равенством

$$\omega_1 \wedge \omega_2 = Q(\omega_1, \omega_2) e_1 \wedge e_2 \wedge e_3 \wedge e_4.$$

Найдите сигнатуру этой квадратичной формы.

## 5.3. Пфаффиан

### 5.3.1. Определитель кососимметрической матрицы

Если  $A = \|a_{ij}\|$  — кососимметрическая матрица, то  $\det A$  является многочленом от независимых переменных  $a_{ij}$ , где  $i < j$ ; обозначим этот многочлен  $P(a_{ij})$ . Для  $n$  нечётного  $P \equiv 0$  (см. задачу 1.1); если  $n$  чётно, то матрицу  $A$  можно представить в виде  $A = XJX^T$ , где элементы матрицы  $X$  являются рациональными функциями от переменных  $a_{ij}$  и  $J = \operatorname{diag} \left( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$  (см. §3.2). Напомним, что это разложение следует рассматривать как разложение матриц, элементы которых

являются рациональными функциями от  $a_{ij}$ ; для некоторых значений  $a_{ij}$  эти рациональные функции могут иметь нулевые знаменатели. Так как  $\det X = f(a_{ij})/g(a_{ij})$ , где  $f$  и  $g$  — многочлены, то  $P = \det(XJX^T) = (f/g)^2$ . Поэтому  $f^2 = Pg^2$ , т.е.  $f^2$  делится на  $g^2$ , а значит,  $f$  делится на  $g$ , т.е.  $f/g = Q$  — некоторый многочлен. В итоге получаем, что  $P = Q^2$ , где  $Q$  — многочлен от  $a_{ij}$ , т.е. определитель кососимметрической матрицы, рассматриваемый как многочлен от  $a_{ij}$ , где  $i < j$ , является полным квадратом.

Этот результат можно получить и другим способом, дающим для многочлена  $Q$  явное выражение. Пусть в пространстве  $V$  задан базис  $e_1, \dots, e_{2n}$ . Сопоставим кососимметрической матрице  $A = \|a_{ij}\|_1^{2n}$  элемент  $\omega = \sum_{i < j} a_{ij} e_i \wedge e_j \in \Lambda^2(V)$ , а элементу  $\omega$  сопоставим  $\Lambda^n \omega = f(A) e_1 \wedge \dots \wedge e_{2n} \in \Lambda^{2n}(V)$ . Функция  $f(A)$  выражается через элементы матрицы  $A$ .

Запишем теперь элементы  $\omega$  и  $\Lambda^n \omega$  в базисе  $\varepsilon_j = \sum x_{ij} e_i$ . Можно проверить, что  $\sum_{i < j} a_{ij} e_i \wedge e_j = \sum_{i < j} b_{ij} \varepsilon_i \wedge \varepsilon_j$ , где  $A = XB X^T$ , и  $\varepsilon_1 \wedge \dots \wedge \varepsilon_{2n} = (\det X) e_1 \wedge \dots \wedge e_{2n}$ . Поэтому  $f(A) e_1 \wedge \dots \wedge e_{2n} = f(B) \varepsilon_1 \wedge \dots \wedge \varepsilon_{2n} = (\det X) f(B) e_1 \wedge \dots \wedge e_{2n}$ , т.е.  $f(XB X^T) = (\det X) f(B)$ . Если  $A$  — невырожденная кососимметрическая матрица, то её можно представить в виде  $A = XJX^T$ , где  $J = \text{diag} \left( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$ . Поэтому  $f(A) = f(XJX^T) = (\det X) f(J)$  и  $\det A = (\det X)^2 = (f(A)/f(J))^2$ .

Докажем, что  $f(A) = n! \sum (-1)^\sigma a_{i_1 i_2} a_{i_3 i_4} \dots a_{i_{2n-1} i_{2n}}$ , где  $\sigma = \begin{pmatrix} 1 & \dots & 2n \\ i_1 & \dots & i_{2n} \end{pmatrix}$  и суммирование ведётся по всем разбиениям множества  $\{1, \dots, 2n\}$  на пары  $\{i_k, i_{k+1}\}$ , где  $i_k < i_{k+1}$  (но не по всем перестановкам  $\sigma$ ). Пусть  $\omega_{ij} = a_{ij} e_i \wedge e_j$ , тогда  $\omega_{ij} \wedge \omega_{kl} = \omega_{kl} \wedge \omega_{ij}$ , причём  $\omega_{ij} \wedge \omega_{kl} = 0$ , если среди индексов  $i, j, k, l$  есть совпадающие. Поэтому  $\Lambda^n(\sum \omega_{ij}) = \sum \omega_{i_1 i_2} \wedge \dots \wedge \omega_{i_{2n-1} i_{2n}} = \sum a_{i_1 i_2} \dots a_{i_{2n-1} i_{2n}} e_{i_1} \wedge \dots \wedge e_{i_{2n}} = \sum (-1)^\sigma a_{i_1 i_2} \dots a_{i_{2n-1} i_{2n}} e_1 \wedge \dots \wedge e_{2n}$ , причём слагаемых, соответствующих коэффициенту  $a_{i_1 i_2} \dots a_{i_{2n-1} i_{2n}}$ , ровно  $n!$ ; в самом деле, каждый из  $n$  элементов  $\omega_{i_1}, \dots, \omega_{i_{2n-1} i_{2n}}$  можно выбрать в любом из  $n$  сомножителей  $\Lambda^n(\sum \omega_{ij})$ , причём в каждом сомножителе выбирается ровно один элемент. В частности,  $f(J) = n!$ .

Многочлен  $\text{Pf}(A) = f(A)/f(J) = \pm \sqrt{\det A}$ , рассматриваемый как многочлен от переменных  $a_{ij}$ , где  $i < j$ , называют *пфаффианом*. Легко проверить, что для матриц порядка 2 и 4 соответственно пфаффиан равен  $a_{12}$  и  $a_{12}a_{34} - a_{13}a_{23}$ .

### 5.3.2. Пфаффиан матриц специального вида

Пусть  $1 \leq \sigma_1 < \dots < \sigma_{2k} \leq 2n$ . Множество  $\{\sigma_1, \dots, \sigma_{2k}\}$  можно дополнить до множества  $\{1, 2, \dots, 2n\}$  множеством  $\{\bar{\sigma}_1, \dots, \bar{\sigma}_{2(n-k)}\}$ , где

$\bar{\sigma}_1 < \dots < \bar{\sigma}_{2(n-k)}$ . В результате множеству  $\{\bar{\sigma}_1, \dots, \bar{\sigma}_{2k}\}$  сопоставим перестановку  $\sigma = (\sigma_1 \dots \sigma_{2k} \bar{\sigma}_1 \dots \bar{\sigma}_{2(n-k)})$ . Легко проверить, что  $(-1)^\sigma = (-1)^a$ , где  $a = (\sigma_1 - 1) + (\sigma_2 - 2) + \dots + (\sigma_{2k} - 2k)$ .

Пфаффиан подматрицы кососимметрической матрицы  $M = \|m_{ij}\|_1^{2n}$ , где  $m_{ij} = (-1)^{i+j+1}$ , обладает следующим свойством.

**Теорема 5.3.2.1.** Пусть  $P_{\sigma_1 \dots \sigma_{2k}} = \text{Pf} \left( M \begin{pmatrix} \sigma_1 & \dots & \sigma_{2k} \\ \sigma_1 & \dots & \sigma_{2k} \end{pmatrix} \right)$ . Тогда  $P_{\sigma_1 \dots \sigma_{2k}} = (-1)^\sigma$ .

**Доказательство.** Применим индукцию по  $k$ . Ясно, что  $P_{\sigma_1 \sigma_2} = m_{\sigma_1 \sigma_2} = (-1)^{\sigma_1 + \sigma_2 + 1}$ . Знак перестановки, соответствующей  $\{\sigma_1, \sigma_2\}$ , равен  $(-1)^a$ , где  $a = (\sigma_1 - 1) + (\sigma_2 - 2) \equiv (\sigma_1 + \sigma_2 + 1) \pmod{2}$ . Воспользовавшись результатом задачи 3.1, легко проверить, что  $P_{\sigma_1 \dots \sigma_{2k}} =$

$$\sum_{i=2}^{2k} (-1)^i P_{\sigma_1 \sigma_i} P_{\sigma_2 \dots \widehat{\sigma_i} \dots \sigma_{2k}}. \text{ По предположению индукции } P_{\sigma_2 \dots \widehat{\sigma_i} \dots \sigma_{2k}} = (-1)^\tau,$$

где  $\tau = (\sigma_2 \dots \widehat{\sigma_i} \dots \sigma_{2k} 1 2 \dots 2m)$ . Знаки перестановок  $\sigma$  и  $\tau$  равны  $(-1)^a$  и  $(-1)^b$ , где  $a = (\sigma_1 - 1) + \dots + (\sigma_{2k} - 2k)$  и  $b = (\sigma_2 - 1) + (\sigma_3 - 2) + \dots + (\sigma_{i-1} - i + 2) + (\sigma_{i+1} - i + 1) + \dots + (\sigma_{2k} - 2k + 2)$ , т.е.  $(-1)^\tau = (-1)^\sigma (-1)^{\sigma_1 + \sigma_i + 1}$ . Значит,

$$P_{\sigma_1 \dots \sigma_{2k}} = \sum_{i=2}^{2k} (-1)^i (-1)^{\sigma_1 + \sigma_i + 1} (-1)^\sigma (-1)^{\sigma_1 + \sigma_i + 1} = (-1)^\sigma \sum_{i=2}^{2k} (-1)^i = (-1)^\sigma.$$

□

**Теорема 5.3.2.2.** Пусть  $M$  — определённая выше матрица,  $A$  — кососимметрическая матрица порядка  $2n$ . Тогда  $\text{Pf}(A + \lambda^2 M) = \sum_{k=0}^n \lambda^{2k} P_k$ , где  $P_k = \sum_{\sigma} A \begin{pmatrix} \sigma_1 & \dots & \sigma_{2(n-k)} \\ \sigma_1 & \dots & \sigma_{2(n-k)} \end{pmatrix}$ .

**Доказательство.** [Ка2] Матрицы  $A$  и  $M$  будем рассматривать как элементы  $\sum_{i < j} a_{ij} e_i \wedge e_j$  и  $\sum_{i < j} m_{ij} e_i \wedge e_j$  в  $\Lambda^2 V$ . Так как  $A \wedge M = M \wedge A$ , справедлива

$$\text{формула бинома Ньютона: } \Lambda^n(A + \lambda^2 M) = \sum_{k=0}^n \binom{n}{k} \lambda^{2k} (\Lambda^k M) \wedge (\Lambda^{n-k} A) =$$

$$\sum_{k=0}^n \binom{n}{k} \lambda^{2k} \sum (k! P_{\sigma_1 \dots \sigma_{2k}}) ((n-k)! P_k) e_{\sigma_1} \wedge \dots \wedge e_{\sigma_k} \wedge \dots. \text{ Согласно теореме 5.3.2.1}$$

$$P_{\sigma_1 \dots \sigma_{2k}} = (-1)^\sigma; \text{ ясно также, что } e_{\sigma_1} \wedge \dots \wedge e_{\sigma_k} \wedge \dots = (-1)^\sigma e_1 \wedge \dots \wedge e_{2n}.$$

Поэтому  $\Lambda^n(A + \lambda^2 M) = n! \sum_{k=0}^n \lambda^{2k} P_k e_1 \wedge \dots \wedge e_n$ , а значит,  $\text{Pf}(A + \lambda^2 M) =$

$$\sum_{k=0}^n \lambda^{2k} P_k. \quad \square$$

## Задачи

**3.1.** Пусть  $\text{Pf}(A) = a_{pq} C_{pq} + f$ , где многочлен  $f$  не зависит от  $a_{pq}$ ;  $A_{pq}$

— матрица, полученная из  $A$  вычёркиванием строк и столбцов с номерами  $p$  и  $q$ . Докажите, что  $C_{pq} = (-1)^{p+q+1} \text{Pf}(A_{pq})$ .

**3.2.** Пусть  $X$  — матрица порядка  $2n$ , строками которой являются координаты векторов  $x_1, \dots, x_{2n}$ ;  $g_{ij} = \{x_i, x_j\}$ , где величина  $\{a, b\}$  для векторов  $a = (a_1, \dots, a_{2n})$  и  $b = (b_1, \dots, b_{2n})$  равна  $\sum_{k=1}^n (a_{2k-1}b_{2k} - a_{2k}b_{2k-1})$ . Докажите, что  $\det X = \text{Pf}(G)$  где  $G = \|g_{ij}\|_1^{2n}$ .

## 5.4. Разложимые тензоры

### 5.4.1. Разложимый тензор определяет подпространство

Кососимметрический тензор  $\omega \in \Lambda^k(V)$  называют *разложимым*, если его можно представить в виде  $\omega = x_1 \wedge \dots \wedge x_k$ , где  $x_i \in V$ .

Симметрический тензор  $T \in S^k(V)$  называют *разложимым*, если его можно представить в виде  $T = S(x_1 \otimes \dots \otimes x_k)$ , где  $x_i \in V$ .

**Теорема 5.4.1.1.** Если  $x_1 \wedge \dots \wedge x_k = y_1 \wedge \dots \wedge y_k \neq 0$ , то  $\langle x_1, \dots, x_k \rangle = \langle y_1, \dots, y_k \rangle$ .

**Доказательство.** Предположим, например, что  $y_1 \notin \langle x_1, \dots, x_k \rangle$ . Тогда векторы  $e_1 = x_1, \dots, e_k = x_k$  и  $e_{k+1} = y_1$  можно дополнить до базиса. Разложив векторы  $y_1, \dots, y_k$  по этому базису, получим равенство  $e_1 \wedge \dots \wedge e_k = e_{k+1} \wedge (\sum a_{i_2 \dots i_k} e_{i_2} \wedge \dots \wedge e_{i_k})$ . Это равенство противоречит линейной независимости векторов  $e_{i_1} \wedge \dots \wedge e_{i_k}$ .  $\square$

**Следствие.** Любому разложимому кососимметрическому тензору  $\omega = x_1 \wedge \dots \wedge x_k$  можно сопоставить  $k$ -мерное подпространство  $\langle x_1, \dots, x_k \rangle$ ; это подпространство не зависит от разложения, а зависит лишь от самого тензора  $\omega$ .

**Теорема 5.4.1.2.** Пусть  $W_1$  и  $W_2$  — подпространства в  $V$  размерности  $p_1$  и  $p_2$ ; им соответствуют  $\omega_1 \in \Lambda^{p_1}(V)$  и  $\omega_2 \in \Lambda^{p_2}(V)$ .

а)  $W_1 \subset W_2$  тогда и только тогда, когда существует  $\omega \in \Lambda^{p_2-p_1}(V)$ , для которого  $\omega_1 \wedge \omega = \omega_2$ .

б)  $W_1 \cap W_2 = 0$  тогда и только тогда, когда  $\omega_1 \wedge \omega_2 \neq 0$ . При этом  $\omega_1 \wedge \omega_2$  соответствует подпространству  $W_1 \oplus W_2$ .

**Доказательство.** а) Пусть  $W_1 \subset W_2$ . Выберем в  $W_1$  базис  $e_1, \dots, e_{p_1}$  и дополним его до базиса  $e_1, \dots, e_{p_2}$  пространства  $W_2$ . Тогда  $\omega_1 = e_1 \wedge \dots \wedge e_{p_1}$

$\cdots \wedge e_{p_1}$ ,  $\omega_2 = e_1 \wedge \cdots \wedge e_{p_2}$  и  $\omega = e_{p_1} \wedge \cdots \wedge e_{p_2}$ . Обратное утверждение очевидно.

б) Пусть  $W_1 = \langle e_1, \dots, e_p \rangle$  и  $W_2 = \langle e_{p+1}, \dots, e_q \rangle$ . Векторы  $e_1, \dots, e_{p+q}$  линейно независимы тогда и только тогда, когда  $e_1 \wedge \cdots \wedge e_{p+q} \neq 0$ .  $\square$

**Теорема 5.4.1.3.** Если  $S(x_1 \otimes \cdots \otimes x_k) = S(y_1 \otimes \cdots \otimes y_k) \neq 0$ , то  $\langle x_1, \dots, x_k \rangle = \langle y_1, \dots, y_k \rangle$ .

**Доказательство.** [Me] Предположим, например, что  $y_1 \notin \langle x_1, \dots, x_k \rangle$ . Пусть  $T = S(x_1 \otimes \cdots \otimes x_k)$  — ненулевой тензор. Любой полилинейной функции  $f: V \times \cdots \times V \rightarrow K$  соответствует линейная функция  $\tilde{f}: V \otimes \cdots \otimes V \rightarrow K$ . Тензор  $T$  ненулевой, поэтому существует такая линейная функция  $\tilde{f}$ , что  $\tilde{f}(T) \neq 0$ . Полилинейная функция  $f$  является линейной комбинацией произведений линейных функций, поэтому существуют такие линейные функции  $g_1, \dots, g_k$ , что  $\tilde{g}(T) \neq 0$ , где  $g = g_1 \cdots g_k$ . Рассмотрим линейные функции  $h_1, \dots, h_k$ , совпадающие с  $g_1, \dots, g_k$  на подпространстве  $\langle x_1, \dots, x_k \rangle$  и принимающие значение 0 на векторе  $y_1$ ; пусть  $h = h_1 \cdots h_k$ . Тогда  $\tilde{h}(T) = \tilde{g}(T) \neq 0$ . С другой стороны,  $T = S(y_1 \otimes \cdots \otimes y_k)$ , поэтому  $\tilde{h}(T) = \sum_{\sigma} h_1(y_{\sigma(1)}) \cdots h_k(y_{\sigma(k)}) = 0$ , так как в любое произведение входит  $h_i(y_1) = 0$ . Получено противоречие, поэтому  $y_1 \in \langle x_1, \dots, x_k \rangle$ . Аналогичные рассуждения показывают, что  $\langle y_1, \dots, y_k \rangle \subset \langle x_1, \dots, x_k \rangle$  и  $\langle x_1, \dots, x_k \rangle \subset \langle y_1, \dots, y_k \rangle$ .  $\square$

### 5.4.2. Соотношения Плюккера

Исходя непосредственно из определения разложимого кососимметрического тензора, нельзя за конечное число операций выяснить, разложим ли данный кососимметрический тензор  $\sum_{i_1 < \cdots < i_k} a_{i_1 \dots i_k} e_{i_1} \wedge \cdots \wedge e_{i_k}$ . Мы покажем, что условие разложимости кососимметрического тензора эквивалентно некоторой системе уравнений для его координат  $a_{i_1 \dots i_k}$  (соотношения Плюккера). Предварительно сделаем несколько замечаний.

Пусть  $\alpha \in \Lambda^p V^*$ . Зададим отображение  $i(\alpha): \Lambda^k V \rightarrow \Lambda^{k-p} V$  так, чтобы для всех  $T \in \Lambda^k V$  и  $f \in \Lambda^{k-p} V^*$  выполнялось равенство  $\langle i(\alpha)T, f \rangle = \langle T, \alpha \wedge f \rangle$ . Аналогично для  $\omega \in \Lambda^p V$  зададим отображение  $i(\omega): \Lambda^k V^* \rightarrow \Lambda^{k-p} V^*$  так, чтобы для всех  $f \in \Lambda^k V^*$  и  $T \in \Lambda^{k-p} V$  выполнялось равенство  $\langle i(\omega)f, T \rangle = \langle f, \omega \wedge T \rangle$ .

Будем говорить, что подпространство  $W \subset U$  обёртывает кососимметрический тензор  $\omega \in \Lambda^k V$ , если  $\omega \in \Lambda^k W$ , т.е.  $\omega$  представляется в виде суммы внешних произведений элементов  $W$ . Если  $\omega$  — разложимый кососимметрический тензор, соответствующий подпространству  $U$ , то

$W$  обёртывает  $\omega$  тогда и только тогда, когда  $W \supset U$ , т.е.  $U$  — минимальное обёртывающее подпространство для  $\omega$ . Покажем, что для любого  $\omega \in \Lambda^k V$  существует минимальное обёртывающее подпространство  $W$ , причём  $\dim W = k$  тогда и только тогда, когда  $\omega$  разложим (если  $\omega$  неразложим, то  $\dim W > k$ ).

**Теорема 5.4.2.1.** *Пространство  $W$  обёртывает  $\omega \in \Lambda^k V$  тогда и только тогда, когда  $i(\alpha)\omega \in W$  для любого  $\alpha \in \Lambda^{k-1} V^*$ .*

**Доказательство.** Пусть  $e_1, \dots, e_m$  — базис пространства  $W$ . Дополним его до базиса  $e_1, \dots, e_n$  пространства  $V$  и рассмотрим двойственный базис  $\varepsilon_1, \dots, \varepsilon_n$ .

Пусть  $\omega = \sum x_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k}$  и  $\alpha = \varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_{k-1}}$ . Тогда

$$\begin{aligned} \langle i(\alpha)\omega, \varepsilon_j \rangle &= \langle \omega, \alpha \wedge \varepsilon_j \rangle = \\ &= \left\langle \sum x_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k}, \varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_{k-1}} \wedge \varepsilon_j \right\rangle = \\ &= \pm \frac{1}{k!} x_{i_1 \dots i_k}, \end{aligned}$$

если  $j$  дополняет набор  $j_1, \dots, j_{k-1}$  до  $i_1, \dots, i_k$ ; в противном случае получаем 0. Таким образом,  $k!i(\alpha)\omega = \sum \pm x_{j_1 \dots j_{k-1} j} e_j$ . В этой сумме каждый вектор  $e_j$ , где  $j \notin \{j_1, \dots, j_{k-1}\}$ , встречается только один раз, поскольку  $j_1 < \dots < j_{k-1} < j < \dots < j_{k-1}$ . Теперь ясно, что  $i(\alpha)\omega \in W$  для всех  $\alpha \in \Lambda^{k-1} V^*$  тогда и только тогда, когда  $x_{i_1 \dots i_k} = 0$ , если хотя бы один из индексов  $i_1, \dots, i_k$  больше  $m$ . Последнее условие эквивалентно тому, что подпространство  $W$  обёртывает  $\omega$ .  $\square$

**Следствие.** *Подпространство, состоящее из элементов вида  $i(\alpha)\omega$ , где  $\omega \in \Lambda^k V$  фиксирован, а  $\alpha$  пробегает пространство  $\Lambda^{k-1} V^*$ , является минимальным обёртывающим подпространством для  $\omega \in \Lambda^k V$ .*

Ясно, что размерность минимального обёртывающего подпространства для  $\omega \in \Lambda^k V$  не меньше  $k$ , причём она равна  $k$  тогда и только тогда, когда  $\omega$  разложим.

**Теорема 5.4.2.2.** *Пусть  $\omega \in \Lambda^k V$  и  $W$  — его минимальное обёртывающее пространство. Кососимметрический тензор  $\omega$  разложим тогда и только тогда, когда  $w \wedge \omega = 0$  для любого  $w \in W$ .*

**Первое доказательство.** Предположим сначала, что  $\omega = v_1 \wedge \dots \wedge v_k \neq 0$ . Тогда  $W = \langle v_1, \dots, v_k \rangle$  и  $w \wedge \omega = 0$  для любого вектора  $w \in W$ .

Предположим теперь, что  $w \wedge \omega = 0$  для любого  $w \in W$ . Пусть  $e_1, \dots, e_m$  — базис  $W$  и  $\omega = \sum x_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k}$ . По предположению  $e_j \wedge \omega = 0$

для  $j = 1, \dots, m$ , т.е.  $\sum_{i_1 < \dots < i_k} x_{i_1 \dots i_k} e_j \wedge e_{i_1} \wedge \dots \wedge e_{i_k} = 0$ . Пусть  $x_{i_1 \dots i_k} \neq 0$  для некоторого набора индексов  $i_1 < \dots < i_k$ . Предположим, что  $m > k$ . Тогда можно выбрать индекс  $j$ , отличный от  $i_1, \dots, i_k$ , а в таком случае  $\sum_{i_1 < \dots < i_k} x_{i_1 \dots i_k} e_j \wedge e_{i_1} \wedge \dots \wedge e_{i_k} \neq 0$ , поскольку в этой сумме есть по крайней мере одно ненулевое слагаемое и нет ненулевых линейно независимых слагаемых. Полученное противоречие показывает, что  $m = k$ , т.е.  $\omega$  разложим.  $\square$

**Второе доказательство.** Для разложимого тензора утверждение очевидно (см. первое доказательство). Предположим теперь, что кососимметрический тензор  $\omega$  неразложим. Тогда  $\dim W = m > k$ . Операция внешнего произведения индуцирует невырожденное отображение  $\Lambda^k W \otimes \Lambda^{m-k} W \rightarrow \Lambda^m W$  (см. теорему 5.2.4.3), поэтому  $\omega \wedge \alpha \neq 0$  для некоторого  $\alpha \in \Lambda^{m-k} W$ . А так как  $\alpha = \sum a_{i_1 \dots i_{m-k}} w_{i_1} \wedge \dots \wedge w_{i_{m-k}}$ , то  $\omega \wedge w_i \neq 0$  для некоторого  $w_i \in W$ .  $\square$

**Следствие (соотношения Плюккера).** Кососимметрический тензор  $\omega = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k}$  разложим тогда и только тогда, когда

$$\left( \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k} \right) \wedge \left( \sum_j a_{j_1 \dots j_{k-1} j} e_{j_1} \wedge \dots \wedge e_{j_{k-1}} \right) = 0$$

для любого набора чисел  $j_1 < \dots < j_{k-1}$ .

**Доказательство.** Пусть  $W$  — минимальное обертывающее пространство для  $\omega$ . Вычисление из теоремы 5.4.2.1 показывает, что  $W$  состоит из векторов  $\sum_j a_{j_1 \dots j_{k-1} j} e_j$  (подразумевается, что коэффициент  $a_{j_1 \dots j_{k-1} j}$  меняет знак при перестановке любых двух индексов). Согласно теореме 5.4.2.2 кососимметрический тензор  $\omega$  разложим тогда и только тогда, когда  $\omega \wedge w = 0$  для всех  $w \in W$ .  $\square$

Рассмотрим подробнее случай  $k = 2$ . Коэффициент при  $e_i \wedge e_j \wedge e_p$  для фиксированного  $j_1 = q$  равен  $a_{ij}a_{qp} + a_{pi}a_{qj} + a_{jp}a_{qi}$ . Легко также проверить, что соотношение  $a_{ij}a_{qp} + a_{pi}a_{qj} + a_{jp}a_{qi} = 0$  нетривиально, только если индексы  $i, j, p, q$  попарно различны.

Критерий разложимости кососимметрического тензора (теорема 5.4.2.2) можно записать и по-другому.

**Теорема 5.4.2.3.** а) Кососимметрический тензор  $\omega \in \Lambda^k V$  разложим тогда и только тогда, когда  $i(\alpha)\omega \wedge \omega = 0$  для любого  $\alpha \in \Lambda^{k-1} V^*$ .

б) Кососимметрический тензор  $\omega \in \Lambda^k V$  разложим тогда и только тогда, когда  $i(i(\omega)\beta)\omega = 0$  для любого  $\beta \in \Lambda^{k+1} V^*$ .

**Доказательство.** а) Согласно теореме 5.4.2.1 минимальное обёртывающее пространство  $W$  состоит из векторов вида  $i(\alpha)\omega$ .

б) Это условие эквивалентно условию а). Действительно,  $i(\omega)\beta \in \Lambda^1 V^*$  и  $i(i(\omega)\beta)\omega \in \Lambda^{k-1} V$ . Поэтому для  $\alpha \in \Lambda^{k-1} V^*$  можно рассмотреть

$$\begin{aligned} \langle i(i(\omega)\beta)\omega, \alpha \rangle &= \langle \omega, (i(\omega)\beta) \wedge \alpha \rangle = \\ &= (-1)^{k-1} \langle \omega, \alpha \wedge (i(\omega)\beta) \rangle = \\ &= (-1)^{k-1} \langle i(\alpha)\omega, i(\omega)\beta \rangle = \\ &= (-1)^{k-1} \langle \omega \wedge i(\alpha)\omega, \beta \rangle. \end{aligned}$$

Таким образом,  $i(i(\omega)\beta)\omega = 0$  для всех  $\beta \in \Lambda^{k+1} V^*$  тогда и только тогда, когда  $i(\alpha)\omega \wedge \omega = 0$  для всех  $\alpha \in \Lambda^{k-1} V^*$ .  $\square$

### 5.4.3. Альтернатива соотношениям Плюккера

Соотношения Плюккера — это система уравнений степени 2, задающая множество разложимых тензоров в  $\Lambda^m V$ . То же самое множество можно задать системой уравнений степени  $m$ . Делается это следующим образом.

Для  $\alpha_1, \dots, \alpha_m \in V^*$  можно рассмотреть линейные отображения  $c_j : \Lambda^m V \rightarrow V$ ,  $j = 1, \dots, m$ , и  $c : \Lambda^m V \rightarrow \mathbb{R}$ , которые задаются следующими формулами:

$$\begin{aligned} c_j(v_1 \wedge \dots \wedge v_m) &= \begin{vmatrix} \alpha_1(v_1) & \dots & \alpha_{j-1}(v_1) & v_1 & \alpha_{j+1}(v_1) & \dots & \alpha_m(v_1) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_1(v_m) & \dots & \alpha_{j-1}(v_m) & v_m & \alpha_{j+1}(v_m) & \dots & \alpha_m(v_m) \end{vmatrix}, \\ c(v_1 \wedge \dots \wedge v_m) &= \begin{vmatrix} \alpha_1(v_1) & \dots & \alpha_m(v_1) \\ \dots & \dots & \dots \\ \alpha_1(v_m) & \dots & \alpha_m(v_m) \end{vmatrix}. \end{aligned}$$

Здесь выражение для  $c_j$  нужно понимать как формальное разложение определителя по  $j$ -му столбцу, т.е.  $c_j(v_1 \wedge \dots \wedge v_m)$  — это линейная комбинация векторов  $v_1, \dots, v_m$  с соответствующими коэффициентами.

**Теорема 5.4.3.1 ([Ro1]).** Кососимметрический тензор  $\omega \in \Lambda^m V$  разложим тогда и только, когда для любых  $\alpha_1, \dots, \alpha_m \in V^*$  имеет место равенство

$$c_1(\omega) \wedge \dots \wedge c_m(\omega) = (c(\omega))^{m-1} \omega, \quad (5.1)$$

где  $c_1, \dots, c_m$  и  $c$  определены по  $\alpha_1, \dots, \alpha_m$ , как указано выше.



**Доказательство.** Если равенство (5.1) имеет место для всех  $\alpha_1, \dots, \alpha_m$ , то для данного  $\omega \neq 0$  мы сразу получаем явное разложение, поскольку  $\alpha_1, \dots, \alpha_m$  можно выбрать так, что  $c(\omega) = 1$ .

Предположим теперь, что  $\omega = v_1 \wedge \dots \wedge v_m$ . Равенство (5.1) можно доказать разными способами.

*Первый способ.* Пусть  $a_{ij} = \alpha_i(v_j)$  и  $M_{ij}$  — определитель матрицы, полученной из матрицы  $A = \|a_{ij}\|_1^m$  вычёркиванием  $i$ -й строки и  $j$ -го столбца. Тогда  $c_1(\omega) \wedge \dots \wedge c_m(\omega) = w_1 \wedge \dots \wedge w_m$ , где  $w_k = \sum_{i=1}^m (-1)^{i_k+k} M_{ki} v_{i_k}$ . Поэтому

$$c_1(\omega) \wedge \dots \wedge c_m(\omega) = \sum_{(i_1, \dots, i_m) \in S_m} (-1)^{(i_1, \dots, i_m)} M_{1i_1} \dots M_{mi_m} v_{i_1} \wedge \dots \wedge v_{i_m}.$$

Сумма  $\sum (-1)^{(i_1, \dots, i_m)} M_{1i_1} \dots M_{mi_m}$  равна определителю матрицы  $\|M_{ij}\|$ . Согласно задаче 1.29 этот определитель равен определителю присоединённой матрицы  $\text{adj } A$ , а определитель присоединённой матрицы равен  $|A|^{m-1}$ .

*Второй способ.* Каждый вектор  $c_j(\omega)$  является линейной комбинацией векторов  $v_1, \dots, v_m$ , поэтому

$$c_1(\omega) \wedge \dots \wedge c_m(\omega) = \lambda v_1 \wedge \dots \wedge v_m. \quad (5.2)$$

Непосредственно из определения  $c_j$  видно, что  $\alpha_i(c_j(\omega)) = \delta_{ij} c(\omega)$ . Поэтому, применив  $c$  к обеим частям равенства (5.2), получим  $(c(\omega))^m = \lambda c(\omega)$ , т.е.  $c(\omega)(\lambda - (c(\omega))^{m-1}) = 0$ . Если  $c(\omega) \neq 0$ , то мы получаем  $\lambda = (c(\omega))^{m-1}$ , что и требовалось. Если же  $c(\omega) = 0$ , то матрица  $A = \|\alpha_j(v_i)\|$  вырожденная, а потому присоединённая матрица  $\text{adj } A$  тоже вырожденная. Из этого легко вывести, что векторы  $c_1(\omega), \dots, c_m(\omega)$  линейно зависимы, поскольку эти векторы являются линейными комбинациями векторов  $v_1, \dots, v_m$ , коэффициентами которых служат элементы матрицы  $\text{adj } A$ . Таким образом, если  $c(\omega) = 0$ , то  $c_1(\omega) \wedge \dots \wedge c_m(\omega) = 0$ , поэтому равенство (5.1) снова выполняется.  $\square$

**Замечание.** Записав равенство (5.1) в координатах, мы получим систему уравнений степени  $m$ .

#### 5.4.4. Свойства отображения $d = i(v)$

Обсудим теперь некоторые свойства рассмотренного выше отображения  $d = i(v): \Lambda V^* \rightarrow \Lambda V^*$ . Это отображение обладает тем свойством,



**4.6.** Пусть тензоры  $z_1 = x_1 \wedge \cdots \wedge x_r$  и  $z_2 = y_1 \wedge \cdots \wedge y_r$  не пропорциональны,  $X = \langle x_1, \dots, x_r \rangle$  и  $Y = \langle y_1, \dots, y_r \rangle$ . Докажите, что пространство  $\langle z_1, z_2 \rangle$  состоит из разложимых кососимметрических тензоров тогда и только тогда, когда  $\dim(X \cap Y) = r - 1$ .

**4.7.** Пусть подпространство  $W \subset \Lambda^k V$  состоит из разложимых кососимметрических тензоров. Сопоставим каждому элементу  $w = x_1 \wedge \cdots \wedge x_k \in W$  подпространство  $[w] = \langle x_1, \dots, x_k \rangle \subset V$ . Докажите, что либо все подпространства  $[w]$  имеют общее  $(k-1)$ -мерное подпространство, либо все они лежат в одном  $(k+1)$ -мерном подпространстве.

**4.8.** Сопоставим подпространству  $\Lambda \subset \Lambda^k V$  подпространство  $\Lambda^\perp = \{v^* \mid i(v^*)\Lambda = 0\}$  в  $V^*$ . Докажите, что  $W = (\Lambda^\perp)^\perp$  является минимальным подпространством в  $V$ , для которого  $\Lambda$  лежит в подпространстве  $\Lambda^k W \subset \Lambda^k V$ .

## 5.5. Тензорный ранг

### 5.5.1. Тензорный ранг в $V \otimes W$

Пространство  $V \otimes W$  состоит из линейных комбинаций элементов  $v \otimes w$ , но не любой элемент этого пространства можно представить в виде  $v \otimes w$ . Рангом элемента  $T \in V \otimes W$  называют наименьшее число  $k$ , для которого  $T = v_1 \otimes w_1 + \cdots + v_k \otimes w_k$ ; ранг элемента  $T$  обозначим  $\text{rk } T$ .

**Теорема 5.5.1.1.** Если  $T = \sum a_{ij} e_i \otimes \varepsilon_j$ , где  $\{e_i\}$  и  $\{\varepsilon_i\}$  — базисы пространств  $V$  и  $W$ , то  $\text{rk } T = \text{rk } \|a_{ij}\|$ .

**Доказательство.** Пусть  $v_p = \sum \alpha_i^p e_i$ ,  $w_p = \sum \beta_j^p \varepsilon_j$ ,  $\alpha^p$  — столбец  $(\alpha_1^p, \dots, \alpha_n^p)$  и  $\beta^p$  — строка  $(\beta_1^p, \dots, \beta_m^p)$ . Тогда  $\|a_{ij}\| = \alpha^1 \beta^1 + \cdots + \alpha^k \beta^k$ . Наименьшее число  $k$ , для которого возможно такое разложение матрицы  $\|a_{ij}\|$ , равно рангу этой матрицы (см. §2.4.2).  $\square$

**Следствие 1.** Множество  $\{T \in V \otimes W \mid \text{rk } T \leq k\}$  замкнуто; в частности, если  $\lim_{i \rightarrow \infty} T_i = T$  и  $\text{rk } T_i \leq k$ , то  $\text{rk } T \leq k$ .

Замкнутость является следствием того, что это множество задаётся алгебраическими уравнениями.

**Следствие 2.** Ранг элемента вещественного пространства  $V \otimes W$  не изменяется при переходе к комплексификации.

Для элемента  $T \in V_1 \otimes \cdots \otimes V_p$  ранг можно определить аналогичным образом, т.е. как наименьшее число  $k$ , для которого  $T = v_1^1 \otimes \cdots \otimes v_p^1 + \cdots + v_1^k \otimes \cdots \otimes v_p^k$ . Оказывается, что при  $p \geq 3$  неверны свойства, сформулированные в следствиях 1 и 2. Но прежде чем перейти к изучению свойств тензорного ранга, расскажем, чем был вызван интерес к нему.

### 5.5.2. Алгоритм Штрассена

В пространстве матриц порядка  $n$  выберем базис  $e_{\alpha\beta} = \|\delta_{i\alpha}\delta_{j\beta}\|_1^n$ ; пусть  $\varepsilon_{\alpha\beta}$  — двойственный базис. Тогда  $A = \sum_{i,j} a_{ij}e_{ij}$ ,  $B = \sum_{i,j} b_{ij}e_{ij}$  и  $AB = \sum_{i,j,k} a_{ik}b_{kj}e_{ij} = \sum_{i,j,k} \varepsilon_{ik}(A)\varepsilon_{kj}(B)e_{ij}$ . Таким образом, вычисление произведения двух матриц порядка  $n$  сводится к вычислению  $n^3$  произведений линейных функций  $\varepsilon_{ik}(A)\varepsilon_{kj}(B)$ . Но является ли число  $n^3$  минимальным? Оказывается, что нет. Например, для матриц порядка 2 можно указать 7 пар линейных функций  $f_p$  и  $g_p$  и 7 матриц  $E_p$ , для которых  $AB = \sum_{p=1}^7 f_p(A)g_p(B)E_p$ . Это разложение было построено в работе [St1].

Вычисление наименьшего числа таких троек  $(f_p, g_p, E_p)$  эквивалентно вычислению ранга тензора  $\sum_{i,j,k} \varepsilon_{ik} \otimes \varepsilon_{kj} \otimes e_{ij} = \sum_p f_p \otimes g_p \otimes E_p$ .

Отождествим векторы и ковекторы и введём для краткости обозначения  $a = e_{11}$ ,  $b = e_{12}$ ,  $c = e_{21}$  и  $d = e_{22}$ . Тогда для матриц порядка 2 получаем  $\sum_{i,j,k} \varepsilon_{ik} \otimes \varepsilon_{kj} \otimes e_{ij} = (a \otimes a + b \otimes c) \otimes a + (a \otimes b + b \otimes d) \otimes b + (c \otimes a + d \otimes c) \otimes c + (c \otimes b + d \otimes d) \otimes d$ . Разложение Штрассена имеет вид  $\sum_{p=1}^7 \varepsilon_{ik} \otimes \varepsilon_{kj} \otimes e_{ij} = \sum_{p=1}^7 T_p$ , где  $T_1 = (a - d) \otimes (a - d) \otimes (a + d)$ ,  $T_2 = d \otimes (a + c) \otimes (a + c)$ ,  $T_3 = (a - b) \otimes d \otimes (a - b)$ ,  $T_4 = a \otimes (b + d) \otimes (b + d)$ ,  $T_5 = (c - d) \otimes a \otimes (c - d)$ ,  $T_6 = (b - d) \otimes (c + d) \otimes a$ ,  $T_7 = (c - a) \otimes (a + b) \otimes d$ . Оно приводит к следующему алгоритму вычисления произведения матриц  $A = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$  и  $B = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ :

$$\begin{aligned} S_1 &= a_1 - d_1, & S_2 &= a_2 - d_2, & S_3 &= a_1 - b_1, \\ S_4 &= b_1 - d_1, & S_5 &= c_2 + d_2, & S_6 &= a_2 + c_2, \\ S_7 &= b_2 + d_2, & S_8 &= c_1 - d_1, & S_9 &= c_1 - a_1, & S_{10} &= a_2 + b_2; \\ P_1 &= S_1 S_2, & P_2 &= S_3 d_2, & P_3 &= S_4 S_5, & P_4 &= d_1 S_6, \\ P_5 &= a_1 S_7, & P_6 &= S_8 a_2, & P_7 &= S_9 S_{10}; \\ S_{11} &= P_1 + P_2, & S_{12} &= S_{11} + P_3, & S_{13} &= S_{12} + P_4, \\ S_{14} &= P_5 - P_2, & S_{15} &= P_4 + P_6, & S_{16} &= P_1 + P_5, \\ S_{17} &= S_{16} - P_6, & S_{18} &= S_{17} + P_7. \end{aligned}$$

Тогда  $AB = \begin{pmatrix} S_{13} & S_{14} \\ S_{15} & S_{18} \end{pmatrix}$ . При вычислении  $AB$  по алгоритму Штрассена требуется 7 умножений и 18 сложений (или вычитаний).

### 5.5.3. Тензорный ранг в $V_1 \otimes V_2 \otimes V_3$

Пусть  $V$  — двумерное пространство с базисом  $\{e_1, e_2\}$ . Рассмотрим тензор  $T = e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2$ .

**Теорема 5.5.3.1.** *Ранг тензора  $T$  равен 3, но существует последовательность тензоров ранга не более 2, сходящаяся к  $T$ .*

**Доказательство.** Пусть  $T_\lambda = \lambda^{-1}[e_1 \otimes e_1 \otimes (-e_2 + \lambda e_1) + (e_1 + \lambda e_2) \otimes (e_1 + \lambda e_2) \otimes e_2]$ . Тогда  $T_\lambda - T = \lambda e_2 \otimes e_2 \otimes e_2$ , поэтому  $\lim_{\lambda \rightarrow 0} |T_\lambda - T| = 0$ .

Предположим, что  $T = a \otimes b \otimes c + u \otimes v \otimes w = (\alpha_1 e_1 + \alpha_2 e_2) \otimes b \otimes c + (\lambda_1 e_1 + \lambda_2 e_2) \otimes v \otimes w = e_1 \otimes (\alpha_1 b \otimes c + \lambda_1 v \otimes w) + e_2 \otimes (\alpha_2 b \otimes c + \lambda_2 v \otimes w)$ . Тогда  $e_1 \otimes e_1 + e_2 \otimes e_2 = \alpha_1 b \otimes c + \lambda_1 v \otimes w$  и  $e_1 \otimes e_2 = \alpha_2 b \otimes c + \lambda_2 v \otimes w$ . Следовательно, линейно независимые тензоры  $b \otimes c$  и  $v \otimes w$  ранга 1 принадлежат пространству  $\langle e_1 \otimes e_1 + e_2 \otimes e_2, e_1 \otimes e_2 \rangle$ . Последнее пространство можно отождествить с пространством матриц вида  $\begin{pmatrix} x & y \\ 0 & x \end{pmatrix}$ . Но все такие матрицы ранга 1 линейно зависимы. Получено противоречие.  $\square$

**Следствие.** *Подмножество тензоров ранга не более 2 в  $T_0^3(V)$  не замкнуто, поэтому его нельзя задать системой алгебраических уравнений.*

Рассмотрим теперь тензор  $T_1 = e_1 \otimes e_1 \otimes e_1 - e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2$ . Пусть  $\text{rk}_{\mathbb{R}} T_1$  — ранг тензора  $T_1$  над полем  $\mathbb{R}$ ,  $\text{rk}_{\mathbb{C}} T_1$  — ранг тензора  $T_1$  над полем  $\mathbb{C}$ .

**Теорема 5.5.3.2.**  $\text{rk}_{\mathbb{R}} T_1 \neq \text{rk}_{\mathbb{C}} T_1$ .

**Доказательство.** Легко проверить, что  $T_1 = (a_1 \otimes a_1 \otimes a_2 + a_2 \otimes a_2 \otimes a_1)/2$ , где  $a_1 = e_1 + ie_2$  и  $a_2 = e_1 - ie_2$ . Поэтому  $\text{rk}_{\mathbb{C}} T_1 \leq 2$ . Предположим теперь, что  $\text{rk}_{\mathbb{R}} T_1 \leq 2$ . Тогда, как и при доказательстве теоремы 5.5.3.1, получим, что линейно независимые тензоры  $b \otimes c$  и  $v \otimes w$  ранга 1 принадлежат пространству  $\langle e_1 \otimes e_1 + e_2 \otimes e_2, e_1 \otimes e_2 - e_2 \otimes e_1 \rangle$ , которое можно отождествить с пространством матриц вида  $\begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ . Но над полем  $\mathbb{R}$  среди таких матриц нет матриц ранга 1.  $\square$

### Задачи

**5.1.** Пусть  $U \subset V$  и  $T \in T_0^p(U) \subset T_0^p(V)$ . Докажите, что ранг тензора  $T$  не зависит от того, рассматривается ли он в  $T_0^p(U)$  или в  $T_0^p(V)$ .

**5.2.** Пусть  $e_1, \dots, e_k$  — линейно независимые векторы;  $e_i^{\otimes p} = e_i \otimes \dots \otimes e_i$ ,  $e_i \in T_0^p(V)$ , где  $p \geq 2$ . Докажите, что ранг тензора  $e_1^{\otimes p} + \dots + e_k^{\otimes p}$  равен  $k$ .

## 5.6. Линейные отображения пространств матриц

Тензорное произведение  $V_1 \otimes \dots \otimes V_p$  является линейным пространством; в этом пространстве есть ещё и дополнительная структура — каждому элементу можно сопоставить его ранг. Поэтому можно, например, рассмотреть линейные преобразования, переводящие тензоры ранга  $k$  в тензоры ранга  $k$ . Наиболее интересен случай отображения пространства  $\text{Hom}(V_1, V_2) = V_1^* \otimes V_2$  в себя. Отметим, также, что если  $\dim V_1 = \dim V_2 = n$ , то невырожденным отображениям из  $\text{Hom}(V_1, V_2)$  соответствуют тензоры ранга  $n$ , т.е. равенство  $\det A = 0$  можно интерпретировать на языке ранга тензоров.

### 5.6.1. Отображения, сохраняющие ранг 1

Если  $A: U \rightarrow U$  и  $B: V \rightarrow V$  — невырожденные линейные операторы, то линейный оператор  $T = A \otimes B: U \otimes V \rightarrow U \otimes V$  сохраняет ранг элементов  $U \otimes V$ . В случае, когда  $\dim U = \dim V$ , есть ещё один тип преобразований, сохраняющих ранг элементов  $U \otimes V$ . Возьмём произвольный изоморфизм  $\varphi: U \rightarrow V$  и зададим отображение  $S: U \otimes V \rightarrow U \otimes V$  формулой  $S(u \otimes v) = \varphi^{-1}v \otimes \varphi u$ . Тогда любое преобразование вида  $TS$ , где  $T = A \otimes B$  — преобразование первого типа, сохраняет ранг элементов  $U \otimes V$ .

**Замечание.** Легко проверить, что  $S$  — инволюция.

В матричном виде преобразование первого типа запишется как  $X \mapsto AXB$ , а преобразование второго типа как  $X \mapsto AX^T B$ . Преобразование второго типа не сводится к преобразованию первого типа (см. задачу 6.1).

**Теорема 5.6.1.1.** Пусть линейное отображение  $T: U \otimes V \rightarrow U \otimes V$  переводит любой элемент ранга 1 в элемент ранга 1. Тогда  $T = A \otimes B$  или  $T = (A \otimes B)S$ , причём второй случай возможен, только если  $\dim U = \dim V$ .

**Доказательство.** [ГЗ] Нам потребуется следующее утверждение.

**Лемма.** Пусть элементы  $\alpha_1, \alpha_2 \in U \otimes V$  таковы, что  $\text{rk}(t_1\alpha_1 + t_2\alpha_2) \leq 1$  для любых чисел  $t_1$  и  $t_2$ . Тогда  $\alpha_i$  можно представить в виде  $\alpha_i = u_i \otimes v_i$ , где  $u_1 = u_2$  или  $v_1 = v_2$ .

**Доказательство.** Предположим, что  $\alpha_i = u_i \otimes v_i$ ,  $\alpha_1 + \alpha_2 = u \otimes v$  и  $\langle u_1 \rangle \neq \langle u_2 \rangle$ ,  $\langle v_1 \rangle \neq \langle v_2 \rangle$ . Тогда можно считать, что  $\langle u \rangle \neq \langle u_1 \rangle$ . С одной стороны,  $(f \otimes g)(u \otimes v) = f(u)g(v)$ . С другой стороны,  $(f \otimes g)(u \otimes v) = (f \otimes g)(u_1 \otimes v_1 + u_2 \otimes v_2) = f(u_1)g(v_1) + f(u_2)g(v_2)$ . Поэтому, выбрав  $f \in U^*$  и  $g \in V^*$  так, что  $f(u) = 0$ ,  $f(u_1) \neq 0$  и  $g(v_2) = 0$ ,  $g(v_1) \neq 0$ , приходим к противоречию.  $\square$

В дальнейшем будем считать, что  $\dim V \geq \dim U \geq 2$ . Кроме того, для удобства фиксированные векторы будем обозначать  $a$  и  $b$ , а векторы, в качестве которых можно взять любые векторы пространств  $U$  и  $V$ , будем обозначать  $u$  и  $v$ . Применив лемму к  $T(a \otimes b_1)$  и  $T(a \otimes b_2)$ , где  $\langle b_1 \rangle \neq \langle b_2 \rangle$ , получим, что  $T(a \otimes b_i) = a' \otimes b'_i$  или  $T(a \otimes b_i) = a'_i \otimes b'$ . Так как  $T(a \otimes (\lambda b_1 - \mu b_2)) = 0$  только при  $\lambda = \mu = 0$ , то  $\langle b'_1 \rangle \neq \langle b'_2 \rangle$  (соответственно  $\langle a'_1 \rangle \neq \langle a'_2 \rangle$ ).

Легко проверить, что в первом случае  $T(a \otimes v) = a' \otimes v'$  для любого  $v \in V$ . Для доказательства достаточно применить лемму к  $T(a \otimes b_1)$  и  $T(a \otimes v)$ ,  $T(a \otimes b_2)$  и  $T(a \otimes v)$ . В самом деле, случай  $T(a \otimes v) = c' \otimes b'_1$ , где  $\langle c' \rangle \neq \langle a' \rangle$ , невозможен. Аналогично во втором случае  $T(a \otimes v) = f(v) \otimes b'$ , где  $f: V \rightarrow U$  — некоторое отображение (очевидно, линейное). Во втором случае подпространство  $a \otimes V$  мономорфно отображается в  $U \otimes b'$ , поэтому  $\dim V \leq \dim U$ , а значит,  $\dim U = \dim V$ .

Рассмотрим отображение  $T_1$ , равное  $T$  в первом случае и  $TS$  во втором случае. Тогда при фиксированном  $a$  справедливо равенство  $T_1(a \otimes v) = a' \otimes Bv$ , где  $B: V \rightarrow V$  — невырожденный оператор. Пусть  $\langle a_1 \rangle \neq \langle a \rangle$ . Тогда  $T_1(a_1 \otimes v) = a'_1 \otimes Bv$  или  $T_1(a_1 \otimes v) = a' \otimes B_1v$ , где  $\langle B_1 \rangle \neq \langle B \rangle$ . Применив лемму к  $T(a \otimes v)$  и  $T(u \otimes v)$ ,  $T(a_1 \otimes v)$  и  $T(u \otimes v)$ , получим, что в первом случае  $T_1(u \otimes v) = Au \otimes Bv$ , а во втором случае  $T_1(u \otimes v) = a' \otimes f(u, v)$ . Во втором случае пространство  $U \otimes V$  должно мономорфно отображаться в подпространство  $a' \otimes V$ ; этот случай невозможен.  $\square$

**Следствие.** Если линейное отображение  $T: U \otimes V \rightarrow U \otimes V$  любой элемент ранга 1 переводит в элемент ранга 1, то любой элемент ранга  $k$  оно переводит в элемент ранга  $k$ .

### 5.6.2. Отображения, сохраняющие определитель

Пусть  $M_{n,n}$  — пространство матриц порядка  $n$ ,  $T: M_{n,n} \rightarrow M_{n,n}$  — линейное отображение.

**Теорема 5.6.2.1.** *Если  $T$  сохраняет определитель матриц, то  $T$  сохраняет и ранг матриц.*

**Доказательство.** [Ма3] Для удобства введём следующие обозначения:  $I_r$  и  $0_r$  — единичная и нулевая матрицы порядка  $r$ ;  $A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ .

Докажем сначала, что если отображение  $T$  сохраняет определитель, то  $T$  невырожденное. Предположим, что  $T(A) = 0$  и  $A \neq 0$ . Тогда  $0 < \text{rk } A < n$ . Существуют такие невырожденные матрицы  $M$  и  $N$ , что  $MAN = I_r \oplus 0_{n-r}$ , где  $r = \text{rk } A$  (см. теорему 2.2.3.2). Для любой матрицы  $X$  порядка  $n$  выполняется равенство  $|MAN + X| \cdot |MN|^{-1} = |A + M^{-1}XN^{-1}| = |T(A + M^{-1}XN^{-1})| = |T(M^{-1}XN^{-1})| = |X| \cdot |MN|^{-1}$ . Следовательно,  $|MAN + X| = |X|$ . Положив  $X = 0_r \oplus I_{n-r}$ , приходим к противоречию.

Пусть  $\text{rk } A = r$  и  $\text{rk } T(A) = s$ . Тогда существуют такие невырожденные матрицы  $M_1, N_1$  и  $M_2, N_2$ , что  $M_1AN_1 = I_r \oplus 0_{n-r} = Y_1$  и  $M_2T(A)N_2 = I_s \oplus 0_{n-s} = Y_2$ . Рассмотрим отображение  $f: M_{n,n} \rightarrow M_{n,n}$ , заданное формулой  $f(X) = M_2T(M_1^{-1}XN_1^{-1})N_2$ . Это отображение линейно и  $|f(X)| = k|X|$ , где  $k = |M_2M_1^{-1}N_1^{-1}N_2|$ ; кроме того,  $f(Y_1) = M_2T(A)N_2 = Y_2$ . Рассмотрим матрицу  $Y_3 = 0_r \oplus I_{n-r}$ . Тогда  $|\lambda Y_1 + Y_3| = \lambda^r$  для всех  $\lambda$ . С другой стороны,  $|f(\lambda Y_1 + Y_3)| = |\lambda Y_2 + f(Y_3)| = p(\lambda)$ , где  $p$  — многочлен степени не более  $s$ . Следовательно,  $r \leq s$ . А так как  $|B| = |TT^{-1}(B)| = |T^{-1}(B)|$ , то отображение  $T^{-1}$  тоже сохраняет определитель. Поэтому  $s \leq r$ .  $\square$

### 5.6.3. Отображения, сохраняющие собственные значения

Будем говорить, что линейное отображение  $T: M_{n,n} \rightarrow M_{n,n}$  сохраняет собственные значения, если наборы собственных значений матриц  $X$  и  $T(X)$  совпадают для любой матрицы  $X$ .

**Теорема 5.6.3.1.** *а) Если отображение  $T$  сохраняет собственные значения, то  $T(X) = AXA^{-1}$  или  $T(X) = AX^T A^{-1}$ .*

*б) Если отображение  $T$ , определённое над  $\mathbb{C}$ , сохраняет собственные значения эрмитовых матриц, то  $T(X) = AXA^{-1}$  или  $T(X) = AX^T A^{-1}$ .*

**Доказательство.** [Ма3] а) Если  $T$  сохраняет собственные значения, то  $T$  сохраняет ранг, поэтому  $T(X) = AXB$  или  $T(X) = AX^T B$  (см. §5.6.1).



Остаётся доказать, что  $T(I) = I$ . Определитель матрицы равен произведению её собственных значений, поэтому  $T$  сохраняет определитель. Следовательно,  $|X - \lambda I| = |T(X) - \lambda T(I)| = |CT(X) - \lambda I|$ , где  $C = T(I)^{-1}$ , а значит, собственные значения матриц  $X$  и  $CT(X)$  совпадают; кроме того, собственные значения матриц  $X$  и  $T(X)$  совпадают по условию. Отображение  $T$  невырожденное (см. доказательство теоремы 5.6.2.1), поэтому любую матрицу  $Y$  можно представить в виде  $T(X)$ , а значит, собственные значения матриц  $Y$  и  $CY$  совпадают. Матрицу  $C$  можно представить в виде  $C = SU$ , где  $U$  — унитарная матрица,  $S$  — эрмитова положительно определённая матрица. Собственные значения матриц  $U^{-1}$  и  $CU^{-1} = S$  совпадают, но собственные значения унитарной матрицы  $U^{-1}$  имеют вид  $e^{i\varphi}$ , а собственные значения матрицы  $S$  положительны. Следовательно,  $S = U = I$  и  $C = I$ , т.е.  $T(I) = I$ .

б) Достаточно доказать, что если  $T$  сохраняет собственные значения эрмитовых матриц, то  $T$  сохраняет собственные значения всех матриц. Любую матрицу  $X$  можно представить в виде  $X = P + iQ$ , где  $P$  и  $Q$  — эрмитовы матрицы. Для любого действительного  $x$  матрица  $A = P + xQ$  эрмитова. Если собственные значения матрицы  $A$  равны  $\lambda_1, \dots, \lambda_n$ , то собственные значения матрицы  $A^m$  равны  $\lambda_1^m, \dots, \lambda_n^m$ . Поэтому  $\operatorname{tr} A^m = \operatorname{tr}(T(A)^m)$ ; обе части этого равенства являются многочленами от  $x$  степени не более  $m$ . Если два многочлена принимают равные значения при всех действительных  $x$ , то они совпадают, а потому их значения при  $x = i$  тоже равны. Следовательно,  $\operatorname{tr} X^m = \operatorname{tr}(T(X)^m)$  для любой матрицы  $X$ . Остаётся воспользоваться результатом задачи 3.2.  $\square$

#### 5.6.4. Отображения, сохраняющие невырожденность

**Теорема 5.6.4.1.** а) Пусть  $T: M_{n,n} \rightarrow M_{n,n}$  — линейное отображение, переводящее невырожденные матрицы в невырожденные. Тогда  $T$  — невырожденное отображение.

б) Если, кроме того,  $T(I) = I$ , то  $T$  сохраняет собственные значения.

**Доказательство.** [Ма4] а) Если  $|T(X)| = 0$ , то  $|X| = 0$ . Для  $X = A - \lambda I$  получаем, что если  $|T(A - \lambda I)| = |T(I)||T(I)^{-1}T(A) - \lambda I| = 0$ , то  $|A - \lambda I| = 0$ . Поэтому собственные значения матрицы  $T(I)^{-1}T(A)$  являются собственными значениями матрицы  $A$ .

Предположим, что  $A \neq 0$  и  $T(A) = 0$ . Для матрицы  $A$  можно найти такую матрицу  $X$ , что матрицы  $X$  и  $X + A$  не имеют общих собственных значений (см. задачу 5.2), а значит, матрицы  $T(I)^{-1}T(A + X)$  и  $T(I)^{-1}T(X)$  не имеют общих собственных значений. С другой стороны,

эти матрицы совпадают, так как  $T(A + X) = T(X)$ . Получено противоречие.

б) Если  $T(I) = I$ , то из доказательства а) следует, что собственные значения  $T(A)$  являются собственными значениями  $A$ . Поэтому если собственные значения матрицы  $B = T(A)$  не кратные, то собственные значения матрицы  $B$  совпадают с собственными значениями матрицы  $A = T^{-1}(B)$ . Для матрицы  $B$  с кратными собственными значениями можно рассмотреть сходящуюся к ней последовательность матриц  $B_i$  с не кратными собственными значениями (см. теорему 8.4.5.2) и заметить, что собственные значения матриц  $B_i$  сходятся к собственным значениям матрицы  $B$  (см. задачу 1.6).  $\square$

## Задачи

**6.1.** Пусть  $X$  — матрица размера  $m \times n$ , где  $mn > 1$ . Докажите, что преобразование  $X \mapsto X^T$  нельзя представить в виде  $X \mapsto AXB$ .

**6.2.** Пусть  $f: M_{n,n} \rightarrow M_{n,n}$  — невырожденное отображение, причём  $f(XY) = f(X)f(Y)$  для любых матриц  $X$  и  $Y$ . Докажите, что  $f(X) = AXA^{-1}$ , где  $A$  — фиксированная матрица.

## 5.7. Образ полилинейного отображения

Любое полилинейное отображение  $f$  представляется в виде композиции

$$V_1 \times \dots \times V_k \xrightarrow{p} V_1 \otimes \dots \otimes V_k \xrightarrow{\tilde{f}} V,$$

где  $p(v_1, \dots, v_k) = v_1 \otimes \dots \otimes v_k$ . Образ отображения  $p$  отнюдь не является линейным подпространством. Например, если  $V_1 = V_2 = \mathbb{R}^2$ , то образ отображения  $p$  устроен так. Пусть  $(x, y)$  — координаты в  $V_1$ ,  $(z, t)$  — координаты в  $V_2$ . Координаты вектора  $(x, y) \otimes (z, t) = (xz, xt, yz, yt)$  удовлетворяют соотношению  $x_1x_4 = x_2x_3$ . Легко также проверить, что любой вектор в  $\mathbb{R}^4 \cong \mathbb{R}^2 \otimes \mathbb{R}^2$ , координаты которого удовлетворяют соотношению  $x_1x_4 = x_2x_3$ , лежит в образе отображения  $p$ . Действительно, если одна из координат отлична от нуля, например,  $x_1 \neq 0$ , то мы полагаем  $x = 1$ ,  $y = x_3/x_1$ ,  $z = x_1$  и  $t = x_2$ . Если же все координаты равны нулю, то мы полагаем  $x = y = z = t = 0$ .

Итак, образ полилинейного отображения  $f$  является образом довольно сложно устроенного подмножества при линейном отображении

$\tilde{f}$ . Поэтому он, вообще говоря, устроен сложно. В частности, он не всегда является линейным подпространством. Приводимые в этом параграфе сведения об образе полилинейного отображения взяты из статьи [Be4].

### 5.7.1. Два примера

**Пример 5.7.1.1.** Образ билинейного отображения  $f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , заданного формулой

$$f((x_1, x_2), (y_1, y_2)) = (x_1 y_1, x_1 y_2, x_2 y_1),$$

представляет собой пространство  $\mathbb{R}^3$ , из которого в координатной плоскости  $x = 0$  выброшено всё, кроме осей координат (заданных уравнением  $yz = 0$ ).

**Доказательство.** Если  $x \neq 0$ , то система уравнений  $x_1 y_1 = x$ ,  $x_1 y_2 = y$ ,  $x_2 y_1 = z$  всегда совместна: достаточно положить  $x_1 = 1$ ,  $y_1 = x$ ,  $y_2 = y$  и  $x_2 = z/x$ . Если  $x = 0$ , то эта система уравнений совместна тогда и только тогда, когда  $yz = 0$ .  $\square$

**Пример 5.7.1.2.** Образ билинейного отображения  $f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , заданного формулой

$$f((x_1, x_2), (y_1, y_2)) = (x_1 y_1 + x_2 y_2, x_1 y_2, x_2 y_1),$$

представляет собой внешнюю часть конуса:  $x^2 \geq 4yz$ .

**Доказательство.** Если система  $x_1 y_1 + x_2 y_2 = x$ ,  $x_1 y_2 = y$ ,  $x_2 y_1 = z$  совместна, то  $x^2 - 4yz = (x_1 y_1 - x_2 y_2)^2 \geq 0$ . Предположим теперь, что  $x^2 \geq 4yz$ . Если  $y = 0$ , то можно положить  $x_1 = 0$ ,  $y_1 = z$ ,  $x_2 = 1$ ,  $y_2 = x$ . Если же  $y \neq 0$ , то можно положить  $x_1 = 1$ ,  $y_1 = \frac{x+\Delta}{2}$ ,  $x_2 = \frac{x-\Delta}{2y}$ ,  $y_2 = y$ , где  $\Delta = \sqrt{x^2 - 4yz}$ .  $\square$

Этими двумя примерами в действительности исчерпываются образы билинейных отображений в  $\mathbb{R}^3$ , которые не являются подпространствами (см. [Be4]).

### 5.7.2. Образ отображения в 2-мерное подпространство

Образ полилинейного отображения в 1-мерное подпространство, очевидно, является линейным подпространством. Оказывается, что это

верно и для полилинейных отображений в 2-мерное пространство. Это следует из теоремы 5.7.2.1, которую мы сейчас сформулируем и докажем.

Если фиксировать номер  $i$  ( $1 \leq i \leq k$ ) и векторы  $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k$ , то полилинейному отображению  $f : V_1 \times \dots \times V_k \rightarrow V$  можно сопоставить ассоциированное линейное отображение  $f_i : V_i \rightarrow V$ , заданное формулой

$$f_i(v) = f(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k).$$

**Теорема 5.7.2.1.** Пусть полилинейное отображение  $f : V_1 \times \dots \times V_k \rightarrow V$  таково, что размерность образа любого ассоциированного линейного отображения не превосходит 1. Тогда образ отображения  $f$  является линейным подпространством, размерность которого не превосходит 1.

**Доказательство.** Пусть  $f(u_1, \dots, u_k) \neq 0$  и  $f(v_1, \dots, v_k) \neq 0$ . Тогда в силу линейности в каждой строке таблицы

$f(u_1, u_2, \dots, u_k)$	$f(v_1, u_2, \dots, u_k)$	$f(u_1 + v_1, u_2, \dots, u_k)$
$f(u_1, v_2, \dots, v_k)$	$f(v_1, v_2, \dots, v_k)$	$f(u_1 + v_1, v_2, \dots, v_k)$

не более одного нулевого вектора, поэтому найдётся столбец с ненулевыми векторами. Следовательно, можно выбрать  $t_1 \in V_1$  так, что  $f(t_1, u_2, \dots, u_k) \neq 0$  и  $f(t_1, v_2, \dots, v_k) \neq 0$ . Пространство  $\langle f(u_1, u_2, \dots, u_k) \rangle$  имеет размерность 1, поэтому  $\langle f(u_1, u_2, \dots, u_k) \rangle = \langle f(t_1, u_2, \dots, u_k) \rangle$ . Аналогично  $\langle f(v_1, v_2, \dots, v_k) \rangle = \langle f(t_1, v_2, \dots, v_k) \rangle$ . Затем аналогично выбираем  $t_2 \in V_2$  так, что  $\langle f(t_1, u_2, \dots, u_k) \rangle = \langle f(t_1, t_2, u_3, \dots, u_k) \rangle$  и  $\langle f(t_1, v_2, \dots, v_k) \rangle = \langle f(t_1, t_2, v_3, \dots, v_k) \rangle$ . Продолжая такие рассуждения, получаем

$$\langle f(u_1, u_2, \dots, u_k) \rangle = \langle f(t_1, t_2, \dots, t_k) \rangle = \langle f(v_1, v_2, \dots, v_k) \rangle,$$

что и требовалось.  $\square$

**Следствие.** Если  $\dim V = 2$ , то образ полилинейного отображения  $f : V_1 \times \dots \times V_k \rightarrow V$  является линейным подпространством.

**Доказательство.** Если размерность хотя бы одного ассоциированного линейного отображения равна 2, то образ полилинейного отображения  $f$  совпадает с  $V$ . Если же размерности всех ассоциированных линейных отображений меньше 2, то можно применить теорему 5.7.2.1.  $\square$

## Решения

**1.1.** Дополним векторы  $v$  и  $w$  до базисов пространств  $V$  и  $W$ . Если  $v' \otimes w' = v \otimes w$ , то разложения векторов  $v'$  и  $w'$  по этим базисам имеют вид  $\lambda v$  и  $\mu w$ . Ясно также, что  $\lambda v \otimes \mu w = \lambda \mu (v \otimes w)$ , т.е.  $\mu = 1/\lambda$ .

1.2. а) Утверждение очевидным образом следует из определения.

б) Возьмём базисы пространств  $\text{Im } A_1$  и  $\text{Im } A_2$  и дополним их до базисов  $\{e_i\}$  и  $\{\varepsilon_j\}$  пространств  $W_1$  и  $W_2$ . Пространство  $\text{Im } A_1 \otimes W_2$  порождено векторами  $e_i \otimes \varepsilon_j$ , где  $e_i \in \text{Im } A_1$ , а пространство  $W_1 \otimes \text{Im } A_2$  порождено векторами  $e_i \otimes \varepsilon_j$ , где  $\varepsilon_j \in \text{Im } A_2$ . Поэтому пространство  $(\text{Im } A_1 \otimes W_2) \cap (W_1 \otimes \text{Im } A_2)$  порождено векторами  $e_i \otimes \varepsilon_j$ , где  $e_i \in \text{Im } A_1$  и  $\varepsilon_j \in \text{Im } A_2$ , т.е. это пространство совпадает с  $\text{Im } A_1 \otimes \text{Im } A_2$ .

в) Возьмём базисы пространства  $\text{Ker } A_1$  и  $\text{Ker } A_2$  и дополним их до базисов  $\{e_i\}$  и  $\{\varepsilon_j\}$  пространств  $V_1$  и  $V_2$ . Отображение  $A_1 \otimes A_2$  переводит элемент  $e_i \otimes \varepsilon_j$  в нуль, если  $e_i \in \text{Ker } A_1$  или  $\varepsilon_j \in \text{Ker } A_2$ ; остальные элементы вида  $e_i \otimes \varepsilon_j$  это отображение переводит в базис пространства  $\text{Im } A_1 \otimes \text{Im } A_2$ , т.е. в линейно независимые элементы.

1.3. Выберем в пространстве  $V_1 \cap V_2$  базис  $\{v_i\}$  и дополним его до базисов  $\{v_j^1\}$  и  $\{v_k^2\}$  пространств  $V_1$  и  $V_2$ . В результате получим базис  $\{v_i, v_j^1, v_k^2\}$  пространства  $V_1 + V_2$ . Аналогично построим базис  $\{w_\alpha, w_\beta^1, w_\gamma^2\}$  пространства  $W_1 + W_2$ . Тогда  $\{v_i \otimes w_\alpha, v_i \otimes w_\gamma^2, v_k^2 \otimes w_\alpha, v_k^2 \otimes w_\gamma^2\}$  — базисы пространств  $V_1 \otimes W_1$  и  $V_2 \otimes W_2$ , причём элементы этих базисов содержатся в базисе пространства  $(V_1 + V_2) \otimes (W_1 + W_2)$ , т.е. они в совокупности линейно независимы. Следовательно,  $\{v_i \otimes w_\alpha\}$  — базис пространства  $(V_1 \otimes W_1) \cap (V_2 \otimes W_2)$ .

1.4. Пусть  $B: V \rightarrow V_1$  и  $D: W \rightarrow W_2$ . Достаточно проверить, что указанные операторы одинаково действуют на тензоры вида  $v \otimes w$ , где  $v \in V$  и  $w \in W$ . Ясно, что

$$\begin{aligned}(A \otimes C)(B \otimes D)(v \otimes w) &= (A \otimes C)(Bv \otimes Dw) = \\ &= (ABv \otimes CDw) = ((AB) \otimes (CD))(v \otimes w).\end{aligned}$$

1.5. Требуется доказать, что  $a_{ij}B = a_{ji}B^T$ . Для симметрических матриц это очевидно, а для кососимметрических матриц это следует из того, что  $a_{ij} = -a_{ji}$  и  $B^T = -B$ .

1.6. а) Пусть у матриц  $A$  и  $B$  количества положительных собственных значений равны  $p$  и  $p'$ , а отрицательных —  $q$  и  $q'$ . Тогда у матрицы  $A \otimes B$  есть  $pq + p'q'$  положительных собственных значений и  $pq' + p'q$  отрицательных. Поэтому её сигнатура равна  $pq + p'q' - pq' - p'q = (p - q)(p' - q')$ .

б) С помощью замен  $A \mapsto X^T A X$  и  $B \mapsto Y^T B Y$  матрицы  $A$  и  $B$  можно привести к блочно диагональному виду с блоками  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  (и, возможно, нулями).

Поэтому матрицу  $A \otimes B$  можно привести к блочно диагональному виду с бло-

ками  $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$  (и нулями). Легко проверить, что сигнатура каждого

блока равна нулю. Действительно, квадратичная форма, соответствующая такому блоку, имеет вид  $2x_1x_4 - 2x_2x_3$ ; каждое из этих двух слагаемых имеет нулевую сигнатуру.

**1.7.** Ясно, что  $Ax = x - 2(a, x)$ , т.е.  $Aa = -a$  и  $Ax = x$  для  $x \in a^\perp$ .

**1.8.** Фиксируем  $a \neq 0$ ;  $A(a, x)$  — линейная функция, поэтому  $A(a, x) = (b, x)$ , причём  $b = B(a)$ , где  $B$  — некоторое линейное отображение. Если  $x \perp a$ , то  $A(a, x) = 0$ , т.е.  $(b, x) = 0$ . Поэтому  $a^\perp \subset b^\perp$ , а значит,  $B(a) = b = \lambda(a)$ . А так как  $A(u + v, x) = A(u, x) + A(v, x)$ , то  $\lambda(u + v)(u + v) = \lambda(u)u + \lambda(v)v$ . Если векторы  $u$  и  $v$  линейно независимы, то  $\lambda(u) = \lambda(v) = \lambda$ ; любой другой вектор  $w$  линейно независим с одним из векторов  $u$  и  $v$ , поэтому  $\lambda(w) = \lambda$ . Для одномерного пространства утверждение очевидно.

**1.9.** Ясно, что  $T_p^q(V) = (V^*)^{\otimes p} \otimes V^{\otimes q}$  и  $(T_p^q(V))^* = (V^*)^{\otimes q} \otimes V^{\otimes p}$ . Поэтому

$$\begin{aligned} T_{p'}^{q'}(T_p^q(V)) &= ((V^*)^{\otimes q} \otimes V^{\otimes p})^{\otimes p'} \otimes ((V^*)^{\otimes p} \otimes V^{\otimes q})^{\otimes q'} = \\ &= (V^*)^{\otimes qp'} \otimes V^{\otimes pp'} \otimes (V^*)^{\otimes pq'} \otimes V^{\otimes qq'} = \\ &= (V^*)^{\otimes pq' + qp'} \otimes V^{\otimes pp' + qq'} = T_{pq' + qp'}^{pp' + qq'}(V). \end{aligned}$$

**2.1.** Будем поочерёдно менять местами два первых аргумента и два последних:  $f(x, y, z) = f(y, x, z) = -f(y, z, x) = -f(z, y, x) = f(z, x, y) = f(x, z, y) = -f(x, y, z)$ , поэтому  $2f(x, y, z) = 0$ .

**2.2.** Продолжим отображение  $f$  до билинейного отображения  $\mathbb{C}^m \times \mathbb{C}^m \rightarrow \mathbb{C}^n$ . Рассмотрим уравнение  $f(z, z) = 0$ , т.е. систему квадратных уравнений  $f_1(z, z) = 0, \dots, f_n(z, z) = 0$ . Предположим, что  $n < m$ . Тогда эта система имеет ненулевое решение  $z = x + iy$ . Из второго условия следует, что  $y \neq 0$ . Ясно также, что  $0 = f(z, z) = f(x + iy, x + iy) = f(x, x) - f(y, y) + 2if(x, y)$ , поэтому  $f(x, x) = f(y, y) \neq 0$  и  $f(x, y) = 0$ . Это противоречит первому условию.

**2.3.** Элементы  $\alpha_i = e_{2i-1} \wedge e_{2i}$  лежат в  $\Lambda^2(V)$ , поэтому  $\alpha_i \wedge \alpha_j = \alpha_j \wedge \alpha_i$  и  $\alpha_i \wedge \alpha_i = 0$ . Следовательно,  $\Lambda^n \omega = \sum_{i_1, \dots, i_n} \alpha_{i_1} \wedge \dots \wedge \alpha_{i_n} = n! \alpha_1 \wedge \dots \wedge \alpha_n = n! e_1 \wedge \dots \wedge e_{2n}$ .

**2.4.** Пусть на диагонали жордановой нормальной формы матрицы  $A$  стоят числа  $\lambda_1, \dots, \lambda_l$ . Тогда

$$|A - \lambda I| = (\lambda_1 - \lambda) \dots (\lambda_n - \lambda) = \sum_{k=0}^n (-1)^k \sigma_{n-k}(\lambda_1, \dots, \lambda_n) \lambda^k.$$

Если  $e_1, \dots, e_n$  — жорданов базис, то матрица  $\Lambda^{n-k} A$  в базисе  $e_{i_1} \wedge \dots \wedge e_{i_{n-k}}$ , упорядоченном лексикографически, треугольная (теорема 5.2.5.2); на диагонали этой матрицы стоят числа  $\lambda_{i_1} \dots \lambda_{i_{n-k}}$ , поэтому

$$\text{tr}(\Lambda^{n-k} A) = \sum_{i_1 < \dots < i_{n-k}} \lambda_{i_1} \dots \lambda_{i_{n-k}} = \sigma_{n-k}(\lambda_1, \dots, \lambda_n).$$

**2.5.** На всё пространство это скалярное произведение распространяется по линейности; его симметричность очевидна. Поэтому нужно лишь проверить, что если  $x \wedge y \neq 0$ , то

$$\begin{pmatrix} (x, z) & (x, t) \\ (y, z) & (y, t) \end{pmatrix} = (x, x)(y, y) - (x, y)^2 > 0.$$

Это неравенство очевидно.

На  $\Lambda^k V$  скалярное произведение задаётся формулой

$$(x_1 \wedge \dots \wedge x_k, y_1 \wedge \dots \wedge y_k) = \det \begin{pmatrix} (x_1, y_1) & \dots & (x_1, y_k) \\ \vdots & \ddots & \vdots \\ (x_k, y_1) & \dots & (x_k, y_k) \end{pmatrix}.$$

При  $x_1 = y_1, \dots, x_k = y_k$  этот определитель является определителем Грама, поэтому если  $x_1 \wedge \dots \wedge x_k \neq 0$ , то он положителен.

**2.6.** Если  $A = \|a_{ij}\|_1^n$ , то матрица рассматриваемой системы уравнений равна  $S^2(A)$ . Кроме того,  $\det S^2(A) = (\det A)^r$ , где  $r = \frac{2}{n} \binom{n+2-1}{2} = n+1$  (см. теорему 5.2.5.3).

**2.7.** Легко проверить, что  $\sigma_k = \text{tr}(\Lambda^k A)$ . Если в жордановом базисе диагональ матрицы  $A$  имеет вид  $(\lambda_1, \dots, \lambda_n)$ , то  $s_k = \lambda_1^k + \dots + \lambda_n^k$  и  $\sigma_k = \sum \lambda_{i_1} \dots \lambda_{i_k}$ . Требуемое тождество для функций  $s_k$  и  $\sigma_k$  доказано в §1.4.1.

**2.8.** Пусть  $e_j$  и  $\varepsilon_j$ , где  $1 \leq j \leq m$ , — двойственные базисы. Величину  $n! \langle v_1 \wedge \dots \wedge v_n, f_1 \wedge \dots \wedge f_n \rangle$ , где  $v_i = \sum a_{ij} e_j$  и  $f_i = \sum b_{ji} \varepsilon_j$ , можно вычислить двумя способами. С одной стороны, она равна

$$\begin{vmatrix} f_1(v_1) & \dots & f_n(v_1) \\ \dots & \dots & \dots \\ f_1(v_n) & \dots & f_n(v_n) \end{vmatrix} = \begin{vmatrix} \sum a_{1j} b_{j1} & \dots & \sum a_{1j} b_{jn} \\ \dots & \dots & \dots \\ \sum a_{nj} b_{j1} & \dots & \sum a_{nj} b_{jn} \end{vmatrix} = \det AB.$$

С другой стороны, она равна

$$\begin{aligned} n! \left\langle \sum_{k_1 \dots k_n} a_{1k_1} \dots a_{nk_n} e_{k_1} \wedge \dots \wedge e_{k_n}, \right. \\ \left. \sum_{l_1 \dots l_n} b_{l_1 1} \dots b_{l_n n} \varepsilon_{l_1} \wedge \dots \wedge \varepsilon_{l_n} \right\rangle = \\ = \sum_{k_1 \leq \dots \leq k_n} A_{k_1 \dots k_n} B^{l_1 \dots l_n} n! \langle e_{k_1} \wedge \dots \wedge e_{k_n}, \varepsilon_{l_1} \wedge \dots \wedge \varepsilon_{l_n} \rangle = \\ = \sum_{k_1 < \dots < k_n} A_{k_1 \dots k_n} B^{l_1 \dots l_n}. \end{aligned}$$

**2.9.** Легко проверить, что следующие формы образуют ортогональный базис для  $Q$ :

$$\begin{aligned} \omega_1 &= \frac{e_1 \wedge e_2 + e_3 \wedge e_4}{\sqrt{2}}, & \omega_4 &= \frac{e_1 \wedge e_2 - e_3 \wedge e_4}{\sqrt{2}}, \\ \omega_2 &= \frac{e_1 \wedge e_3 - e_2 \wedge e_4}{\sqrt{2}}, & \omega_5 &= \frac{e_1 \wedge e_3 + e_2 \wedge e_4}{\sqrt{2}}, \\ \omega_3 &= \frac{e_1 \wedge e_4 + e_2 \wedge e_3}{\sqrt{2}}, & \omega_6 &= \frac{e_1 \wedge e_4 - e_2 \wedge e_3}{\sqrt{2}}. \end{aligned}$$

При этом  $Q(\omega_i, \omega_i) = 1$  при  $i \leq 3$  и  $Q(\omega_i, \omega_i) = -1$  при  $i \geq 4$ . Таким образом, сигнатура формы  $Q$  равна  $(3, 3)$ .

**3.1.** Так как  $\text{Pf}(A) = \sum (-1)^\sigma a_{i_1 i_2} \dots a_{i_{2n-1} i_{2n}}$ , причём суммирование ведётся во всем разбиениям множества  $\{1, \dots, 2n\}$  на пары  $\{i_{2k-1}, i_{2k}\}$ , где  $i_{2k-1} < i_{2k}$ , то  $a_{i_1 i_2} C_{i_1 i_2} = a_{i_1 i_2} \sum (-1)^\sigma a_{i_3 i_4} \dots a_{i_{2n-1} i_{2n}}$ . Остаётся заметить, что знаки перестановок  $\sigma = \begin{pmatrix} 1 & 2 & \dots & 2n \\ i_1 & i_2 & \dots & i_{2n} \end{pmatrix}$  и  $\tau = \begin{pmatrix} i_1 & i_2 & 1 & 2 & \dots & i_1 & i_2 & \dots & 2n \\ i_1 & i_2 & i_3 & i_4 & \dots & \dots & \dots & i_{2n} \end{pmatrix}$  отличаются на  $(-1)^{i_1+i_2+1}$ . В самом деле, для перемещения элемента  $i_1$  в верхней строке  $\sigma$  на первое место требуется  $i_1 - 1$  транспозиций, а для перемещения  $i_2$  на второе место требуется  $i_2 - 2$  транспозиций.

**3.2.** Легко проверить, что  $G = XJX^T$ , где  $J = \text{diag} \left( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$ . Поэтому  $\text{Pf}(G) = \det X$ .

**4.1.** Ясно, что если  $\omega = \omega_1 \wedge e_1 \wedge \dots \wedge e_r$ , то  $\omega \wedge e_i = 0$ . Предположим теперь, что  $\omega \wedge e_i = 0$  при  $i = 1, \dots, r$  и  $e_1 \wedge \dots \wedge e_r \neq 0$ . Дополним векторы  $e_1, \dots, e_r$  до базиса  $e_1, \dots, e_n$  пространства  $V$ . Тогда  $\omega = \sum a_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k}$ , причём  $\sum a_{i_1 \dots i_k} e_{i_1} \wedge \dots \wedge e_{i_k} \wedge e_i = \omega \wedge e_i = 0$  при  $i = 1, \dots, r$ . Если ненулевые тензоры  $e_{i_1} \wedge \dots \wedge e_{i_k} \wedge e_i$  линейно зависимы, то тензоры  $e_{i_1} \wedge \dots \wedge e_{i_k}$  тоже линейно зависимы. Поэтому  $a_{i_1 \dots i_k} = 0$  при  $i \notin \{i_1, \dots, i_k\}$ . Следовательно,  $a_{i_1 \dots i_k} \neq 0$ , только если  $\{1, \dots, r\} \subset \{i_1, \dots, i_k\}$ , а значит,

$$\omega = \left( \sum b_{i_1 \dots i_{k-r}} e_{i_1} \wedge \dots \wedge e_{i_{k-r}} \right) \wedge e_1 \wedge \dots \wedge e_r.$$

**4.2.** Пусть  $e_1, \dots, e_n$  — базис  $V$ ,  $\varepsilon_1, \dots, \varepsilon_n$  — двойственный базис  $V^*$ . Можно считать, что  $\omega = e_1 \wedge \dots \wedge e_n$ . Пусть  $\alpha = \varepsilon_{i_1} \wedge \dots \wedge \varepsilon_{i_p}$ , где  $i_1 < \dots < i_p$ . Тогда

$$\begin{aligned} & \left\langle i(\alpha)\omega, \sum_{j_1 < \dots < j_p} x_{j_1 \dots j_{n-p}} \varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_{n-p}} \right\rangle = \\ & = \left\langle e_1 \wedge \dots \wedge e_n, \sum_{j_1 < \dots < j_p} x_{j_1 \dots j_{n-p}} \varepsilon_{i_1} \wedge \dots \wedge \varepsilon_{i_p} \wedge \varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_{n-p}} \right\rangle = \\ & = \frac{1}{n!} x_{j_1 \dots j_{n-p}}, \end{aligned}$$

где набор индексов  $j_1, \dots, j_{n-p}$  дополняет набор  $i_1, \dots, i_p$ . Таким образом,  $i(\varepsilon_{i_1} \wedge \dots \wedge \varepsilon_{i_p})\omega = \varepsilon_{j_1} \wedge \dots \wedge \varepsilon_{j_{n-p}}$ . Ясно, что такое отображение является изоморфизмом.

Любой разложимый элемент  $\alpha$  можно представить в виде  $\alpha = \varepsilon_1 \wedge \dots \wedge \varepsilon_p$ . Дополним векторы  $\varepsilon_1, \dots, \varepsilon_p$  до базиса  $\varepsilon_1, \dots, \varepsilon_n$ . Пусть  $e_1, \dots, e_n$  — двойственный базис. Если  $\omega = e_1 \wedge \dots \wedge e_n$ , то  $i(\alpha)\omega = e_{p+1} \wedge \dots \wedge e_n$ .

**4.3.** Первое решение. Согласно задаче 4.2 пространство  $\Lambda^{n-1}V$  изоморфно  $\Lambda^1 V^* \cong V^*$ , причём при этом изоморфизме разложимые элементы  $\Lambda^1 V^*$  соответствуют разложимым элементам  $\Lambda^{n-1}V$ . Но в  $\Lambda^1 V^*$  все элементы разложимые.



Второе решение. Рассмотрим линейное отображение  $f: V \rightarrow \Lambda^n(V)$ , заданное формулой  $f(v) = v \wedge \omega$ . Так как  $\dim \Lambda^n(V) = 1$ , то  $\dim \text{Ker } f \geq n - 1$ . Поэтому существуют линейно независимые векторы  $e_1, \dots, e_{n-1}$ , лежащие в  $\text{Ker } f$ , т.е.  $e_i \wedge \omega = 0$  при  $i = 1, \dots, n - 1$ . Согласно задаче 4.1  $\omega = \lambda e_1 \wedge \dots \wedge e_{n-1}$ .

**4.4.** Для кососимметрического тензора  $\omega \in \Lambda^2(\mathbb{R}^4)$  условие  $\omega \wedge \omega = 0$  эквивалентно равенству  $a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23} = 0$ . В свою очередь, это равенство является единственным соотношением Плюккера в рассматриваемой ситуации.

**4.5.** Пусть  $W_1 = \langle e_1, \dots, e_{2n} \rangle$ . Докажем, что  $W = W_1$ . Пространство  $W$  является наименьшим пространством, для которого  $\Lambda \subset \Lambda^2 W$  (теорема 5.4.2.1). Ясно, что  $\Lambda \subset \Lambda^2 W_1$ , поэтому  $W \subset W_1$  и  $\dim W \leq \dim W_1 = 2n$ . С другой стороны,  $\Lambda^n \omega = n! e_1 \wedge \dots \wedge e_{2n}$  (см. задачу 2.3) и  $\Lambda^2 \omega \in \Lambda^{2n} W$ . Поэтому  $\Lambda^{2n} W \neq 0$  т.е.  $\dim W \geq 2n$ .

**4.6.** При замене базисов пространств  $X$  и  $Y$  тензоры  $z_1$  и  $z_2$  заменяются на пропорциональные им тензоры, поэтому можно считать, что  $z_1 + z_2 = (v_1 \wedge \dots \wedge v_k) \wedge (x_1 \wedge \dots \wedge x_{r-k} + y_1 \wedge \dots \wedge y_{r-k})$ , где  $v_1, \dots, v_k$  — базис пространства  $X \cap Y$ , а векторы  $x_1, \dots, x_{r-k}$  и  $y_1, \dots, y_{r-k}$  дополняют его до базисов пространств  $X$  и  $Y$ . Предположим, что  $z_1 + z_2 = u_1 \wedge \dots \wedge u_r$ . Пусть  $u = v + x + y$ , где  $v \in \langle v_1, \dots, v_k \rangle$ ,  $x \in \langle x_1, \dots, x_{r-k} \rangle$  и  $y \in \langle y_1, \dots, y_{r-k} \rangle$ . Тогда  $(z_1 + z_2) \wedge u = (v_1 \wedge \dots \wedge v_k) \wedge (x_1 \wedge \dots \wedge x_{r-k} \wedge y + y_1 \wedge \dots \wedge y_{r-k} \wedge x)$ . Если  $r - k > 1$ , то ненулевые тензоры  $x_1 \wedge \dots \wedge x_{r-k} \wedge y$  и  $y_1 \wedge \dots \wedge y_{r-k} \wedge x$  будут линейно независимы. Значит, в этом случае из равенства  $(z_1 + z_2) \wedge u = 0$  следует, что  $u \in \langle v_1, \dots, v_k \rangle$ , т.е.  $\langle u_1, \dots, u_r \rangle \subset \langle v_1, \dots, v_k \rangle$  и  $r \leq k$ . Получено противоречие, поэтому  $r - k = 1$ .

**4.7.** Любые два подпространства  $[w_1]$  и  $[w_2]$  имеют общее  $(k - 1)$ -мерное подпространство (см. задачу 4.6). Остаётся воспользоваться теоремой 2.5.9.1.

**4.8.** Если  $\Lambda \subset \Lambda^k W_1$  и  $v^* \in W_1^\perp$ , то  $i(v^*)\Lambda = 0$ , поэтому  $W_1^\perp \subset \Lambda^\perp$ , а значит,  $W_1 \supset (\Lambda^\perp)^\perp$ . Остаётся доказать, что  $\Lambda \subset \Lambda^k W$ . Пусть  $V = W \oplus U$ ,  $u_1, \dots, u_a$  — базис пространства  $U$ ,  $w_1, \dots, w_{n-a}$  — базис пространства  $W$ ,  $u_1^*, \dots, u_{n-a}^*$  — двойственный базис пространства  $V^*$ . Тогда  $u_j^* \in W^\perp = \Lambda^\perp$ , поэтому  $i(u_j^*)\Lambda = 0$ . Если  $j \in \{j_1, \dots, j_b\}$  и  $\{j'_1, \dots, j'_{b-1}\} \cup \{j\} = \{j_1, \dots, j_b\}$ , то отображение  $i(u_j^*)$  переводит тензор  $w_{i_1} \wedge \dots \wedge w_{i_{k-b}} \wedge u_{j_1} \wedge \dots \wedge u_{j_b}$  в  $\lambda w_{i_1} \wedge \dots \wedge w_{i_{k-b}} \wedge u_{j'_1} \wedge \dots \wedge u_{j'_{b-1}}$ ; в противном случае отображение  $i(u_j^*)$  переводит этот тензор в нуль. Поэтому  $i(u_j^*)\Lambda^{k-b} W \otimes \Lambda^b U \subset \Lambda^{k-b} W \otimes \Lambda^{b-1} U$ . Пусть  $\sum_\alpha \Lambda_\alpha \otimes u_\alpha$  — компонента некоторого элемента пространства  $\Lambda$ , лежащая в пространстве  $\Lambda^{k-1} W \otimes \Lambda^1 U$ . Тогда  $i(u_\beta^*)(\sum_\alpha \Lambda_\alpha \otimes u_\alpha) = 0$ , а значит,  $0 = \langle i(u_\beta^*) \sum_\alpha \Lambda_\alpha \otimes u_\alpha, f \rangle = \langle \sum_\alpha \Lambda_\alpha \otimes u_\alpha, u_\beta^* \wedge f \rangle = \langle \Lambda_\beta \otimes u_\beta, u_\beta^* \wedge f \rangle$  для всех  $f$ . Но если  $\Lambda_\beta \otimes u_\beta \neq 0$ , то  $f$  можно подобрать так, что  $\langle \Lambda_\beta \otimes u_\beta, u_\beta^* \wedge f \rangle \neq 0$ . Аналогично доказывается, что компоненты любого элемента пространства  $\Lambda$  в слагаемых  $\Lambda^{k-i} W \otimes \Lambda^i U$  нулевые при  $i > 0$ .

**5.1.** Дополним базис  $e_1, \dots, e_k$  пространства  $U$  до базиса  $e_1, \dots, e_n$  пространства  $V$ . Пусть  $T = \sum \alpha_i v_1^i \otimes \dots \otimes v_p^i$ . Каждый элемент  $v_j^i \in V$  можно представить в виде  $v_j^i = u_j^i + w_j^i$ , где  $u_j^i \in U$  и  $w_j^i \in \langle e_{k+1}, \dots, e_n \rangle$ . Поэтому  $T = \sum \alpha_i u_1^i \otimes \dots \otimes u_p^i + \dots$ ; разложив элементы, обозначенные многоточием, по

базису  $e_1, \dots, e_n$ , легко проверить, что любая их линейная комбинация, не равная нулю, не может лежать в  $T_0^p(U)$ . А так как  $T \in T_0^p(U)$ , то  $T = \sum \alpha_i u_1^i \otimes \dots \otimes u_p^i$ , т.е. ранг элемента  $T$  в  $T_0^p(U)$  не больше, чем в  $T_0^p(V)$ . Обратное неравенство очевидно.

**5.2.** Пусть  $e_1^{\otimes p} + \dots + e_k^{\otimes p} = u_1^1 \otimes \dots \otimes u_p^1 + \dots + u_1^r \otimes \dots \otimes u_p^r$ . Согласно задаче 5.1 можно считать, что  $u_j^i \in \langle e_1, \dots, e_k \rangle$ . Тогда  $u_1^i = \sum_j \alpha_{ij} e_j$ , т.е.  $\sum_i u_1^i \otimes \dots \otimes u_p^i =$

$\sum_j e_j \otimes \left( \sum_i \alpha_{ij} u_2^i \right) \otimes \dots \otimes u_p^i$ . Следовательно,  $\sum_i \alpha_{ij} u_2^i \otimes \dots \otimes u_p^i = e_j^{\otimes p-1}$ , поэтому  $k$  линейно независимых тензоров  $e_1^{\otimes p-1}, \dots, e_k^{\otimes p-1}$  лежат в пространстве  $\langle u_2^1 \otimes \dots \otimes u_p^1, \dots, u_2^r \otimes \dots \otimes u_p^r \rangle$ , размерность которого не превосходит  $r$ . Следовательно,  $r \geq k$ .

**6.1.** Предположим, что  $AXB \equiv X^T$ . Тогда матрицы  $A$  и  $B$  имеют размер  $n \times m$  и  $\sum_{k,s} a_{ik} x_{ks} b_{sj} \equiv x_{ji}$ . Следовательно,  $a_{ij} b_{ij} = 1$  и  $a_{ik} b_{sj} = 0$ , если  $k \neq j$  или  $s \neq i$ . Из первого набора равенств следует, что все элементы матриц  $A$  и  $B$  ненулевые, но тогда второй набор равенств выполняться не может.

**6.2.** Пусть  $B, X \in M_{n,n}$ . Уравнение  $BX = \lambda X$  имеет ненулевое решение  $X$  тогда и только тогда, когда  $\lambda$  — собственное значение матрицы  $B$ . Если  $\lambda$  — собственное значение матрицы  $B$ , то  $BX = \lambda X$  для некоторой ненулевой матрицы  $X$ . Поэтому  $f(B)f(X) = \lambda f(X)$ , а значит,  $\lambda$  — собственное значение матрицы  $f(B)$ . Пусть  $B = \text{diag}(\beta_1, \dots, \beta_n)$ , где  $\beta_i$  — попарно различные ненулевые числа. Тогда матрица  $f(B)$  подобна  $B$ , т.е.  $f(B) = A_1 B A_1^{-1}$ .

Пусть  $g(X) = A_1^{-1} f(X) A_1$ . Тогда  $g(B) = B$ . Если  $X = \|x_{ij}\|_1^n$ , то  $BX = \|\beta_i x_{ij}\|_1^n$  и  $XB = \|x_{ij} \beta_j\|_1^n$ . Поэтому  $BX = \beta_i X$ , только если все строки матрицы  $X$ , кроме  $i$ -й, нулевые, а  $XB = \beta_j X$ , только если все столбцы матрицы  $X$ , кроме  $j$ -го, нулевые. Пусть  $I_{ij}$  — матрица, единственный ненулевой элемент которой равен 1 и стоит в  $i$ -й строке и  $j$ -м столбце. Тогда  $Bg(I_{ij}) = \beta_i g(I_{ij})$  и  $g(I_{ij})B = \beta_j g(I_{ij})$ , а значит,  $g(I_{ij}) = \alpha_{ij} I_{ij}$ . Легко проверить, что  $I_{ij} = I_{i1} I_{1j}$ . Поэтому  $\alpha_{ij} = \alpha_{i1} \alpha_{1j}$ . Кроме того,  $I_{ii}^2 = I_{ii}$ , поэтому  $\alpha_{ii}^2 = \alpha_{ii}$ , а значит,  $\alpha_{i1} \alpha_{1i} = 1$ , т.е.  $\alpha_{1i} = 1/\alpha_{i1}$ . Следовательно,  $\alpha_{ij} = \alpha_{i1}/\alpha_{1j}$ . Поэтому  $g(X) = A_2 X A_2^{-1}$ , где  $A_2 = \text{diag}(\alpha_{11}, \dots, \alpha_{n1})$ , а значит,  $f(X) = AXA^{-1}$ , где  $A = A_1 A_2$ .

## Глава 6.

# Матричные неравенства

### 6.1. Неравенства для симметрических и эрмитовых матриц

#### 6.1.1. Простейшие неравенства

Пусть  $A$  и  $B$  — симметрические (эрмитовы) матрицы. Будем считать, что  $A > B$  (соответственно  $A \geq B$ ), если матрица  $A - B$  положительно (соответственно неотрицательно) определённая. Неравенство  $A > 0$  означает, что матрица  $A$  положительно определённая.

**Теорема 6.1.1.1.** Если  $A > B > 0$ , то  $A^{-1} < B^{-1}$ .

**Доказательство.** Существует такая матрица  $P$ , что  $A = P^*P$  и  $B = P^*DP$ , где  $D$  — диагональная матрица с положительными элементами на диагонали (см. §4.2.1). Условие  $A > B$  означает, что  $I > D$ . Тогда  $I < D^{-1}$ , а так как  $A^{-1} = QQ^*$  и  $B^{-1} = QD^{-1}Q^*$ , где  $Q = P^{-1}$ , то  $A^{-1} < B^{-1}$   $\square$

**Теорема 6.1.1.2.** Если  $A > 0$ , то  $A + A^{-1} \geq 2I$ .

**Доказательство.** Для матрицы  $A$  существует такой ортонормированный базис, что  $(Ax, x) = \sum \alpha_i |x_i|^2$ . В этом базисе  $((A + A^{-1})x, x) = \sum (\alpha_i + \alpha_i^{-1}) |x_i|^2$  и  $(2Ix, x) = \sum 2 |x_i|^2$ . Остаётся заметить, что  $\alpha_i + \alpha_i^{-1} \geq 2$ , так как  $\alpha_i > 0$ .  $\square$

**Теорема 6.1.1.3.** Если  $A$  — вещественная матрица и  $A > 0$ , то  $(A^{-1}x, x) = \max_y (2(x, y) - (Ay, y))$ .

**Доказательство.** Для матрицы  $A$  существует такой ортонормированный базис, что  $(Ax, x) = \sum \alpha_i x_i^2$ . Так как  $2x_i y_i - \alpha_i y_i^2 = -\alpha_i \left(y_i - \frac{x_i}{\alpha_i}\right)^2 + \frac{x_i^2}{\alpha_i}$ , то  $\max_y (2(x, y) - (Ay, y)) = \sum x_i^2 / \alpha_i = (A^{-1}x, x)$ , причём максимум достигается при  $y_i = x_i / \alpha_i$ .  $\square$

### 6.1.2. Неравенство Адамара

**Теорема 6.1.2.1.** Пусть  $A = \begin{pmatrix} A_1 & B \\ B^* & A_2 \end{pmatrix}$  — положительно определённая матрица. Тогда  $\det A \leq \det A_1 \det A_2$ .

**Доказательство.** Матрицы  $A_1$  и  $A_2$  положительно определены. Легко проверить, что  $\det A = \det A_1 \det(A_2 - B^* A_1^{-1} B)$  (см. §1.3.1). Матрица  $B^* A_1^{-1} B$  положительно определена, поэтому  $\det(A_2 - B^* A_1^{-1} B) \leq \det A_2$  (см. задачу 1.1). Следовательно,  $\det A \leq \det A_1 \det A_2$ ; равенство достигается, только если  $B^* A_1^{-1} B = 0$ , т.е.  $B = 0$ .  $\square$

**Следствие 1 (неравенство Адамара).** Если матрица  $A = \|a_{ij}\|_1^n$  положительно определена, то  $\det A \leq a_{11} a_{22} \dots a_{nn}$ ; равенство достигается, только если матрица  $A$  диагональна.

**Следствие 2.** Если  $X$  — произвольная матрица, то

$$|\det X|^2 \leq \sum_i |x_{1i}|^2 \dots \sum_i |x_{ni}|^2.$$

**Доказательство.** Достаточно применить следствие 1 к матрице  $A = XX^*$ .  $\square$

**Теорема 6.1.2.2.** Матрица  $A = \begin{pmatrix} A_1 & B \\ B^* & A_2 \end{pmatrix}$ , где  $A_1 > 0$  положительно определённая тогда и только тогда, когда  $A_2 - B^* A_1^{-1} B > 0$ .

**Доказательство.** [Ev2] Пусть  $T = \begin{pmatrix} I & -A_1^{-1} B \\ 0 & I \end{pmatrix}$ . Матрица  $T^* A T = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 - B^* A_1^{-1} B \end{pmatrix}$  положительно определена тогда и только тогда, когда  $A_2 - B^* A_1^{-1} B > 0$ . С другой стороны,  $T^* A T > 0$  тогда и только тогда, когда  $A > 0$ .  $\square$

**Теорема 6.1.2.3.** Пусть  $A = \begin{pmatrix} A_1 & B \\ B^* & A_2 \end{pmatrix}$  — положительно определённая матрица, причём матрица  $B$  квадратная. Тогда  $|\det B|^2 \leq \det A_1 \det A_2$ .

**Доказательство.** [Ev2] Согласно теореме 6.1.2.2  $A_2 - B^* A_1^{-1} B > 0$ . Следовательно,  $\det(B^* A_1^{-1} B) \leq \det(B^* A_1^{-1} B) + \det(A_2 - B^* A_1^{-1} B) \leq \det A_2$ , т.е.  $|\det B|^2 = \det(BB^*) \leq \det A_1 \det A_2$ .  $\square$

**Теорема 6.1.2.4 (неравенство Сасса).** Пусть  $A$  — положительно определённая недиагональная матрица порядка  $n$ ;  $P_k$  — произведение всех главных миноров порядка  $k$  матрицы  $A$ . Тогда

$$P_1 > P_2^{a_2} > \dots > P_{n-1}^{a_{n-1}} > P_n, \quad \text{где} \quad a_k = \binom{n-1}{k-1}^{-1}.$$

**Доказательство.** [Mi3] Требуемое неравенство можно переписать в виде  $P_k^{n-k} > P_{k+1}^k$  ( $1 \leq k \leq n-1$ ). При  $n=2$  доказательство очевидно. Ясно также, что для диагональной матрицы  $P_k^{n-k} = P_{k+1}^k$ . Предположим, что  $P_k^{n-k} > P_{k+1}^k$  ( $1 \leq k \leq n-1$ ) для некоторого  $n \geq 2$ . Рассмотрим матрицу  $A$  порядка  $n+1$ . Пусть  $A_r$  — матрица, полученная из матрицы  $A$  вычёркиванием  $r$ -й строки и  $r$ -го столбца;  $P_{k,r}$  — произведение всех главных миноров порядка  $k$  матрицы  $A_r$ . Согласно предположению индукции  $P_{k,r}^{n-k} \geq P_{k+1,r}^k$  для  $1 \leq k \leq n-1$  и  $1 \leq r \leq n+1$ , причём хотя бы одна из матриц  $A_r$  не диагональная, поэтому хотя бы одно из неравенств строгое. Следовательно,  $\prod_{r=1}^{n+1} P_{k,r}^{n-k} > \prod_{r=1}^{n+1} P_{k+1,r}^k$  для  $1 \leq k \leq n-1$ , т.е.  $P^{(n-k)(n+1-k)} > P_{k+1}^{(n-k)k}$ . Сократив на  $n-k$ , при  $n \neq k$  получим требуемое. При  $n=k$  рассмотрим матрицу  $\text{adj } A = \|b_{ij}\|_1^{n+1}$ . Так как  $A > 0$ , то  $B > 0$  (см. задачу 1.4), поэтому  $b_{11} \dots b_{n+1,n+1} > \det B = (\det A)^n$  (см. следствие 1 теоремы 6.1.2.1), т.е.  $P_n > P_{n+1}^n$ .  $\square$

**Замечание.** Неравенство  $P_1 > P_n$  совпадает с неравенством Адамара.

### 6.1.3. Определитель линейной комбинации

**Теорема 6.1.3.1.** Пусть  $\alpha_i > 0$ ,  $\sum \alpha_i = 1$  и  $A_i > 0$ . Тогда

$$|\alpha_1 A_1 + \dots + \alpha_k A_k| \geq |A_1|^{\alpha_1} \dots |A_k|^{\alpha_k}.$$

**Доказательство.** [Mi2] Рассмотрим сначала случай  $k=2$ . Пусть  $A, B > 0$ . Тогда существует такая матрица  $P$ , что  $P * AP = \text{diag}(\lambda_1, \dots, \lambda_n) = \Lambda$  и  $P^* BP = I$ . Поэтому  $P^*(\alpha A + (1-\alpha)B)P = \alpha \Lambda + (1-\alpha)I$ , т.е.  $|\alpha A + (1-\alpha)B| \cdot |P^* P| = |\alpha \Lambda + (1-\alpha)I|$ ; ясно также, что  $|P^* P| = |B|^{-1}$ . Следовательно,  $|\alpha A + (1-\alpha)B| = |B| \prod_{i=1}^n (\alpha \lambda_i + 1 - \alpha)$ . Докажем теперь, что если  $0 < \alpha < 1$ ,

и  $\lambda > 0$ , то  $\alpha\lambda + 1 - \alpha \geq \lambda^\alpha$ . Функция  $f(t) = \lambda^t$  выпукла вниз, так как  $f''(t) = (\ln \lambda)^2 \lambda^t \geq 0$ . Поэтому  $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$ . Остаётся положить  $x = 1$  и  $y = 0$ . В итоге получаем  $|\alpha A + (1 - \alpha)B| \geq |B| \cdot |\Lambda|^\alpha = |B| \cdot |A|^\alpha |B|^{-\alpha} = |A|^\alpha |B|^{1-\alpha}$ .

Дальнейшее доказательство проведём индукцией по  $k$ ; будем считать, что  $k \geq 3$ . Так как  $\alpha_1 A_1 + \dots + \alpha_k A_k = (1 - \alpha_k)B + \alpha_k A_k$ , причём матрица  $B = \alpha_1 A_1 / (1 - \alpha_k) + \dots + \alpha_{k-1} A_{k-1} / (1 - \alpha_k)$  положительно определена, то

$$\begin{aligned} |\alpha_1 A_1 + \dots + \alpha_k A_k| &\geq \\ &\geq \left| \frac{\alpha_1}{1 - \alpha_k} A_1 + \dots + \frac{\alpha_{k-1}}{1 - \alpha_k} A_{k-1} \right|^{1-\alpha_k} |A_k|^{\alpha_k}. \end{aligned}$$

А так как  $\frac{\alpha_1}{1-\alpha_k} + \dots + \frac{\alpha_{k-1}}{1-\alpha_k} = 1$ , то

$$\begin{aligned} \left| \frac{\alpha_1}{1 - \alpha_k} A_1 + \dots + \frac{\alpha_{k-1}}{1 - \alpha_k} A_{k-1} \right| &\geq \\ &\geq |A_1|^{\alpha_1/(1-\alpha_k)} \dots |A_{k-1}|^{\alpha_{k-1}/(1-\alpha_k)}. \end{aligned}$$

□

**Замечание.** Можно проверить, что равенство достигается тогда и только тогда, когда  $A_1 = \dots = A_k$ .

**Теорема 6.1.3.2.** Пусть  $\lambda_i$  — произвольные комплексные числа и  $A_i \geq 0$ . Тогда

$$|\det(\lambda_1 A_1 + \dots + \lambda_k A_k)| \leq \det(|\lambda_1| A_1 + \dots + |\lambda_k| A_k).$$

**Доказательство.** [Fr1] Пусть  $k = 2$ ; можно считать, что  $\lambda_1 = 1$  и  $\lambda_2 = \lambda$ . Существует такая унитарная матрица  $U$ , что матрица  $UA_1 U^{-1} = D$  диагональна. Тогда  $M = UA_2 U^{-1} \geq 0$  и

$$\begin{aligned} \det(A_1 + \lambda A_2) &= \det(D + \lambda M) = \\ &= \sum_{p=0}^n \lambda^p \sum_{i_1 < \dots < i_p} M \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} d_{j_1} \dots d_{j_{n-p}}, \end{aligned}$$

где набор  $(j_1, \dots, j_{n-p})$  дополняет  $(i_1, \dots, i_p)$  до  $(1, \dots, n)$  (см. решение задачи 2.1). Из неотрицательной определённости матриц  $M$  и  $D$  следует, что  $M \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} \geq 0$  и  $d_j \geq 0$ . Поэтому

$$\begin{aligned} |\det(A_1 + \lambda A_2)| &\leq \sum_{p=0}^n |\lambda|^p M \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} d_{j_1} \dots d_{j_{n-p}} = \\ &= \det(D + |\lambda| M) = \det(A_1 + |\lambda| A_2). \end{aligned}$$

Проведём теперь доказательство индуктивного шага. Снова будем считать, что  $\lambda_1 = 1$ . Пусть  $A = A_1$  и  $A' = \lambda_2 A_2 + \dots + \lambda_{k+1} A_{k+1}$ . Существует такая унитарная матрица  $U$ , что матрица  $UAU^{-1} = D$  диагональна; матрицы  $M_j = UA_jU^{-1}$  и  $M = UA'U^{-1}$  неотрицательно определены. Поэтому

$$\begin{aligned} |\det(A + A')| &= |\det(D + M)| \leq \\ &\leq \sum_{p=0}^n \sum_{i_1 < \dots < i_p} M \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} d_{j_1} \dots d_{j_{n-p}}. \end{aligned}$$

Так как  $M = \lambda_2 M_2 + \dots + \lambda_{k+1} M_{k+1}$ , то согласно предположению индукции

$$M \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} \leq \det \left( \sum_{j=2}^{k+1} |\lambda_j| M_j \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} \right).$$

Остаётся заметить, что

$$\begin{aligned} \sum_{p=0}^n \sum_{i_1 < \dots < i_p} d_{j_1} \dots d_{j_{n-p}} \det \left( \sum_{j=2}^{k+1} |\lambda_j| M_j \begin{pmatrix} i_1 & \dots & i_p \\ i_1 & \dots & i_p \end{pmatrix} \right) = \\ = \det(D + |\lambda_2| M_2 + \dots + |\lambda_{k+1}| M_{k+1}) = \det(\sum |\lambda_i| A_i). \end{aligned}$$

□

#### 6.1.4. Отношение миноров

**Теорема 6.1.4.1.** Пусть вещественные матрицы  $A$  и  $B$  положительно определены, а матрицы  $A_1$  и  $B_1$  получаются из  $A$  и  $B$  вычёркиванием первой строки и первого столбца. Тогда

$$\frac{|A + B|}{|A_1 + B_1|} \geq \frac{|A|}{|A_1|} + \frac{|B|}{|B_1|}.$$

**Доказательство.** [Be2] Если  $A > 0$ , то

$$(x, Ax)(y, A^{-1}y) \geq (x, y)^2. \quad (6.1)$$

В самом деле, существует такая унитарная матрица  $U$ , что  $U^*AU = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ , причём  $\lambda_i > 0$ . После замены  $x = Ua$  и  $y = Ub$  перейдём к неравенству Коши-Шварца

$$(\sum \lambda_i a_i^2) (\sum b_i^2 / \lambda_i) \geq (\sum a_i b_i)^2. \quad (6.2)$$

Неравенство (6.2) обращается в равенство при  $a_i = b_i/\lambda_i$ , поэтому

$$f(A) = \frac{1}{(y, A^{-1}y)} = \min_x \frac{(x, Ax)}{(x, y)^2}.$$

Докажем теперь, что если  $y = (1, 0, 0, \dots, 0) = e_1$ , то  $f(A) = |A|/|A_1|$ . В самом деле,  $(e_1, A^{-1}e_1) = e_1 A^{-1} e_1^T = (e_1 \operatorname{adj} A e_1^T)/|A| = (\operatorname{adj} A)_{11}/|A| = |A_1|/|A|$ .

Остаётся заметить, что  $\min_x g(x) + \min_x h(x) \leq \min_x (g(x) + h(x))$  для любых функций  $g$  и  $h$ , и положить  $g(x) = (x, Ax)/(x, e_1)$  и  $h(x) = (x, Bx)/(x, e_1)$ .  $\square$

### Задачи

**1.1.** Пусть  $A$  и  $B$  — матрицы порядка  $n$  ( $n > 1$ ), причём  $A > 0$  и  $B \geq 0$ . Докажите, что  $|A + B| \geq |A| + |B|$ , причём равенство достигается только при  $B = 0$ .

**1.2.** Матрицы  $A$  и  $B$  эрмитовы, причём  $A > 0$ . Докажите, что  $|\det(A + iB)| \geq \det A$ , причём равенство достигается только при  $B = 0$ .

**1.3.** Пусть  $A_k$  и  $B_k$  — угловые подматрицы порядка  $k$  положительно определённых матриц  $A$  и  $B$ , причём  $A > B$ . Докажите, что  $|A_k| > |B_k|$ .

**1.4.** Пусть  $A \geq B \geq 0$ . Верно ли, что  $A^2 \geq B^2$ ?

**1.5.** Матрицы  $A$  и  $B$  вещественные симметрические, причём  $A \geq 0$ . Докажите, что если матрица  $C = A + iB$  вырожденная, то  $Cx = 0$  для некоторого ненулевого вещественного вектора  $x$ .

**1.6.** Вещественная симметрическая матрица  $A$  положительно определена. Докажите, что

$$\det \begin{pmatrix} 0 & x_1 & \dots & x_n \\ x_1 & & & \\ \vdots & & A & \\ x_n & & & \end{pmatrix} \leq 0.$$

**1.7.** Докажите, что если  $A > 0$ , то  $|A|^{1/n} = \min \operatorname{tr}(AB)/n$ , где  $n$  — порядок матрицы  $A$  и минимум берётся по всем положительно определённым матрицам  $B$  с определителем 1.

**1.8.** Пусть  $A$  — вещественная положительно определённая матрица с собственными значениями  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . Докажите, что для любого



вектора  $x$  единичной длины выполняется *неравенство Канторовича*

$$(Ax, x)(A^{-1}x, x) \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}.$$

## 6.2. Неравенства для собственных значений

### 6.2.1. Неравенство Шура

**Теорема 6.2.1.1 (неравенство Шура).** Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения матрицы  $A = \|a_{ij}\|_1^n$ . Тогда  $\sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i,j=1}^n |a_{ij}|^2$ , причём равенство достигается тогда и только тогда, когда  $A$  — нормальная матрица.

**Доказательство.** Существует такая унитарная матрица  $U$ , что матрица  $T = U^*AU$  верхняя треугольная, причём матрица  $T$  диагональна тогда и только тогда, когда матрица  $A$  нормальна (теорема 3.7.1.1). Так как  $^* = U^*A^*U$ , то  $TT^* = U^*AA^*U$ , а значит,  $\text{tr}(TT^*) = \text{tr}(AA^*)$ . Остаётся заметить, что  $\text{tr}(AA^*) = \sum_{i,j=1}^n |a_{ij}|^2$  и  $\text{tr}(TT^*) = \sum_{i=1}^n |\lambda_i|^2 + \sum_{i < j} |t_{ij}|^2$ .  $\square$

**Теорема 6.2.1.2.** Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения матрицы  $A = B + iC$ , где матрицы  $B$  и  $C$  эрмитовы. Тогда

$$\sum_{i=1}^n |\text{Re } \lambda_i|^2 \leq \sum_{i,j=1}^n |b_{ij}|^2 \text{ и } \sum_{i=1}^n |\text{Im } \lambda_i|^2 \leq \sum_{i,j=1}^n |c_{ij}|^2.$$

**Доказательство.** Пусть, как и в доказательстве теоремы 6.2.1.1,  $T = U^*AU$ . Так как  $B = (A + A^*)/2$  и  $iC = (A - A^*)/2$ , то  $U^*BU = (T + T^*)/2$  и  $U^*(iC)U = (T - T^*)/2$ . Поэтому

$$\sum_{i,j=1}^n |b_{ij}|^2 = \text{tr}(BB^*) = \frac{1}{4} \text{tr}(T + T^*)^2 = \sum_{i=1}^n |\text{Re } \lambda_i|^2 + \frac{1}{2} \sum_{i < j} |t_{ij}|^2$$

и

$$\sum_{i,j=1}^n |c_{ij}|^2 = \sum_{i=1}^n |\text{Im } \lambda_i|^2 + \frac{1}{2} \sum_{i < j} |t_{ij}|^2.$$

$\square$

### 6.2.2. Сумма двух эрмитовых матриц

**Теорема 6.2.2.1.** Пусть  $A$  и  $B$  — эрмитовы матрицы,  $C = A + B$ . Упорядочим собственные значения этих матриц в порядке возрастания:  $\alpha_1 \leq \dots \leq \alpha_n$ ,  $\beta_1 \leq \dots \leq \beta_n$ ,  $\gamma_1 \leq \dots \leq \gamma_n$ . Тогда:

- а)  $\gamma_i \geq \alpha_j + \beta_{i-j+1}$  при  $i \geq j$ ;
- б)  $\gamma_i \leq \alpha_j + \beta_{i-j+n}$  при  $i \leq j$ .

**Доказательство.** Выберем для матриц  $A$ ,  $B$  и  $C$  такие ортонормированные базисы  $\{a_i\}$ ,  $\{b_i\}$  и  $\{c_i\}$ , что  $Aa_i = \alpha_i a_i$  и т.д. Предположим сначала, что  $i \geq j$ . Рассмотрим подпространства  $V_1 = \langle a_j, \dots, a_n \rangle$ ,  $V_2 = \langle b_{i-j+1}, \dots, b_n \rangle$  и  $V_3 = \langle c_1, \dots, c_i \rangle$ . Так как  $\dim V_1 = n - j + 1$ ,  $\dim V_2 = n - i + j$  и  $\dim V_3 = i$ , то  $\dim(V_1 \cap V_2 \cap V_3) \geq \dim V_1 + \dim V_2 + \dim V_3 - 2n = 1$ . Поэтому подпространство  $V_1 \cap V_2 \cap V_3$  содержит некоторый вектор  $x$  единичной длины. Ясно, что  $\alpha_j + \beta_{i-j+1} \leq (x, Ax) + (x, Bx) = (x, Cx) \leq \gamma_i$ .

Заменив матрицы  $A$ ,  $B$  и  $C$  на  $-A$ ,  $-B$  и  $-C$ , неравенство б) можно свести к неравенству а).  $\square$

**Теорема 6.2.2.2.** Пусть  $A = \begin{pmatrix} B & C \\ C^* & D \end{pmatrix}$  — эрмитова матрица. Упорядочим собственные значения матриц  $A$  и  $B$  в порядке возрастания:  $\alpha_1 \leq \dots \leq \alpha_n$ ,  $\beta_1 \leq \dots \leq \beta_m$ . Тогда  $\alpha_i \leq \beta_i \leq \alpha_{i+n-m}$ .

**Доказательство.** Выберем для операторов  $A$  и  $B$  ортонормированные собственные базисы  $\{a_i\}$  и  $\{b_i\}$ ; можно считать, что операторы  $A$  и  $B$  действуют в пространствах  $V$  и  $U$ , где  $U \subset V$ . Рассмотрим подпространства  $V_1 = \langle a_i, \dots, a_n \rangle$  и  $V_2 = \langle b_1, \dots, b_i \rangle$ . Подпространство  $V_1 \cap V_2$  содержит некоторый единичный вектор  $x$ . Ясно, что  $\alpha_i \leq (x, Ax) = (x, Bx) \leq \beta_i$ . Применяя это неравенство к матрице  $-A$ , получим  $-\alpha_{n-i+1} \leq -\beta_{m-i+1}$ , т.е.  $\beta_j \leq \alpha_{j+n-m}$ .  $\square$

### 6.2.3. Произведение двух эрмитовых проекторов

**Теорема 6.2.3.1.** Пусть  $A$  и  $B$  — эрмитовы проекторы, т.е.  $A^2 = A$  и  $B^2 = B$ . Тогда собственные значения матрицы  $AB$  вещественны и заключены между 0 и 1.

**Доказательство.** [Af] Собственные значения матрицы  $AB = (AAB)B$  совпадают с собственными значениями матрицы  $B(AAB) = (AB)^*AB$  (теорема 3.1.6.1). Последняя матрица неотрицательно определённая, поэтому её собственные значения вещественны и неотрицательны. Если все собственные значения матрицы  $AB$  нулевые, то эрмитова матрица

$(AB)^*AB$  нулевая, а значит,  $AB = 0$  (см. задачу 4.3). Предположим теперь, что  $ABx = \lambda x \neq 0$ . Тогда  $Ax = \lambda^{-1}AABx = \lambda^{-1}ABx = x$ , а значит,  $(x, Bx) = (Ax, Bx) = (x, ABx) = \lambda(x, x)$ , т.е.  $\lambda = (x, Bx)/(x, x)$ . Для матрицы  $B$  существует такой ортонормированный базис, что  $(x, Bx) = \beta_1|x_1|^2 + \dots + \beta_n|x_n|^2$ , где  $\beta_i = 0$  или 1. Поэтому  $\lambda \leq 1$ .  $\square$

#### 6.2.4. Сингулярные значения

Пусть  $A$  — матрица размером  $m \times n$ . Тогда  $A^*A$  — эрмитова неотрицательно определённая матрица порядка  $n$  с неотрицательными собственными значениями  $\lambda_1, \dots, \lambda_n$ . Числа  $\sigma_i = \sqrt{\lambda_i}$  называют *сингулярными значениями* матрицы  $A$ .

Если  $A$  — вектор-столбец (матрица размером  $m \times 1$ ), то сингулярное значение  $A$  — длина этого вектора.

Для эрмитовой неотрицательно определённой матрицы сингулярные значения совпадают с собственными значениями.

Если  $A = SU$  — полярное разложение квадратной матрицы  $A$ , то сингулярные значения матрицы  $A$  совпадают с собственными значениями матрицы  $S$ . Для матрицы  $S$  существует такая унитарная матрица  $V$ , что  $S = V\Lambda V^*$ , где матрица  $\Lambda$  диагональная. Поэтому любую квадратную матрицу  $A$  можно представить в виде  $A = V\Lambda W$ , где матрицы  $V$  и  $W$  унитарные, а  $\Lambda = \text{diag}(\sigma_1, \dots, \sigma_n)$ .

**Теорема 6.2.4.1.** Пусть  $A$  — квадратная матрица, у которой наибольшее и наименьшее сингулярные значения равны  $\sigma_{\max}$  и  $\sigma_{\min}$ . Тогда

$$\sigma_{\min}\|x\| \leq \|Ax\| \leq \sigma_{\max}\|x\|$$

для любого вектора  $x$ .

**Доказательство.** Ясно, что  $\|Ax\| = \|V\Lambda Wx\| = \|\Lambda y\|$ , причём  $\|x\| = \|y\|$ . Если вектор  $y$  имеет координаты  $y_1, \dots, y_n$ , то вектор  $\Lambda y$  имеет координаты  $\sigma_1 y_1, \dots, \sigma_n y_n$ . Поэтому  $\sigma_{\min}\|y\| \leq \|\Lambda y\| \leq \sigma_{\max}\|y\|$ .  $\square$

**Теорема 6.2.4.2.** Пусть  $\sigma_1, \dots, \sigma_n$  — сингулярные значения матрицы  $A$ , причём  $\sigma_1 \geq \dots \geq \sigma_n$ , а  $\lambda_1, \dots, \lambda_n$  — собственные значения матрицы  $A$ , причём  $|\lambda_1| \geq \dots \geq |\lambda_n|$ . Тогда  $|\lambda_1| \dots |\lambda_m| \leq \sigma_1 \dots \sigma_m$  при  $m \leq n$ .

**Доказательство.** Пусть  $Ax = \lambda_1 x$ . Тогда  $|\lambda_1|^2(x, x) = (Ax, Ax) = (x, A^*Ax) \leq \sigma_1^2(x, x)$ , так как  $\sigma_1^2$  — наибольшее собственное значение эрмитова оператора  $A^*A$ . Поэтому  $|\lambda_1| \leq \sigma_1$ ; при  $m = 1$  неравенство доказано. Применим полученное неравенство к операторам  $\Lambda^m(A)$

и  $\Lambda^m(A^*A)$  (см. §5.2.5). Их собственные значения равны  $\lambda_{i_1} \dots \lambda_{i_m}$  и  $\sigma_{i_1}^2 \dots \sigma_{i_m}^2$ , поэтому  $|\lambda_1 \dots \lambda_m| \leq \sigma_1 \dots \sigma_m$ .

Ясно также, что  $|\lambda_1 \dots \lambda_n| = |\det A| = \sqrt{\det(A^*A)} = \sigma_1 \dots \sigma_n$ .  $\square$

**Теорема 6.2.4.3.** Пусть  $\sigma_1 \geq \dots \geq \sigma_n$  — сингулярные значения матрицы  $A$ ,  $\tau_1 \geq \dots \geq \tau_n$  — сингулярные значения матрицы  $B$ . Тогда  $|\operatorname{tr}(AB)| \leq \sum_{i=1}^n \sigma_i \tau_i$ .

**Доказательство.** [Mi4] Пусть  $A = U_1 S V_1$  и  $B = U_2 T V_2$ , где  $U_i$  и  $V_i$  — унитарные матрицы,  $S = \operatorname{diag}(\sigma_1, \dots, \sigma_n)$  и  $T = \operatorname{diag}(\tau_1, \dots, \tau_n)$ . Тогда  $\operatorname{tr}(AB) = \operatorname{tr}(U_1 S V_1 U_2 T V_2) = \operatorname{tr}(V_2 U_1 S V_1 V_2 T) = \operatorname{tr}(U^T S V T)$ , где  $U = (V_2 U_1)^T$  и  $V = V_1 U_2$ . Поэтому  $|\operatorname{tr}(AB)| = |\sum u_{ij} v_{ij} \sigma_i \tau_j| \leq (\sum |u_{ij}|^2 \sigma_i \tau_j + \sum |v_{ij}|^2 \sigma_i \tau_j)/2$ . Матрицы с элементами  $|u_{ij}|^2$  и  $|v_{ij}|^2$  дважды стохастические, поэтому  $\sum |u_{ij}|^2 \sigma_i \tau_j \leq \sum \sigma_i \tau_i$  и  $\sum |v_{ij}|^2 \sigma_i \tau_j \leq \sum \sigma_i \tau_i$  (см. задачу 6.1).  $\square$

### 6.2.5. Круги Гершгорина

Следующее легко доказываемое утверждение даёт полезную оценку для собственных значений матрицы.

**Теорема 6.2.5.1 (Гершгорин [Г1]).** Каждое собственное значение матрицы  $\|a_{ij}\|_1^n$  принадлежит одному из кругов  $|a_{kk} - z| \leq \rho_k$ , где  $\rho_k = \sum_{j \neq k} |a_{kj}|$ .

**Доказательство.** Пусть  $\lambda$  — собственное значение данной матрицы. Тогда система  $\sum a_{ij} x_j = \lambda x_i$  ( $i = 1, \dots, n$ ) имеет ненулевое решение  $(x_1, \dots, x_n)$ . Выберем среди чисел  $x_1, \dots, x_n$  наибольшее по модулю; пусть это будет  $x_k$ . Так как  $a_{kk} x_k - \lambda x_k = - \sum_{j \neq k} a_{kj} x_j$ , то  $|a_{kk} x_k - \lambda x_k| \leq \sum_{j \neq k} |a_{kj} x_j| \leq \rho_k |x_k|$ , т.е.  $|a_{kk} - \lambda| \leq \rho_k$ .  $\square$

**Следствие.** Если  $|a_{kk}| > \sum_{j \neq k} |a_{kj}|$  для всех  $k$ , то матрица  $A$  невырожденная.

**Доказательство.** Число  $z = 0$  не удовлетворяет ни одному из неравенств  $|a_{kk} - z| \leq \rho_k$ , поэтому 0 не является собственным значением данной матрицы.  $\square$

**Замечание.** Аналогично доказывается, что каждое собственное значение матрицы  $\|a_{ij}\|_1^n$  принадлежит одному из кругов  $|a_{kk} - z| \leq \rho'_k$ , где  $\rho'_k = \sum_{j \neq k} |a_{jk}|$ .

Набор кругов  $|a_{kk} - z| \leq \rho_k$ , где  $\rho_k = \sum_{j \neq k} |a_{kj}|$ , называют *кругами Гершгорина*. (Набор кругов  $|a_{kk} - z| \leq \rho'_k$  тоже называют кругами Гершгорина.)

Круги Гершгорина можно с успехом применять для того, чтобы получать различные оценки для корней многочленов. Следующие ниже примеры таких оценок взяты из статьи [Be1].

**Теорема 6.2.5.2.** Пусть  $A_1, \dots, A_n$  — произвольные положительные числа,  $A_0 = 0$  и  $A_1 = 1$ . Тогда корни многочлена

$$p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0 \quad (6.1)$$

по абсолютной величине не превосходят  $\max_{0 \leq i \leq n-1} \left\{ \frac{A_i + |a_i|}{A_{i+1}} \right\}$ .

**Доказательство.** Рассмотрим для многочлена  $p(z)$  его сопровождающую матрицу

$$C = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix}.$$

Собственные значения матрицы  $C$  — это корни многочлена  $p(z)$  (см. §1.1.9). Возьмём диагональную матрицу  $D = \text{diag}(A_1, \dots, A_n)$  и рассмотрим матрицу

$$DCD^{-1} = \begin{pmatrix} 0 & A_1/A_2 & 0 & \dots & 0 \\ 0 & 0 & A_2/A_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_{n-1}/A_n \\ -a_0/A_1 & -a_1/A_2 & -a_2/A_3 & \dots & -a_{n-1}/A_n \end{pmatrix},$$

имеющую те же самые собственные значения, что и матрица  $C$ . Согласно теореме 6.2.5.1 каждое собственное значение матрицы  $DCD^{-1}$  принадлежит одному из кругов  $|z| \leq \frac{|a_0|}{A_1}$ ,  $|z| \leq \frac{A_1 + |a_1|}{A_2}$ ,  $\dots$ ,  $|z| \leq \frac{A_{n-2} + |a_{n-2}|}{A_{n-1}}$ ,  $\left| z + \frac{a_{n-1}}{A_n} \right| \leq \frac{A_{n-1}}{A_n}$ . Остаётся заметить, что последний круг содержится в круге  $|z| \leq \frac{A_{n-1} + |a_{n-1}|}{A_n}$ .  $\square$

**Следствие 1 (Коши).** Корни многочлена (6.1) по абсолютной величине не превосходят наибольшего из чисел  $|a_0|$ ,  $1 + |a_1|$ ,  $\dots$ ,  $1 + |a_{n-1}|$ .

**Доказательство.** Достаточно положить  $A_1 = A_2 = \dots = A_{n-1} = 1$ .  $\square$

**Следствие 2.** Корни многочлена (6.1) по абсолютной величине не превосходят наибольшего из чисел  $\frac{|a_0|}{|a_1|}$ ,  $2\frac{|a_1|}{|a_2|}$ , ...,  $2\frac{|a_{n-2}|}{|a_{n-1}|}$ ,  $2|a_{n-1}|$ .

**Доказательство.** Положим  $A_i = |a_i|$  для  $i = 1, \dots, n-1$ . Тогда  $\frac{A_i + |a_i|}{A_{i+1}} = 2\frac{|a_i|}{|a_{i+1}|}$  при  $i \geq 1$  (мы полагаем  $a_n = 1$ ).  $\square$

**Теорема 6.2.5.3 (Коши).** Корни многочлена (6.1) по абсолютной величине не превосходят положительного корня многочлена

$$z^n - |a_{n-1}|z^{n-1} - |a_{n-2}|z^{n-2} - \dots - |a_0|. \quad (6.2)$$

**Доказательство.** Как и при доказательстве теоремы 6.2.5.2, рассмотрим матрицу  $DCD^{-1}$ . Пусть  $\rho$  — положительный корень многочлена (6.2). Положим  $A_i = \rho^{n-i}$ ,  $i = 1, 2, \dots, n-1$ , т.е.  $D = \text{diag}(\rho^{n-1}, \rho^{n-2}, \dots, \rho, 1)$ . В таком случае

$$DCD^{-1} = \begin{pmatrix} 0 & \rho & 0 & \dots & 0 \\ 0 & 0 & \rho & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \rho \\ -a_0/\rho^{n-1} & -a_1/\rho^{n-2} & -a_2/\rho^{n-3} & \dots & -a_{n-1} \end{pmatrix}. \quad (6.3)$$

На этот раз применим теорему Гершгорина не для столбцов матрицы, а для строк. В результате получим неравенства  $|z| \leq \rho$ , ...,  $|z| \leq \rho$ ,  $|z| \leq |a_{n-1}| + \frac{|a_{n-2}|}{\rho} + \dots + \frac{|a_0|}{\rho^{n-1}} = \rho$ .  $\square$

**Теорема 6.2.5.4.** Пусть  $\lambda_1, \dots, \lambda_n$  — положительные числа, причём  $\lambda_1 + \dots + \lambda_n = 1$ . Тогда корни многочлена (6.1) по абсолютной величине не превосходят  $\max_{1 \leq j \leq n} \left| \frac{a_{n-j}}{\lambda_j} \right|^{1/j}$ .

**Доказательство.** Рассмотрим такую же матрицу  $DCD^{-1}$ , как и при доказательстве теоремы 6.2.5.3, но только теперь положим  $\rho = \max_{1 \leq j \leq n} \left| \frac{a_{n-j}}{\lambda_j} \right|^{1/j}$ .

В результате снова получим неравенства  $|z| \leq \rho$  и  $|z| \leq |a_{n-1}| + \frac{|a_{n-2}|}{\rho} + \dots + \frac{|a_0|}{\rho^{n-1}}$ . Из неравенства  $|a_{n-j}| \leq \rho^j \lambda_j$  следует неравенство  $\frac{|a_{n-j}|}{\rho^{j-1}} \leq \rho \lambda_j$ . Поэтому

$$|a_{n-1}| + \frac{|a_{n-2}|}{\rho} + \dots + \frac{|a_0|}{\rho^{n-1}} \leq \rho(\lambda_1 + \dots + \lambda_n) = \rho.$$

$\square$

**Теорема 6.2.5.5.** *Корни многочлена (6.1) лежат в круге*

$$\left| z + \frac{a_{n-1}}{2} \right| \leq \frac{|a_{n-1}|}{2} + |a_{n-2}|^{1/2} + |a_{n-3}|^{1/3} + \dots + |a_0|^{1/n}. \quad (6.4)$$

**Доказательство.** Будем действовать так же, как и при доказательстве теоремы 6.2.5.4, только теперь положим  $\rho = \max_{2 \leq j \leq n} |a_{n-j}|^{1/j}$ . В результате получим неравенства  $|z| \leq \rho$  и

$$\begin{aligned} |z + a_{n-1}| &\leq \frac{|a_{n-2}|}{\rho} + \dots + \frac{|a_0|}{\rho^{n-1}} \leq \\ &\leq \frac{|a_{n-2}|}{|a_{n-2}|^{1/2}} + \dots + \frac{|a_0|}{|a_0|^{(n-1)/n}} \leq \\ &\leq |a_{n-2}|^{1/2} + |a_{n-3}|^{1/3} + \dots + |a_0|^{1/n}. \end{aligned}$$

Из каждого из этих неравенств следует неравенство (6.4), поскольку  $\left| z + \frac{a_{n-1}}{2} \right| \leq |z + a_{n-1}| + \frac{|a_{n-1}|}{2}$ ,  $\left| z + \frac{a_{n-1}}{2} \right| \leq |z| + \frac{|a_{n-1}|}{2}$  и  $\rho \leq |a_{n-2}|^{1/2} + |a_{n-3}|^{1/3} + \dots + |a_0|^{1/n}$ .  $\square$

Круги Гершгорина можно использовать аналогичным образом и для оценок снизу абсолютных величин корней многочленов.

**Теорема 6.2.5.6.** *Фиксируем положительное число  $t$ . Пусть  $M = \max_{1 \leq i \leq n} |a_i|t^i$ . Тогда корни многочлена (6.1) по абсолютной величине не меньше  $\frac{|a_0|t}{|a_0|+M}$ .*

**Доказательство.** Рассмотрим матрицу, обратную матрице (6.3) из доказательства теоремы 6.2.5.3 для  $\rho = t$ :

$$\begin{pmatrix} -\frac{a_1}{a_0} & -\frac{ta_2}{a_0} & \dots & -\frac{t^{n-2}a_{n-1}}{a_0} & -\frac{t^{n-1}}{a_0} \\ 1/t & 0 & \dots & 0 & 0 \\ 0 & 1/t & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/t & 0 \end{pmatrix}.$$

Каждое собственное значение  $\lambda$  этой матрицы удовлетворяет неравенству  $|\lambda| \leq \max \left\{ \frac{|a_0|+|a_i|t^i}{|a_0|t} \right\} \leq \frac{|a_0|+M}{|a_0|t}$ . Ясно также, что корнями многочлена (6.1) являются числа  $1/\lambda$ .  $\square$

### Задачи

**2.1.** Докажите, что если  $U$  — унитарная матрица, а  $S \geq 0$ , то  $|\operatorname{tr}(US)| \leq \operatorname{tr} S$ .

**2.2.** Докажите, что если  $A$  и  $B$  — неотрицательно определённые матрицы, то  $\operatorname{tr}(AB) \leq \operatorname{tr} A \operatorname{tr} B$ .

**2.3.** [Cu] Докажите, что  $\lim_{k \rightarrow \infty} A^k = 0$  тогда и только тогда, когда выполнено одно из следующих условий:

- а) все собственные значения матрицы  $A$  по модулю меньше 1;
- б) существует положительно определённая матрица  $H$ , для которой  $H - A^*HA > 0$ .

**2.4.** Пусть  $A = \|a_{ij}\|_1^n$  и  $B = \|b_{ij}\|_1^n$ , где  $a_{ij}, b_{ij} \in \mathbb{C}$ . Докажите, что

$$\operatorname{Re} \operatorname{tr}(AB) \leq \frac{1}{2}(\operatorname{tr} AA^* + \operatorname{tr} BB^*).$$

**2.5.** Матрицы  $A$  и  $B$  эрмитовы. Докажите, что  $\operatorname{tr}(AB)^2 \leq \operatorname{tr}(A^2B^2)$ .

**2.6.** [Hu] Пусть  $\alpha_1 \geq \dots \geq \alpha_n$  и  $\beta_1 \geq \dots \geq \beta_n$  — собственные значения положительно определённых матриц  $A$  и  $B$ , а  $\lambda$  — любое собственное значение матрицы  $AB$ . Докажите, что

$$\frac{2}{n} \cdot \frac{\alpha_n^2 \beta_n^2}{\alpha_n^2 + \beta_n^2} \leq \lambda \leq \frac{n}{2}(\alpha_1^2 + \beta_1^2).$$

**2.7.** [Y1] Матрицы  $A$  и  $B$  положительно определённые. Докажите, что  $\sqrt{\operatorname{tr}(AB)} \leq \frac{\operatorname{tr} A + \operatorname{tr} B}{2}$ .

### Сингулярные значения

**2.8.** Докажите, что для любой квадратной матрицы  $A$  собственные значения матриц  $A^*A$  и  $AA^*$  одинаковые.

**2.9.** Докажите, что если все сингулярные значения матрицы  $A$  равны, то  $A = \lambda U$ , где  $U$  — унитарная матрица.

**2.10.** Докажите, что если сингулярные значения матрицы  $A$  равны  $\sigma_1, \dots, \sigma_n$ , то сингулярные значения матрицы  $\operatorname{adj} A$  равны  $\prod_{i \neq 1} \sigma_i, \dots, \prod_{i \neq n} \sigma_i$ .

**2.11.** Пусть  $\sigma_1, \dots, \sigma_n$  — сингулярные значения матрицы  $A$ . Докажите, что собственными значениями матрицы  $\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$  являются числа  $\sigma_1, \dots, \sigma_n, -\sigma_1, \dots, -\sigma_n$ .



**2.12.** Пусть  $A$  — сопровождающая матрица многочлена  $p(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_0$ ,  $\sigma_1 \geq \dots \geq \sigma_n$  — её сингулярные значения. Докажите, что  $\sigma_2 = \dots = \sigma_{n-1} = 1$ , а  $\sigma_1$  и  $\sigma_n$  являются корнями квадратного уравнения

$$z^2 - (|a_0|^2 + \dots + |a_{n-1}|^2 + 1)z + |a_0|^2 = 0.$$

**2.13.** [Da] Пусть  $A$  — квадратная матрица порядка  $n$ , а  $B$  — матрица размером  $n \times m$ , где  $n > m$ . Пусть, далее,  $\alpha_1 \geq \dots \geq \alpha_n$  — сингулярные значения матрицы  $A$ ,  $\beta_1 \geq \dots \geq \beta_m$  — сингулярные значения матрицы  $B$ ,  $\gamma_1 \geq \dots \geq \gamma_m$  — сингулярные значения матрицы  $C = AB$ . Докажите, что

$$\alpha_{n-m+1} \dots \alpha_n \beta_1 \dots \beta_m \leq \gamma_1 \dots \gamma_m \leq \alpha_1 \dots \alpha_m \beta_1 \dots \beta_m.$$

**2.14.** [Da] Пусть  $H$  — положительно определённая эрмитова матрица порядка  $n$  с собственными значениями  $\lambda_1, \dots, \lambda_n$ ,  $M$  — матрица размером  $n \times m$ , где  $m < n$ . Докажите, что

$$\lambda_{n-m+1} \dots \lambda_n \det(M^* M) \leq \det(M^* H M) \leq \lambda_1 \dots \lambda_m \det(M^* M).$$

## 6.3. Неравенства для норм матриц

### 6.3.1. Операторная норма

*Операторной* (или *спектральной*) *нормой* матрицы  $A$  называют величину  $\|A\|_s = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ ; число  $\rho(A) = \max |\lambda_i|$ , где  $\lambda_1, \dots, \lambda_n$  — собственные значения матрицы  $A$ , называют *спектральным радиусом* матрицы  $A$ . Так как существует такой ненулевой вектор  $x$ , что  $Ax = \lambda_i x$ , то  $\|A\|_s \geq \rho(A)$ .

Легко проверить, что если  $U$  — унитарная матрица, то  $\|A\|_s = \|AU\|_s = \|UA\|_s$ . Для этого достаточно заметить, что  $\|AUx\|/\|x\| = \|Ay\|/\|U^{-1}y\| = \|Ay\|/\|y\|$ , где  $y = Ux$ , и  $\|UAx\|/\|x\| = \|Ax\|/\|x\|$ .

**Теорема 6.3.1.1.**  $\|A\|_s = \sqrt{\rho(A^* A)}$ .

**Доказательство.** Если  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ , то  $(|\Lambda x|/|x|)^2 = \sum |\lambda_i x_i|^2 / \sum |x_i|^2 \leq \max |\lambda_i|^2$ . Пусть  $|\lambda_j| = \max |\lambda_i|$  и  $\Lambda x = \lambda_j x$ . Тогда  $|\Lambda x|/|x| = |\lambda_j|$ . Поэтому  $\|\Lambda\|_s = \rho(\Lambda)$ .

Любую матрицу  $A$  можно представить в виде  $A = U\Lambda V$ , где  $U$  и  $V$  — унитарные матрицы, а матрица  $\Lambda$  диагональна, причём на её диагонали стоят сингулярные числа матрицы (см. §2.3). Поэтому  $\|A\|_s = \|\Lambda\|_s = \rho(\Lambda) = \sqrt{\rho(A^*A)}$ .  $\square$

**Теорема 6.3.1.2.** Если матрица  $A$  нормальна, то  $\|A\|_s = \rho(A)$ .

**Доказательство.** Нормальную матрицу  $A$  можно представить в виде  $A = U^*\Lambda U$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  и  $U$  — унитарная матрица. Поэтому  $A^*A = U^*\Lambda\bar{\Lambda}U$ . Пусть  $Ae_i = \lambda_i e_i$  и  $x_i = U^{-1}e_i$ . Тогда  $A^*Ax_i = |\lambda_i|^2 x_i$ , а значит,  $\rho(A^*A) = \rho(A)^2$ .  $\square$

### 6.3.2. Евклидова норма

Евклидовой нормой матрицы  $A$  называют величину  $\|A\|_e = \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{\text{tr}(A^*A)}$ , где  $\sigma_i$  — сингулярные числа матрицы  $A$ .

Если  $U$  — унитарная матрица, то  $\|AU\|_e = \sqrt{\text{tr}(U^*A^*AU)} = \sqrt{\text{tr}(A^*A)} = \|A\|_e$  и  $\|UA\|_e = \|A\|_e$ .

**Теорема 6.3.2.1.** Если  $A$  — матрица порядка  $n$ , то

$$\|A\|_s \leq \|A\|_e \leq \sqrt{n}\|A\|_s.$$

**Доказательство.** Пусть  $\sigma_1, \dots, \sigma_n$  — сингулярные числа матрицы  $A$ , причём  $\sigma_1 \geq \dots \geq \sigma_n$ . Тогда  $\|A\|_s^2 = \sigma_1^2$  и  $\|A\|_e^2 = \sigma_1^2 + \dots + \sigma_n^2$ . Ясно, что  $\sigma_1^2 \leq \sigma_1^2 + \dots + \sigma_n^2 \leq n\sigma_1^2$ .  $\square$

**Замечание.** Евклидова и спектральная норма не случайно выражаются через сингулярные значения матрицы; это связано с их инвариантностью относительно действия группы унитарных матриц. Если  $f(A)$  — произвольная функция от матрицы  $A$  и  $f(A) = f(UA) = f(AU)$  для любой унитарной матрицы  $U$ , то  $f$  зависит лишь от сингулярных значений матрицы  $A$ . В самом деле,  $A = U\Lambda V$ , где  $\Lambda = \text{diag}(\sigma_1, \dots, \sigma_n)$ , а  $U$  и  $V$  — унитарные матрицы. Следовательно,  $f(A) = f(\Lambda)$ .

### 6.3.3. Наилучшее приближение эрмитовой или унитарной матрицей

**Теорема 6.3.3.1.** Пусть  $A$  — произвольная матрица,  $S$  — эрмитова матрица. Тогда  $\|A - (A + A^*)/2\|_e \leq \|A - S\|_e$ , причём равенство достигается только при  $S = (A + A^*)/2$ .

**Доказательство.** Легко проверить, что

$$\begin{aligned} \operatorname{tr}(A - S)^*(A - S) - \operatorname{tr}\left(A - \frac{A + A^*}{2}\right)^* \left(A - \frac{A + A^*}{2}\right) &= \\ &= \operatorname{tr}\left(S - \frac{A + A^*}{2}\right)^2. \end{aligned}$$

Поэтому

$$\|A - S\|_e^2 - \left\|A - \frac{A + A^*}{2}\right\|_e^2 = \left\|S - \frac{A + A^*}{2}\right\|_e^2 \geq 0,$$

причём равенство достигается только при  $S = (A + A^*)/2$ .  $\square$

**Теорема 6.3.3.2.** Если  $S \geq 0$  и  $U$  — унитарная матрица, то  $\|S - I\|_e \leq \|S - U\|_e$ , причём если  $S > 0$ , то равенство достигается тогда и только тогда, когда  $U = I$ .

**Доказательство.** Ясно, что  $(S - U)^*(S - U) = S^2 - (U + U^*)S + I$  и  $(S - I)^*(S - I) = S^2 - 2S + I$ . Поэтому  $\|S - U\|_e^2 - \|S - I\|_e^2 = 2 \operatorname{tr} S - \operatorname{tr}((U + U^*)S)$ . Так как  $|\operatorname{tr}(US)| \leq \operatorname{tr} S$  и  $|\operatorname{tr}(U^*S)| \leq \operatorname{tr} S$  (см. задачу 2.1), то  $\operatorname{tr}((U + U^*)S) \leq |\operatorname{tr}(US)| + |\operatorname{tr}(U^*S)| \leq 2 \operatorname{tr} S$ .

Если  $S > 0$ , то равенство может достигаться, только если  $U = e^{i\varphi} I$ . В этом случае  $\operatorname{tr}((U + U^*)S) = (e^{i\varphi} + e^{-i\varphi}) \operatorname{tr} S$ . Ясно, что если  $e^{i\varphi} + e^{-i\varphi} = 2$ , то  $e^{i\varphi} = 1$ .  $\square$

**Теорема 6.3.3.3.** Пусть  $A = US$  — полярное разложение матрицы  $A$ . Тогда для любой унитарной матрицы  $X$  справедливо неравенство  $\|A - U\|_e \leq \|A - X\|_e$ , причём для невырожденной матрицы  $A$  равенство достигается только при  $X = U$ .

**Доказательство.** Ясно, что  $\|A - X\|_e = \|U - X\|_e = \|U(S - U^{-1}X)\|_e = \|S - U^{-1}X\|_e$ . Согласно теореме 6.3.3.2  $\|S - U^{-1}X\|_e \geq \|S - I\|_e = \|A - U\|_e$ , причём если  $S > 0$ , то равенство достигается, только если  $U^{-1}X = I$ , т.е.  $X = U$ .  $\square$

#### 6.3.4. Наилучшее приближение невырожденной матрицы вырожденной матрицей

**Теорема 6.3.4.1.** Пусть  $A$  — невырожденная матрица,  $X$  — вырожденная матрица. Тогда  $\|A - X\|_s \geq \|A^{-1}\|_s^{-1}$ , причём если  $\|A^{-1}\|_s = \rho(A^{-1})$ , то существует вырожденная матрица  $X$ , для которой  $\|A - X\|_s = \|A^{-1}\|_s^{-1}$ .

**Доказательство.** [Fr2] Возьмём такой вектор  $v$ , что  $Xv = 0$  и  $v \neq 0$ . Тогда

$$\begin{aligned}\|A - X\|_s &\geq \frac{\|(A - X)v\|}{\|v\|} = \frac{\|Ax\|}{\|x\|} \geq \min_x \frac{\|Ax\|}{\|x\|} = \\ &= \min_y \frac{\|y\|}{\|A^{-1}y\|} = \|A^{-1}\|_s^{-1}.\end{aligned}$$

Предположим теперь, что  $\|A^{-1}\|_s = |\lambda^{-1}|$  и  $A^{-1}y = \lambda^{-1}y$ , т.е.  $Ay = \lambda y$ . Тогда  $\|A^{-1}\|_s^{-1} = |\lambda| = \|Ay\|/\|y\|$ . Матрица  $X = A - \lambda I$  вырожденная и  $\|A - X\|_s = \|\lambda I\|_s = |\lambda| = \|A^{-1}\|_s^{-1}$ .  $\square$

### Задачи

**3.1.** Докажите, что  $\|A + B\|_s \leq \|A\|_s + \|B\|_s$  и  $\|A + B\|_e \leq \|A\|_e + \|B\|_e$ .

**3.2.** Докажите, что если  $\lambda \neq 0$  — собственное значение матрицы  $A$ , то  $\|A^{-1}\|_s^{-1} \leq |\lambda| \leq \|A\|_s$ .

**3.3.** Докажите, что  $\|AB\|_s \leq \|A\|_s \|B\|_s$  и  $\|AB\|_e \leq \|A\|_e \|B\|_e$ .

**3.4.** Пусть  $P_1$  и  $P_2$  — два проектора, причём  $\|P_1 - P_2\|_s < 1$ . Докажите, что  $\text{rk } P_1 = \text{rk } P_2$ .

**3.5.** Пусть  $A$  — матрица порядка  $n$ . Докажите, что

$$\|\text{adj } A\|_e \leq n^{\frac{2-n}{2}} \|A\|_e^{n-1}.$$

## 6.4. Дополнение по Шуру и произведение Адамара

### 6.4.1. Дополнение по Шуру положительно определённой матрицы

Пусть  $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  — блочная запись матрицы  $A$ , причём  $|A_{11}| \neq 0$ . Напомним, что дополнением по Шуру матрицы  $A_{11}$  в  $A$  называют матрицу  $(A | A_{11}) = A_{22} - A_{21}A_{11}^{-1}A_{12}$ .

**Теорема 6.4.1.1.** Если  $A > 0$ , то  $(A | A_{11}) > 0$ .

**Доказательство.** Пусть  $T = \begin{pmatrix} I & -A_{11}^{-1}B \\ 0 & I \end{pmatrix}$ , где  $B = A_{12} = A_{21}^*$ . Тогда матрица  $T^*AT = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} - B^*A_{11}^{-1}B \end{pmatrix}$  положительно определена, а значит,  $A_{22} - B^*A_{11}^{-1}B > 0$ .  $\square$

**Замечание.** Аналогично можно доказать, что если  $A \geq 0$  и  $|A_{11}| \neq 0$ , то  $(A | A_{11}) \geq 0$ .

**Теорема 6.4.1.2.** Если  $H$  и  $K$  — произвольные положительно определённые матрицы порядка  $n$ , а  $X$  и  $Y$  — произвольные матрицы размера  $n \times m$ , то

$$X^* H^{-1} X + Y^* K Y - (X + Y)^* (H + K)^{-1} (X + Y) \geq 0.$$

**Доказательство.** Пусть  $T = \begin{pmatrix} I_n & H^{-1}X \\ 0 & I_m \end{pmatrix}$ . Матрица  $A = T^* \begin{pmatrix} H & 0 \\ 0 & 0 \end{pmatrix} T = \begin{pmatrix} H & X^* \\ X^* & X^* H^{-1} X \end{pmatrix}$  неотрицательно определённая. Аналогично матрица  $B = \begin{pmatrix} K & Y^* \\ Y^* & Y^* K^{-1} Y \end{pmatrix}$  неотрицательно определённая. Остаётся применить теорему 6.4.1.1 к дополнению по Шуру матрицы  $H + K$  в матрице  $A + B$ .  $\square$

**Теорема 6.4.1.3.** Пусть  $A, B \geq 0$  и  $A_{11}, B_{11} > 0$ . Тогда

$$(A + B | A_{11} + B_{11}) \geq (A | A_{11}) + (B | B_{11}).$$

**Доказательство.** По определению  $(A + B | A_{11} + B_{11}) = (A_{22} + B_{22}) - (A_{21} + B_{21})(A_{11} + B_{11})^{-1}(A_{12} + B_{12})$ , а согласно теореме 6.4.1.2  $A_{21}A_{11}^{-1}A_{12} + B_{21}B_{11}^{-1}B_{12} \geq (A_{21} + B_{21})(A_{11} + B_{11})^{-1}(A_{12} + B_{12})$ . Поэтому  $(A + B | A_{11} + B_{11}) \geq (A_{22} + B_{22}) - (A_{21}A_{11}^{-1}A_{12} + B_{21}B_{11}^{-1}B_{12}) = (A | A_{11}) + (B | B_{11})$ .  $\square$

Полученные результаты можно применить к доказательству следующего утверждения.

**Теорема 6.4.1.4.** Пусть  $A_k$  и  $B_k$  — угловые подматрицы порядка  $k$  положительно определённых матриц  $A$  и  $B$  порядка  $n$ . Тогда

$$|A + B| \geq |A| \left( 1 + \sum_{k=1}^{n-1} \frac{|B_k|}{|A_k|} \right) + |B| \left( 1 + \sum_{k=1}^{n-1} \frac{|A_k|}{|B_k|} \right).$$

**Доказательство.** [На3] Заметим сначала, что согласно теореме 6.4.1.3 и задаче 1.1

$$\begin{aligned} |(A + B | A_{11} + B_{11})| &\geq |(A | A_{11}) + (B | B_{11})| \geq \\ &\geq |(A | A_{11})| + |(B | B_{11})| = |A|/|A_{11}| + |B|/|B_{11}|. \end{aligned}$$

При  $n = 2$  получаем

$$\begin{aligned} |A + B| &= |A_1 + B_1| \cdot |(A + B | A_1 + B_1)| \geq \\ &\geq (|A_1| + |B_1|)(|A|/|A_1| + |B|/|B_1|). \end{aligned}$$

Предположим теперь, что утверждение доказано для матриц порядка  $n-1$  и докажем его для матриц порядка  $n$ . По предположению индукции

$$\begin{aligned} |A_{n-1} + B_{n-1}| &\geq \\ &\geq |A_{n-1}| \left( 1 + \sum_{k=1}^{n-2} \frac{|B_k|}{|A_k|} \right) + |B_{n-1}| \left( 1 + \sum_{k=1}^{n-2} \frac{|A_k|}{|B_k|} \right). \end{aligned}$$

Кроме того, согласно сделанному выше замечанию  $|(A + B \mid A_{n-1} + B_{n-1})| \geq |A|/|A_{n-1}| + |B|/|B_{n-1}|$ . Поэтому

$$\begin{aligned} |A + B| &\geq \left[ |A_{n-1}| \left( 1 + \sum_{k=1}^{n-2} \frac{|B_k|}{|A_k|} \right) + \right. \\ &\quad \left. + |B_{n-1}| \left( 1 + \sum_{k=1}^{n-2} \frac{|A_k|}{|B_k|} \right) \right] \left( \frac{|A|}{|A_{n-1}|} + \frac{|B|}{|B_{n-1}|} \right) \geq \\ &\geq |A| \left( 1 + \sum_{k=1}^{n-2} \frac{|B_k|}{|A_k|} + \frac{|B_{n-1}|}{|A_{n-1}|} \right) + \\ &\quad + |B| \left( 1 + \sum_{k=1}^{n-2} \frac{|A_k|}{|B_k|} + \frac{|A_{n-1}|}{|B_{n-1}|} \right). \end{aligned}$$

□

#### 6.4.2. Произведение Адамара

Если  $A = \|a_{ij}\|_1^n$  и  $B = \|b_{ij}\|_1^n$  — квадратные матрицы, то их *произведением в смысле Адамара* называют матрицу  $C = \|c_{ij}\|_1^n$ , где  $c_{ij} = a_{ij}b_{ij}$ . Произведение в смысле Адамара обозначается  $A \circ B$  (иногда используется обозначение  $A * B$ ).

**Теорема 6.4.2.1 (Шур).** Если  $A, B > 0$ , то  $A \circ B > 0$ .

**Доказательство.** Пусть  $U = \|u_{ij}\|_1^n$  — такая унитарная матрица, что  $A = U^* \Lambda U$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Тогда  $a_{ij} = \sum_p \bar{u}_{pi} \lambda_p u_{pj}$ , а значит,  $\sum_{i,j} a_{ij} b_{ij} \bar{x}_i x_j = \sum_p \lambda_p \sum_{i,j} b_{ij} y_i^p \bar{y}_j^p$ , где  $y_i^p = x_i u_{pi}$ . Все числа  $\lambda_p$  положительны, поэтому остаётся доказать, что если числа  $x_i$  не все нулевые, то числа  $y_i^p$  тоже не все нулевые. Для этого достаточно заметить, что  $\sum_{i,p} |y_i^p|^2 = \sum_{i,p} |x_i u_{pi}|^2 = \sum_i \left( |x_i|^2 \sum_p |u_{pi}|^2 \right) = \sum_i |x_i|^2$ . □

**Теорема 6.4.2.2 (Оппенгейм).** Если  $A, B > 0$ , то

$$\det(A \circ B) \geq (\prod a_{ii}) \det B.$$

**Доказательство.** Для матриц порядка 1 утверждение очевидно. Предположим, что утверждение доказано для матриц порядка  $n - 1$ ; запишем матрицы  $A$  и  $B$  порядка  $n$  в виде

$$A = \begin{pmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

где  $a_{11}$  и  $b_{11}$  — числа. Тогда  $\det(A \circ B) = a_{11}b_{11} \det(A \circ B \mid a_{11}b_{11})$  и  $(A \circ B \mid a_{11}b_{11}) = A_{22} \circ B_{22} - A_{21} \circ B_{21} a_{11}^{-1} b_{11}^{-1} A_{12} \circ B_{12} = A_{22} \circ (B \mid b_{11}) + (A \mid a_{11}) \circ (B_{21} B_{12} b_{11}^{-1})$ . Так как матрицы  $(A \mid a_{11})$  и  $(B \mid b_{11})$  положительно определены (см. теорему 6.4.1.1), то согласно теореме 6.4.2.1 матрицы  $A_{22} \circ (B \mid b_{11})$  и  $(A \mid a_{11}) \circ (B_{21} B_{12} b_{11}^{-1})$  положительно определены. Поэтому  $\det(A \circ B) \geq a_{11}b_{11} \det(A_{22} \circ (B \mid b_{11}))$  (см. задачу 1.1). Согласно предположению индукции  $\det(A_{22} \circ (B \mid b_{11})) \geq a_{22} \dots a_{nn} \det(B \mid b_{11})$ ; ясно также, что  $\det(B \mid b_{11}) = \det B / b_{11}$ .  $\square$

**Замечание.** Равенство достигается, только если матрица  $B$  диагональная.

### Задачи

**4.1.** Докажите, что если  $A$  и  $B$  — положительно определённые матрицы порядка  $n$ , причём  $A \geq B$ , то  $|A + B| \geq |A| + n|B|$ .

**4.2.** [Dj1] Докажите, что любую положительно определённую матрицу  $A$  можно представить в виде  $A = B \circ C$ , где  $B$  и  $C$  — положительно определённые матрицы.

**4.3.** [Dj1] Докажите, что если  $A > 0$  и  $B > 0$ , то  $\text{rk}(A \circ B) \geq \text{rk } B$ .

## 6.5. Неотрицательные матрицы

Понятие положительного числа можно обобщить для матриц разными способами. С одним из этих обобщений — положительно определёнными матрицами — мы уже познакомились. В этом параграфе мы займёмся другим обобщением — положительными матрицами.

Вещественную матрицу  $A = \|a_{ij}\|_1^n$  называют *положительной* (соответственно *неотрицательной*), если все  $a_{ij} > 0$  (соответственно  $a_{ij} \geq 0$ ).

В этом параграфе для положительных матриц будет использовано обозначение  $A > 0$ ; запись  $A > B$  означает, что  $A - B > 0$ . Обратите внимание, что во всех остальных параграфах обозначение  $A > 0$  означает, что  $A$  — эрмитова положительно определённая матрица.

Вектор  $x = (x_1, \dots, x_n)$  называют *положительным*, если  $x_i > 0$ . Обозначение:  $x > 0$ .

### 6.5.1. Неразложимые матрицы

Матрицу  $A$  порядка  $n$  называют *разложимой*, если множество  $\{1, \dots, n\}$  можно разбить на два непустых подмножества  $I$  и  $J$  так, что  $a_{ij} = 0$  при  $i \in I$  и  $j \in J$ . Иными словами, матрица  $A$  разложима, если перестановкой строк и столбцов её можно привести к виду  $\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ , где  $A_{11}$  и  $A_{22}$  — квадратные матрицы.

**Теорема 6.5.1.1.** Если  $A$  — неотрицательная неразложимая матрица порядка  $n$ , то  $(I + A)^{n-1} > 0$ .

**Доказательство.** Для каждого ненулевого неотрицательного вектора  $u$  рассмотрим вектор  $z = (I + A)u = u + Au$ . Предположим, что не все координаты вектора  $u$  положительны. Изменив при необходимости нумерацию векторов базиса, можно считать, что  $u = \begin{pmatrix} u \\ 0 \end{pmatrix}$ , где  $u > 0$ . Тогда  $Au = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} A_{11}u \\ A_{21}u \end{pmatrix}$ . Так как  $u > 0$ ,  $A_{21} \geq 0$  и  $A_{21} \neq 0$ , то  $A_{21}u \neq 0$ . Поэтому у вектора  $z$  на крайней мере на одну положительную координату больше, чем у вектора  $u$ . Следовательно, если  $u \geq 0$  и  $u \neq 0$ , то  $(I + A)^{n-1}u > 0$ . Беря в качестве вектора  $u$  базисные векторы, получаем требуемое.  $\square$

### 6.5.2. Экстремальные векторы

Пусть  $A$  — неотрицательная матрица порядка  $n$  и  $x$  — неотрицательный вектор;  $r_x = \sup\{\rho \geq 0 \mid Ax \geq \rho x\}$ . Пусть, далее,  $r = \sup_{x \geq 0} r_x$ . Супремум достаточно взять не по всем  $x \geq 0$ , а лишь по компактному множеству  $P = \{x \geq 0 \mid \|x\| = 1\}$ . Поэтому существует такой ненулевой неотрицательный вектор  $z$ , что  $Az \geq rz$ , и не существует такого положительного вектора  $w$ , что  $Aw > rw$ .

Неотрицательный вектор  $z$  называют *экстремальным вектором* матрицы  $A$ , если  $Az \geq rz$ .



**Теорема 6.5.2.1.** Если  $A$  — неотрицательная неразложимая матрица, то  $r > 0$  и экстремальный вектор матрицы  $A$  является собственным вектором.

**Доказательство.** Если  $\xi = (1, \dots, 1)$ , то  $A\xi > 0$ , поэтому  $r > 0$ . Пусть  $z$  — экстремальный вектор матрицы  $A$ . Тогда  $Az - rz = \eta \geq 0$ . Предположим, что  $\eta \neq 0$ . Домножив обе части равенства на  $(I + A)^{n-1}$ , получим  $Aw - rw = (I + A)^{n-1}\eta > 0$ , где  $w = (I + A)^{n-1}z > 0$ . Приходим к противоречию.  $\square$

**Замечание 1.** Ненулевой экстремальный вектор  $z$  матрицы  $A$  положителен. В самом деле,  $z \geq 0$  и  $Az = rz$ , поэтому  $(1 + r)^{n-1}z = (I + A)^{n-1}z > 0$ .

**Замечание 2.** Собственный вектор матрицы  $A$ , соответствующий собственному значению  $r$ , единствен с точностью до пропорциональности. В самом деле, пусть  $Ax = rx$  и  $Ay = ry$ , причём  $x > 0$ . Если  $\mu = \min(y_i/x_i)$ , то  $y_i \geq \mu x_i$  и вектор  $z = y - \mu x$  имеет неотрицательные координаты, причём по крайней мере одна из них равна нулю. Предположим, что  $z \neq 0$ . Тогда  $z > 0$ , так как  $z \geq 0$  и  $Az = rz$  (см. замечание 1). Приходим к противоречию.

**Теорема 6.5.2.2.** Пусть  $A$  — неотрицательная неразложимая матрица, а матрица  $B$  такова, что  $|b_{ij}| \leq a_{ij}$ . Тогда если  $\beta$  — собственное значение матрицы  $B$ , то  $|\beta| \leq r$ , причём если  $\beta = re^{i\varphi}$ , то  $|b_{ij}| = a_{ij}$  и  $B = e^{i\varphi} D A D^{-1}$ , где  $D = \text{diag}(d_1, \dots, d_n)$  и  $|d_i| = 1$ .

**Доказательство.** Пусть  $Bu = \beta u$ , где  $u \neq 0$ . Рассмотрим вектор  $y^+ = (|y_1|, \dots, |y_n|)$ . Так как  $\beta y_i = \sum_j b_{ij} y_j$ , то  $|\beta y_i| = \sum_j |b_{ij} y_j| \leq \sum_j a_{ij} |y_j|$ , а значит,  $|\beta| \leq r_{y^+} \leq r$ .

Предположим теперь, что  $\beta = re^{i\varphi}$ . Тогда  $y^+$  — экстремальный вектор матрицы  $A$ , а значит,  $y^+ > 0$  и  $Ay^+ = ry^+$ . Пусть  $B^+ = \|b'_{ij}\|$ , где  $b'_{ij} = |b_{ij}|$ . Тогда  $B^+ \leq A$  и  $B^+ y^+ = ry^+ = Ay^+$ , а так как  $y^+ > 0$ , то  $B^+ = A$ . Рассмотрим матрицу  $D = \text{diag}(d_1, \dots, d_n)$  где  $d_i = y_i/|y_i|$ . Тогда  $y = Dy^+$  и равенство  $Bu = \beta u$  можно переписать в виде  $BDy^+ = \beta Dy^+$ , т.е.  $Cy^+ = ry^+$ , где  $C = e^{-i\varphi} D^{-1} B D$ . Из определения матрицы  $C$  следует, что  $C^+ = B^+ = A$ . Докажем теперь, что  $C^+ = C$ . В самом деле,  $Cy^+ = ry^+ = B^+ y^+ = Ay^+ = C^+ y^+$ , а так как  $C^+ \geq C$  и  $y^+ > 0$ , то  $C^+ y^+ \geq Cy^+$ , причём равенство возможно, только если  $C = C^+ = A$ .  $\square$

### 6.5.3. Канонический вид неразложимой матрицы

**Теорема 6.5.3.1.** Пусть  $A$  — неотрицательная неразложимая матрица,  $k$  — количество её различных собственных значений, по модулю равных наи-

большему собственному значению  $r$ , причём  $k > 1$ . Тогда существует такая матрица перестановки  $P$ , что матрица  $PAP^T$  имеет блочный вид

$$\begin{pmatrix} 0 & A_{12} & 0 & \dots & 0 \\ 0 & 0 & A_{23} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_{k-1,k} \\ A_{k1} & 0 & 0 & \dots & 0 \end{pmatrix}.$$

**Доказательство.** Наибольшие по модулю собственные значения матрицы  $A$  имеют вид  $\alpha_j = r \exp(i\varphi_j)$ . Применив теорему 6.5.2.2 к матрице  $B = A$ , получим  $A = \exp(i\varphi_j) \cdot D_j A D_j^{-1}$ . Поэтому  $p(t) = |A - tI| = |\exp(i\varphi_j) \cdot D_j A D_j^{-1} - tI| = \lambda p(\exp(-i\varphi_j)t)$ , где  $|\lambda| = 1$ . Числа  $\alpha_1, \dots, \alpha_k$  являются корнями многочлена  $p$ , поэтому они инвариантны относительно поворотов на углы  $\varphi_j$ . Учитывая, что собственное значение  $r$  некрратное (см. задачу 5.4), получаем  $\alpha_j = r \exp(2\pi i/k)$ . Пусть  $y_1$  — собственный вектор, соответствующий собственному значению  $\alpha_1 = r \exp(2\pi i/k)$ . Тогда  $y_1^+ > 0$  и  $y_1 = D_1 y_1^+$  (см. доказательство теоремы 6.5.2.2). Существует такая матрица перестановки  $P$ , что  $PD_1 P^T = \text{diag}(e^{i\gamma_1} I_1, \dots, e^{i\gamma_s} I_s)$ , где числа  $e^{i\gamma_1}, \dots, e^{i\gamma_s}$  попарно различны и  $I_1, \dots, I_s$  — единичные матрицы. Если вместо вектора  $y_1$  взять вектор  $e^{-i\gamma_1} y_1$ , то можно считать, что  $\gamma_1 = 0$ .

Разобьём матрицу  $PAP^T$  на блоки  $A_{pq}$  в соответствии с разбиением матрицы  $PD_1 P^T$ . Так как  $A = \exp(i\varphi_j) D_j A D_j^{-1}$ , то

$$PAP^T = \exp(i\varphi_1)(PD_1 P^T)(PAP^T)(PD_1 P^T)^{-1},$$

т.е.  $A_{pq} = \exp[i(\gamma_p - \gamma_q + \frac{2\pi}{k})] A_{pq}$ . Значит, если  $\frac{2\pi}{k} + \gamma_p \not\equiv \gamma_q \pmod{2\pi}$ , то  $A_{pq} = 0$ . Числа  $\gamma_i$  попарно различны, поэтому для любого  $p$  существует не более одного такого числа  $q$ , что  $A_{pq} \neq 0$  (при этом  $q \neq p$ ). Из неразложимости матрицы  $A$  следует, что хотя бы одно такое число  $q$  существует. Следовательно, существует такое отображение  $p \mapsto q(p)$ , что  $A_{p,q(p)} \neq 0$  и  $\frac{2\pi}{k} + \gamma_p \equiv \gamma_{q(p)} \pmod{2\pi}$ . Для  $p = 1$  получаем  $\gamma_{q(1)} \equiv \frac{2\pi}{k} \pmod{2\pi}$ . После перестановки строк и столбцов матрицы  $PAP^T$  можно считать, что  $\gamma_{q(1)} = \gamma_2$ . Продолжив аналогичные рассуждения, можно получить  $\gamma_{q(j-1)} = \gamma_j = 2\pi(j-1)/k$  при  $2 \leq j \leq \min(k, s)$ . Докажем, что  $s = k$ . Предположим сначала, что  $1 < s < k$ . Тогда  $\frac{2\pi}{k} + \gamma_s - \gamma_r \not\equiv 0 \pmod{2\pi}$  при  $1 \leq r \leq s$ . Поэтому  $A_{sr} = 0$  для  $1 \leq r \leq s$ , т.е. матрица  $A$  приводима. Предположим теперь, что  $s > k$ . Тогда  $\gamma_i = 2(i-1)\pi/k$  при  $1 \leq i \leq k$ . Числа  $\gamma_j$  попарно различны при  $1 \leq j \leq s$  и для любого  $i$ , где  $1 \leq i \leq k$ , найдётся такое число  $j$  ( $1 \leq j \leq k$ ), что  $\frac{2\pi}{k} + \gamma_i \equiv \gamma_j \pmod{2\pi}$ . Поэтому  $\frac{2\pi}{k} + \gamma_i \not\equiv \gamma_r \pmod{2\pi}$  при  $1 \leq i \leq k$  и  $k < r \leq s$ , т.е.  $A_{ir} = 0$  при таких  $k$  и

$r$ . В обоих случаях приходим к противоречию, поэтому  $k = s$ . Теперь ясно, что при указанном выборе матрицы перестановки  $P$  матрица  $PAP^T$  имеет требуемый вид.  $\square$

**Следствие.** Если  $A > 0$ , то наибольшее положительное собственное значение матрицы  $A$  по модулю больше всех остальных её собственных значений.

#### 6.5.4. Примитивные матрицы

Неотрицательную матрицу  $A$  называют *примитивной*, если она неразложима и наибольшее по модулю собственное значение единственно.

**Теорема 6.5.4.1.** Если матрица  $A$  примитивна, то  $A^m > 0$  для некоторого  $m$ .

**Доказательство.** [Ma2] Поделив элементы матрицы  $A$  на наибольшее по модулю собственное значение, можно считать, что  $A$  — неразложимая матрица, наибольшее собственное значение которой равно 1, а все остальные собственные значения по модулю меньше 1. Пусть  $S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$  — нормальная жорданова форма матрицы  $A$ . Так как все собственные значения матрицы  $B$  по модулю меньше 1, то  $\lim_{n \rightarrow \infty} B^n = 0$  (см. задачу 2.3 а). Первый столбец  $x^T$  матрицы  $S$  является собственным вектором матрицы  $A$ , соответствующим собственному значению 1 (см. задачу 1.8). Следовательно, этот вектор является экстремальным вектором матрицы  $A$ , а значит,  $x_i > 0$  (см. теорему 6.5.2.1). Аналогично первая строка  $y$  матрицы  $S^{-1}$  состоит из положительных элементов. Поэтому  $\lim_{n \rightarrow \infty} A^n = \lim_{n \rightarrow \infty} S \begin{pmatrix} 1 & 0 \\ 0 & B^n \end{pmatrix} S^{-1} = S \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} S^{-1} = x^T y > 0$ , а значит,  $A^m > 0$  для некоторого  $m$ .  $\square$

**Замечание.** Если  $A \geq 0$  и  $A^m > 0$ , то матрица  $A$  примитивна (см. задачу 5.5).

Неотрицательной матрице  $A$  порядка  $n$  можно сопоставить ориентированный граф с  $n$  вершинами. В этом графе вершины с номерами  $i$  и  $j$  соединены ребром, идущим из  $i$  в  $j$ , если  $a_{ij} > 0$ . Например, матрице

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 1 & 0 & \dots & 0 \end{pmatrix} \quad (6.1)$$

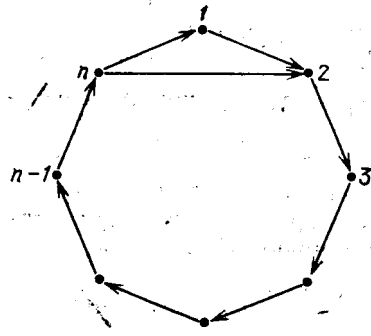


Рис. 6.1.

сопоставляется граф, изображённый на рис. 6.1.

Матрица  $A$  неразложима тогда и только тогда, когда вершины графа можно разбить на множества  $I$  и  $J$  так, что нет рёбер, идущих из вершин множества  $I$  в вершины множества  $J$ .

Индукцией по  $k$  легко проверить, что матрице  $A^k$  сопоставлен граф, в котором вершины с номерами  $i$  и  $j$  соединены ребром, идущим из  $i$  в  $j$ , если для графа, соответствующего матрице  $A$ , существует путь длиной  $k$ , идущий из  $i$  в  $j$  в соответствии с ориентацией рёбер.

**Теорема 6.5.4.2.** Если  $A$  — примитивная матрица порядка  $n$ , то  $A^\gamma > 0$  при  $\gamma = n^2 - 2n + 2$ , причём для матрицы  $A$  порядка  $n \geq 3$ , заданной равенством (6.1), эта оценка точная.

**Доказательство.** Сопоставим матрице  $A$  граф. Пусть  $l_i$  — наименьшая длина пути с началом и концом в вершине  $i$  (если  $a_{ii} > 0$ , то  $l_i = 1$ );  $l$  — наименьшее из чисел  $l_i$ . Для определённости будем считать, что  $l = l_1$  и  $a_{12} > 0, a_{23} > 0, \dots, a_{l1} > 0$ . Пусть  $p$  — произвольная вершина. Длина кратчайшего пути, ведущего из неё в одну из вершин  $1, 2, \dots, l$ , не превосходит  $n - l$ . Его можно продолжить до пути длиной ровно  $n - l$ , ведущего в одну из этих же вершин. Пусть это будет вершина  $j$ . Согласно теореме 6.5.4.1 и замечанию к ней матрица  $A^l$  тоже примитивна. Соединим в соответствующем ей графе вершину  $j$  с вершиной  $q$ . Длина кратчайшего такого пути не превосходит  $n - 1$ . В графе, соответствующем матрице  $A^l$ , вершина  $j$  соединена сама с собой, поэтому путь, соединяющий  $j$  с  $q$ , можно продолжить до пути длиной  $n - 1$ . Следовательно, в графе, соответствующем матрице  $A$ , любые вершины  $p$  и  $q$  можно соединить путём длиной  $n - l + (n - 1)l$ , т.е.  $A^{n-l+(n-1)l} > 0$ . Ясно также, что

$l \leq n - 1$ , так как в противном случае соответствующий граф является циклом и любой степени матрицы  $A$  соответствует цикл. Следовательно,  $n - l + (n - 1)l = (n - 2)l + n \leq (n - 2)(n - 1) + n = n^2 - 2n + 2$ . Остаётся заметить, что если  $A \geq 0$  и  $A^m > 0$ , то  $A^k > 0$  при  $k > m$  (см. задачу 5.1).

Перейдём теперь к доказательству точности нашей оценки. Докажем сначала, что в графе, изображённом на рис. 6.1, нет пути длиной  $n^2 - 2n + 1$ , идущего из вершины 1 в вершину 1, т.е.  $(A^{n^2 - 2n + 1})_{11} = 0$ . Любой путь, идущий из вершины 1 в вершину 1, состоит из  $p$  обходов по большому циклу длиной  $n$  и  $q$  обходов по меньшему циклу длиной  $n - 1$ , причём  $p \geq 1$ . Предположим, что  $pn + q(n - 1) = (n - 1)^2$ . Тогда  $-q \equiv 1 \pmod{n}$ , а значит,  $q \geq n - 1$ . Следовательно,  $q(n - 1) \geq (n - 1)^2$  и  $p \leq 0$ . Получено противоречие.

Докажем теперь, что в рассматриваемом графе для любых двух вершин  $i$  и  $j$  существует путь длиной  $n^2 - 2n + 2$ , идущий из  $i$  в  $j$ . Ясно, что из любой вершины  $i$  в любую вершину  $j \neq i$  ведёт путь длиной  $\alpha$ , где  $0 \leq \alpha \leq n - 1$ . Пройдя из  $i$  в  $j$ , можно дополнительно совершить  $p$  обходов по циклу длиной  $n$  и  $q$  обходов по циклу длиной  $n - 1$ , причём если  $i = j = 1$  и  $q > 0$ , то  $p > 0$ . В случае  $i = j$  положим  $p = 1$  и  $q = n - 2$ . Тогда  $pn + q(n - 1) = n^2 - 2n + 2$ . Докажем теперь, что если  $1 \leq \alpha \leq n - 1$ , то уравнение  $pn + q(n - 1) + \alpha = n^2 - 2n + 2$  имеет неотрицательное целочисленное решение  $(p, q)$ . Так как  $n^2 - 2n + 2 - \alpha \geq n^2 - 3n + 3$ , то достаточно доказать, что уравнение  $pn + q(n - 1) = d$  имеет неотрицательное решение  $(p, q)$  при  $d \geq n^2 - 3n + 3$ . Доказательство проведём индукцией по  $d$ . При  $d = n^2 - 3n + 3$  положим  $p = 1$  и  $q = n - 3$ . Для доказательства индуктивного шага разберём два случая.

1.  $q > 0$ . Тогда  $p'n + q'(n - 1) = d + 1$  для  $p' = p + 1$  и  $q' = q - 1$ .
2.  $q = 0$ . Тогда  $pn = d$  и  $p \geq n - 2$ . Положив  $p' = p - n + 2$  и  $q' = n - 1$ , получим  $p'n + q'(n - 1) = d + 1$ .

Согласно задаче 5.5 рассматриваемая матрица примитивна, поэтому оценка для  $\gamma(A)$  точная.  $\square$

### 6.5.5. Теорема Стильтеса—де Рама

**Теорема 6.5.5.1 (Стильтес—де Рам [Rh]).** Пусть невырожденная матрица  $A = \|a_{ij}\|_1^n$  с вещественными элементами такова, что у неё нет отрицательных собственных значений и  $a_{ij} \leq 0$  при  $i \neq j$ . Тогда матрица  $A^{-1}$  неотрицательная.

**Доказательство.** Пусть  $\Delta = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i > 0, i = 1, \dots, n\}$ ,  $\partial_\Delta$  — граница  $\Delta$ ,  $C_\Delta = \mathbb{R}^n \setminus \Delta$ . Если  $x \in \partial_\Delta$ , то  $x' = Ax \notin \Delta$ . Действительно, если  $x_i = 0$ , то  $x'_i \leq 0$ .

Для  $t \in [0, 1]$  рассмотрим матрицу  $A_t = tA + (1-t)I$ . Внедиагональные элементы матрицы  $A_t$  неотрицательны, поэтому если  $x \in \partial_\Delta$ , то  $Ax \notin \Delta$ . Мы пока не пользовались тем, что у матрицы  $A$  нет отрицательных собственных значений. Это условие нужно для того, чтобы матрица  $A_t$  была невырожденной. Действительно, если  $\det A_t = 0$ , то  $A_t y = 0$  для некоторого  $y \neq 0$ . Равенство  $tAy + (1-t)y = 0$  при  $t \neq 0$  эквивалентно равенству  $Ay = (1-t^{-1})y$  (при  $t = 0$  получаем невырожденную матрицу  $A_0 = I$ ). Таким образом, число  $1-t^{-1} < 0$  является собственным значением матрицы  $A$ , что противоречит условию.

Множество  $A_t \partial_\Delta$  делит пространство  $\mathbb{R}^n$  на множества  $A_t \Delta$  и  $A_t C_\Delta$ . При этом  $A_t \partial_\Delta \cap \Delta = \emptyset$ . Следовательно,  $\Delta \subset A_t \Delta$  или  $\Delta \subset A_t C_\Delta$ . Матрица  $A_0$  единичная, поэтому при  $t = 0$  имеет место первое включение. Соображения непрерывности показывают, что при  $t = 1$  тоже имеет место первое включение, т.е.  $\Delta \subset A\Delta$ . Следовательно,  $A^{-1}\Delta \subset \Delta$ . Но это эквивалентно тому, что все элементы матрицы  $A^{-1}$  неотрицательны.  $\square$

### Задачи

**5.1.** Докажите, что если  $A \geq 0$  и  $A^k > 0$ , то  $A^{k+1} > 0$ .

**5.2.** Докажите, что неотрицательный собственный вектор неразложимой неотрицательной матрицы положителен.

**5.3.** Пусть  $A = \begin{pmatrix} B & C \\ D & E \end{pmatrix}$  — неотрицательная неразложимая матрица, причём  $B$  — квадратная матрица. Докажите, что если  $\alpha$  и  $\beta$  — наибольшие собственные значения матриц  $A$  и  $B$ , то  $\beta < \alpha$ .

**5.4.** Докажите, что если  $A$  — неотрицательная неразложимая матрица, то её наибольшее собственное значение является некратным корнем характеристического многочлена.

**5.5.** Докажите, что если  $A \geq 0$  и  $A^m > 0$ , то матрица  $A$  примитивна.

**5.6.** Докажите, что если  $A$  — неотрицательная неразложимая матрица и  $a_{11} > 0$ , то матрица  $A$  примитивна.

**5.7.** [Si] Матрица  $A$  примитивна. Может ли количество положительных элементов у матрицы  $A$  быть больше, чем у матрицы  $A^2$ ?

### 6.6. Дважды стохастические матрицы

Неотрицательную матрицу  $A = \|a_{ij}\|_1^n$  называют *дважды стохастической*, если  $\sum_{i=1}^n a_{ij} = 1$  и  $\sum_{j=1}^n a_{ij} = 1$  при всех  $i$  и  $j$ .

### 6.6.1. Простейшие свойства

**Теорема 6.6.1.1.** *Произведение дважды стохастических матриц является дважды стохастической матрицей.*

**Доказательство.** Пусть матрицы  $A$  и  $B$  дважды стохастические и  $C = AB$ . Тогда  $\sum_{i=1}^n c_{ij} = \sum_{i=1}^n \sum_{p=1}^n a_{ip} b_{pj} = \sum_{p=1}^n b_{pj} \sum_{i=1}^n a_{ip} = \sum_{p=1}^n b_{pj} = 1$ . Аналогично  $\sum_{j=1}^n c_{ij} = 1$ .  $\square$

**Теорема 6.6.1.2.** *Если матрица  $A = \|a_{ij}\|_1^n$  унитарная, то матрица  $B = \|b_{ij}\|_1^n$ , где  $b_{ij} = |a_{ij}|^2$ , дважды стохастическая.*

**Доказательство.** Достаточно заметить, что  $\sum_{i=1}^n |a_{ij}|^2 = \sum_{j=1}^n |a_{ij}|^2 = 1$ .  $\square$

### 6.6.2. Теорема Биркгофа

**Теорема 6.6.2.1 (Биркгоф).** *Множество всех дважды стохастических матриц порядка  $n$  является выпуклым многогранником с матрицами перестановок в качестве вершин.*

Пусть  $i_1, \dots, i_k$  — некоторые номера строк матрицы  $A$ ,  $j_1, \dots, j_l$  — некоторые номера столбцов. Матрицу  $\|a_{ij}\|$ , где  $i \in \{i_1, \dots, i_k\}$  и  $j \in \{j_1, \dots, j_l\}$ , называют подматрицей матрицы  $A$ . Диагональю матрицы  $A$  называют набор элементов  $a_{1,\sigma(1)}, \dots, a_{n,\sigma(n)}$ , где  $\sigma$  — некоторая перестановка. При доказательстве теоремы Биркгофа нам потребуется следующее утверждение.

**Теорема 6.6.2.2 (Фробениус—Кениг).** *Каждая диагональ матрицы  $A$  порядка  $n$  содержит нулевой элемент тогда и только тогда, когда матрица  $A$  содержит нулевую подматрицу размера  $s \times t$ , где  $s + t = n + 1$ .*

**Доказательство.** Предположим сначала, что на пересечении строк  $i_1, \dots, i_s$  и столбцов  $j_1, \dots, j_t$  стоят нули, причём  $s + t = n + 1$ . Тогда хотя бы одно из  $s$  чисел  $\sigma(i_1), \dots, \sigma(i_s)$  принадлежит множеству  $\{j_1, \dots, j_t\}$ , т.е. соответствующий элемент диагонали нулевой.

Предположим теперь, что каждая диагональ матрицы  $A$  порядка  $n$  содержит нулевой элемент, и докажем, что тогда матрица  $A$  содержит нулевую подматрицу размера  $s \times t$ , где  $s + t = n + 1$ . Доказательство проведём индукцией по  $n$ . При  $n = 1$  утверждение очевидно. Предположим теперь, что утверждение верно для матриц порядка  $n - 1$ , и рассмотрим

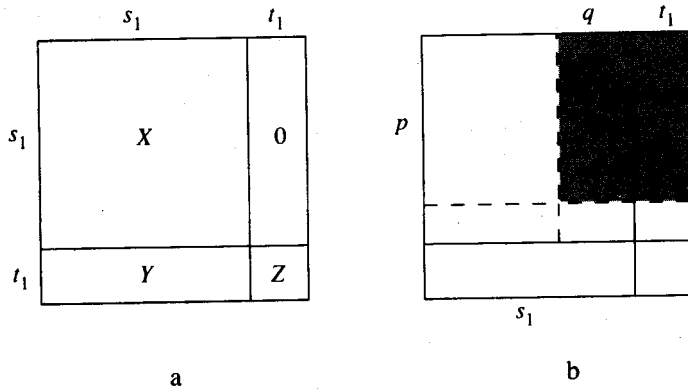


Рис. 6.2.

ненулевую матрицу порядка  $n$ . Выберем в ней ненулевой элемент и выбросим строку и столбец, его содержащие. В полученной матрице порядка  $n - 1$  каждая диагональ содержит нулевой элемент, поэтому у неё есть нулевая подматрица размера  $s_1 \times t_1$ , где  $s_1 + t_1 = n$ . Следовательно, исходную матрицу  $A$  перестановкой строк и столбцов можно привести к блочному виду, изображённом на рис. 6.2 а.

Предположим, что матрица  $X$  имеет диагональ без нулевых элементов. Любую диагональ матрицы  $Z$  можно дополнить этой диагональю до диагонали матрицы  $A$ . Следовательно, любая диагональ матрицы  $Z$  содержит нулевой элемент. В результате получаем, что нулевой элемент содержат либо все диагонали матрицы  $X$ , либо все диагонали матрицы  $Z$ . Пусть, для определённости, все диагонали матрицы  $X$  содержат нулевой элемент. Тогда матрица  $X$  содержит нулевую подматрицу размера  $p \times q$ , где  $p + q = s_1 + 1$ . Следовательно, матрица  $A$  содержит нулевую подматрицу размера  $p \times (t_1 + q)$  (на рис. 6.2 б эта матрица заштрихована). Ясно также, что  $p + (t_1 + q) = s_1 + 1 + t_1 = n + 1$ .  $\square$

**Следствие.** У любой дважды стохастической матрицы есть диагональ, состоящая из положительных элементов.

**Доказательство.** Предположим, что каждая диагональ дважды стохастической матрицы содержит нулевой элемент. Тогда эта матрица содержит нулевую подматрицу размера  $s \times t$ , где  $s + t = n + 1$ . Сумма элементов каждой из рассматриваемых строк и каждого из рассматриваемых столбцов



равна 1; на пересечениях этих строк и столбцов стоят нули, поэтому сумма одних лишь элементов этих строк и столбцов равна  $s + t = n + 1$ , что больше суммы всех элементов, равной  $n$ . Приходим к противоречию.  $\square$

Приступим теперь к доказательству теоремы Биркгофа. Нужно доказать, что любую дважды стохастическую матрицу  $S$  можно представить в виде  $S = \sum \lambda_i P_i$ , где  $P_i$  — матрица перестановки,  $\lambda_i \geq 0$  и  $\sum \lambda_i = 1$ . Применим индукцию по числу  $k$  положительных элементов матрицы  $S$  порядка  $n$ . При  $k = n$  утверждение очевидно, так как в этом случае  $S$  — матрица перестановки. Предположим теперь, что матрица  $S$  не является матрицей перестановки. У этой матрицы есть положительная диагональ (см. следствие теоремы 6.6.2.2); пусть  $P$  — матрица перестановки, соответствующая этой диагонали, и  $x$  — наименьший элемент диагонали. Ясно, что  $x \neq 1$ . Матрица  $T = \frac{1}{1-x}(S - xP)$  дважды стохастическая, причём положительных элементов у неё по крайней мере на один меньше, чем у матрицы  $S$ . По предположению индукции матрицу  $T$  можно представить в требуемом виде, а  $S = xP + (1-x)T$ .  $\square$

**Теорема 6.6.2.3.** *Любая дважды стохастическая матрица  $S$  порядка  $n$  является выпуклой линейной комбинацией не более чем  $n^2 - 2n + 2$  матриц перестановок.*

**Доказательство.** Вычеркнем из матрицы  $S$  последнюю строку и последний столбец. Матрица  $S$  однозначно восстанавливается по оставшимся  $(n-1)^2$  элементам, поэтому множество дважды стохастических матриц порядка  $n$  можно рассматривать как выпуклый многогранник в пространстве размерности  $(n-1)^2$ . Остаётся воспользоваться результатом задачи 3.3.  $\square$

### 6.6.3. Неравенство Г. Вейля

**Теорема 6.6.3.1.** *Пусть  $x_1 \geq x_2 \geq \dots \geq x_n$  и  $y_1 \geq \dots \geq y_n$ , причём  $x_1 + \dots + x_k \leq y_1 + \dots + y_k$  при всех  $k < n$  и  $x_1 + \dots + x_n = y_1 + \dots + y_n$ . Тогда существует такая дважды стохастическая матрица  $S$ , что  $Sy = x$ .*

**Доказательство.** Пусть  $Y = \text{diag}(y_1, \dots, y_n)$ ,  $U = \|u_{ij}\|_1^n$  — ортогональная матрица и  $H = UYU^T$ . Тогда  $h_{ii} = \sum_p |u_{ip}|^2 y_p$ . Таким образом, если  $y$  — столбец  $(y_1, \dots, y_n)$ ,  $h$  — столбец  $(h_{11}, \dots, h_{nn})$  и  $S = \|s_{ij}\|_1^n$ , где  $s_{ij} = |u_{ij}|^2$ , то  $h = Sy$ . Так как  $S$  — дважды стохастическая матрица, то остаётся воспользоваться результатом теоремы 8.10.3.1. Обратите лишь внимание на то, что в формулировке нашей теоремы и теоремы 8.10.3.1 числа упорядочены по-разному.  $\square$

**Теорема 6.6.3.2 (неравенство Г. Вейля).** Пусть  $\alpha_1 \geq \dots \geq \alpha_n$  — модули собственных значений невырожденной матрицы  $A$ ,  $\sigma_1 \geq \dots \geq \sigma_n$  — её сингулярные значения. Тогда  $\alpha_1^s + \dots + \alpha_k^s \leq \sigma_1^s + \dots + \sigma_k^s$  для всех  $k \leq n$  и  $s > 0$ .

**Доказательство.** Согласно теореме 6.2.4.2  $\alpha_1 \dots \alpha_k \leq \sigma_1 \dots \sigma_k$  при  $k \leq n$  и  $\alpha_1 \dots \alpha_n = \sigma_1 \dots \sigma_n$ . Пусть  $x$  и  $y$  — столбцы  $(\ln \alpha_1, \dots, \ln \alpha_n)$  и  $(\ln \sigma_1, \dots, \ln \sigma_n)$ . Согласно теореме 6.6.3.1 существует такая дважды стохастическая матрица  $S$ , что  $x = Sy$ . Фиксируем  $k \leq n$  и для  $u = (u_1, \dots, u_n)$  рассмотрим функцию  $f(u) = f(u_1) + \dots + f(u_k)$ , где  $f(t) = \exp(st)$  — выпуклая функция; функция  $f$  выпукла на множестве векторов с положительными координатами. Фиксируем теперь вектор  $u$  с положительными координатами и рассмотрим функцию  $g(S) = f(Su)$ , определённую на множестве дважды стохастических матриц. Если  $0 \leq \lambda \leq 1$ , то  $g(\lambda S + (1 - \lambda)T) = f(\lambda Su + (1 - \lambda)Tu) \leq \lambda f(Su) + (1 - \lambda)f(Tu) = \lambda g(S) + (1 - \lambda)g(T)$ , т.е. функция  $g$  выпуклая. Выпуклая функция, определённая на выпуклом многограннике, принимает наибольшее значение в одной из его вершин. Следовательно,  $g(S) \leq g(P)$ , где  $P$  — матрица некоторой перестановки  $\pi$  (см. теорему 6.6.2.1). В результате получаем  $f(x) = f(Sy) = g(S) \leq g(P) = f(y_{\pi(1)}, \dots, y_{\pi(n)})$ . Остаётся заметить, что  $f(x) = \exp(s \ln \alpha_1) + \dots + \exp(s \ln \alpha_k) = \alpha_1^s + \dots + \alpha_k^s$  и  $f(y_{\pi(1)}, \dots, y_{\pi(n)}) = \sigma_{\pi(1)}^s + \dots + \sigma_{\pi(k)}^s \leq \sigma_1^s + \dots + \sigma_k^s$ .  $\square$

#### 6.6.4. Теорема Гофмана–Виландта

В качестве примера использования теоремы Биркгофа докажем следующее утверждение.

**Теорема 6.6.4.1 (Гофман–Виландт [Ho]).** Пусть  $A$  и  $B$  — нормальные матрицы,  $\alpha_1, \dots, \alpha_n$  и  $\beta_1, \dots, \beta_n$  — их собственные значения. Тогда

$$\|A - B\|_e^2 \geq \min \sum_{i=1}^n (\alpha_{\sigma(i)} - \beta_i)^2,$$

где минимум берётся по всем перестановкам  $\sigma$ .

**Доказательство.** Пусть  $A = V\Lambda_a V^*$  и  $B = W\Lambda_b W^*$ , где  $\Lambda_a = \text{diag}(\alpha_1, \dots, \alpha_n)$ ,  $\Lambda_b = \text{diag}(\beta_1, \dots, \beta_n)$ ,  $V$  и  $W$  — унитарные матрицы. Тогда

$$\|A - B\|_e^2 = \|V\Lambda_a V^* - W\Lambda_b W^*\|_e^2 = \|U\Lambda_a U^* - \Lambda_b\|_e^2,$$

где  $U = W^*V$ . Кроме того,

$$\begin{aligned}\|U\Lambda_a U^* - \Lambda_b\|_e^2 &= \text{tr}(U\Lambda_a U^* - \Lambda_b)(U\Lambda_a^* U^* - \Lambda_b^*) = \\ &= \text{tr}(\Lambda_a \Lambda_a^* + \Lambda_b \Lambda_b^*) - 2 \text{Re tr}(U\Lambda_a U^* \Lambda_b^*) = \\ &= \sum_{i=1}^n (|\alpha_i|^2 + |\beta_i|^2) - 2 \sum_{i,j=1}^n |u_{ij}|^2 \text{Re}(\bar{\beta}_i \alpha_j).\end{aligned}$$

А так как матрица  $\|c_{ij}\|$ , где  $c_{ij} = |u_{ij}|^2$ , дважды стохастическая, то

$$\|A - B\|_e^2 \geq \sum_{i=1}^n (|\alpha_i|^2 + |\beta_i|^2) - 2 \min \sum_{i,j=1}^n c_{ij} \text{Re}(\bar{\beta}_i \alpha_j),$$

где минимум берётся по всем дважды стохастическим матрицам  $C$ . При фиксированных наборах чисел  $\alpha_i, \beta_j$  нужно найти минимум линейной функции на выпуклом многограннике, вершины которого — матрицы перестановок. Этот минимум достигается в одной из вершин, т.е. для некоторой матрицы  $c_{ij} = \delta_{i,\sigma(i)}$ . В этом случае  $2 \sum_{i,j=1}^n c_{ij} \text{Re}(\bar{\beta}_i \alpha_j) = 2 \sum_{i,j=1}^n c_{ij} \text{Re}(\bar{\beta}_i \alpha_{\sigma(i)})$ . Следовательно,

$$\|A - B\|_e^2 \geq \sum_{i=1}^n (|\alpha_{\sigma(i)}|^2 + |\beta_i|^2 - 2 \text{Re}(\bar{\beta}_i \alpha_{\sigma(i)})) = \sum_{i=1}^n (\alpha_{\sigma(i)} - \beta_i)^2.$$

□

## Задачи

**6.1.** [Mi4] Пусть  $A = \|a_{ij}\|_1^n$  — дважды стохастическая матрица;  $x_1 \geq \dots \geq x_n \geq 0$  и  $y_1 \geq \dots \geq y_n \geq 0$ . Докажите, что  $\sum_{r,s} a_{rs} x_r y_s \leq \sum_r x_r y_r$ .

**6.2.** [Be3] Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения эрмитовой матрицы  $H = \|h_{ij}\|$ . Докажите, что точка с координатами  $(h_{11}, \dots, h_{nn})$  принадлежит выпуклой оболочке точек, координаты которых получаются из чисел  $\lambda_1, \dots, \lambda_n$  всевозможными перестановками.

**6.3.** Пусть  $S$  — стохастическая по строкам матрица, т.е. матрица  $S$  неотрицательная и сумма элементов каждой строки равна 1. Докажите, что для любого натурального  $n$  наибольший элемент каждого столбца матрицы  $I + S + S^2 + \dots + S^n$  лежит на диагонали.

## Решения

**1.1.** Существует такая матрица  $P$ , что  $P^*AP = I$  и  $P^*BP = \text{diag}(\mu_1, \dots, \mu_n)$ , где  $\mu_i \geq 0$ . Поэтому  $|A + B| = d^2 \prod (1 + \mu_i)$ ,  $|A| = d^2$  и  $|B| = d^2 \prod \mu_i$ , где  $d = |\det P|$ . Ясно также, что  $\prod (1 + \mu_i) = 1 + (\mu_1 + \dots + \mu_n) + \dots + \prod \mu_i \geq 1 + \prod \mu_i$ . Неравенство строгое, если  $\mu_1 + \dots + \mu_n > 0$ , т.е. хотя бы одно из чисел  $\mu_1, \dots, \mu_n$  отлично от нуля.

**1.2.** Как и в предыдущей задаче,  $\det(A + iB) = d^2 \prod (\alpha_k + i\beta_k)$  и  $\det A = d^2 \prod \alpha_k$ , где  $\alpha_k > 0$  и  $\beta_k \in \mathbb{R}$ . Так как  $|\alpha_k + i\beta_k|^2 = |\alpha_k|^2 + |\beta_k|^2$ , то  $|\alpha_k + i\beta_k| \geq \alpha_k > 0$ , причём неравенство строгое, если  $\beta_k \neq 0$ .

**1.3.** Так как  $A - B = C > 0$ , то  $A_k = B_k + C_k$ , где  $A_k, B_k, C_k > 0$ . Поэтому  $|A_k| > |B_k| + |C_k|$  (см. задачу 1.1).

**1.4.** Нет, не верно. Пусть  $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  и  $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ . Легко проверить, что  $A \geq B \geq 0$ , но матрица  $A^2 - B^2 = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}$  не является неотрицательно определённой, поскольку её определитель равен  $-1$ .

**1.5.** Пусть  $x + iy$  — ненулевой собственный вектор матрицы  $C$ , соответствующий нулевому собственному значению. Тогда  $(A + iB)(x + iy) = (Ax - By) + i(Bx + Ay) = 0$ , т.е.  $Ax = By$  и  $Ay = -Bx$ . Поэтому  $0 \leq (Ax, x) = (By, x) = (y, Bx) = -(y, Ay) \leq 0$ , т.е.  $(Ax, x) = (Ay, y) = 0$ . Следовательно,  $Ax = Ay = 0$ , а значит,  $Ax = Bx = 0$  и  $Ay = By = 0$ , причём хотя бы один из векторов  $x$  и  $y$  отличен от нуля.

**1.6.** Пусть  $z = (z_1, \dots, z_n)$ . Квадратичная форма  $Q$ , соответствующая рассматриваемой матрице, имеет вид  $2 \sum_{i=1}^n x_i z_0 z_i + (Az, z) = 2z_0(z, x) + (Az, z)$ . Форма  $Q$  положительно определена на подпространстве коразмерности 1, поэтому остаётся доказать, что квадратичная форма  $Q$  не является положительно определённой. Если  $x \neq 0$ , то  $(z, x) \neq 0$  для некоторого  $z$ . Поэтому число  $z_0$  можно подобрать так, что  $2z_0(z, x) + (Az, z) < 0$ .

**1.7.** Существует такая унитарная матрица  $U$ , что  $U^*AU = \text{diag}(\lambda_1, \dots, \lambda_n)$ , где  $\lambda_i \geq 0$ . Кроме того,  $\text{tr}(AB) = \text{tr}(U^*AUB')$ , где  $B' = U^*BU$ . Поэтому можно считать, что  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ . В этом случае  $\text{tr}(AB)/n = (\sum \lambda_i b_{ii})/n \geq (\prod \lambda_i b_{ii})^{1/n} = |A|^{1/n} (\prod b_{ii})^{1/n}$ . Остаётся заметить, что согласно неравенству Адамара  $\prod b_{ii} \geq |B| = 1$ .

**1.8.** Симметрические матрицы  $A - \lambda_n I$ ,  $A - \lambda_1 I$  и  $A^{-1}$  попарно коммутируют, поэтому существует базис, в котором все они диагональны (см. §7.1.2). При этом  $A - \lambda_n I \leq 0$ ,  $A - \lambda_1 I \geq 0$  и  $A^{-1} > 0$ , значит,  $(A - \lambda_n I)(A - \lambda_1 I)A^{-1} \leq 0$ . Для любого вектора  $x$  получаем

$$((A + \lambda_1 \lambda_n A^{-1})x, x) \leq (\lambda_1 + \lambda_n)(x, x).$$

Для вектора  $x$  единичной длины получаем

$$(Ax, x) + \lambda_1 \lambda_n (A^{-1}x, x) \leq \lambda_1 + \lambda_n.$$

Пусть  $u = \lambda_1 \lambda_n (A^{-1}x, x)$ . Домножив неравенство  $(Ax, x) \leq \lambda_1 + \lambda_n - u$  на  $u$ , получим

$$u(Ax, x) \leq (\lambda_1 + \lambda_n)u - u^2 \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n},$$

т.е.

$$(Ax, x)(A^{-1}x, x) \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1 \lambda_n}.$$

**2.1.** Пусть  $S = V^*DV$ , где  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ ,  $V$  — унитарная матрица. Тогда  $\text{tr}(US) = \text{tr}(UV^*DV) = \text{tr}(VUV^*D)$ . Пусть  $VUV^* = W = \|w_{ij}\|_1^n$ . Тогда  $\text{tr}(US) = \sum w_{ii}\lambda_i$ . Так как  $W$  — унитарная матрица, то  $|w_{ii}| \leq 1$ , а значит,  $|\sum w_{ii}\lambda_i| \leq \sum |\lambda_i| = \text{tr} S$ .

Если  $S > 0$ , т.е.  $\lambda_i \neq 0$ , то  $\text{tr} S = \text{tr}(US)$  тогда и только тогда, когда  $w_{ii} = 1$ , т.е.  $W = I$ , а значит,  $U = I$ . Равенство  $\text{tr} S = |\text{tr}(US)|$  для положительно определённой матрицы  $S$  может выполняться, только если  $w_{ii} = e^{i\varphi}$ , т.е.  $U = e^{i\varphi}I$ .

**2.2.** Пусть  $\alpha_1 \geq \dots \geq \alpha_n \geq 0$  и  $\beta_1 \geq \dots \geq \beta_n \geq 0$  — собственные значения матриц  $A$  и  $B$ . Для неотрицательно определённых матриц собственные значения совпадают с сингулярными значениями, поэтому  $|\text{tr}(AB)| \leq \sum \alpha_i \beta_i \leq (\sum \alpha_i)(\sum \beta_i) = \text{tr} A \text{tr} B$  (см. теорему 6.2.4.3).

**2.3.** а) Если  $\lim_{k \rightarrow \infty} A^k = 0$  и  $Ax = \lambda x$ , то  $\lim_{k \rightarrow \infty} \lambda^k = 0$ . Докажем теперь, что если  $|\lambda_i| < 1$  для всех собственных значений матрицы  $A$ , то  $\lim_{k \rightarrow \infty} A^k = 0$ . Это утверждение достаточно доказать для жордановой клетки  $A = \lambda I + N$ . Ясно, что  $(\lambda I + N)^k = \sum \binom{k}{i} \lambda^{k-i} N^i$  и  $N^i = 0$  при  $i \geq n$ , где  $n$  — порядок матрицы  $A$ . Остаётся заметить, что  $\binom{k}{i} < \frac{k^i}{i!}$  и  $\lim_{k \rightarrow \infty} k^i \lambda^k = 0$ .

б) Если  $Ax = \lambda x$  и матрица обладает указанными свойствами, то  $0 < (Hx - A^*HAx, x) = (Hx, x) - (H\lambda x, \lambda x) = (Hx, x)(1 - |\lambda|^2)$ . Поэтому  $|\lambda| < 1$ , так как  $(Hx, x) > 0$ .

Предположим теперь, что  $\lim_{k \rightarrow \infty} A^k = 0$ . Тогда  $\lim_{k \rightarrow \infty} (A^*)^k = 0$  и  $\lim_{k \rightarrow \infty} (A^*)^k A^k = 0$ . Легко проверить, что если  $\lim_{k \rightarrow \infty} B_k = 0$ , то все собственные значения матриц  $B_k$  стремятся к нулю. Поэтому все собственные значения матрицы  $(A^*)^m A^m$  по модулю меньше 1 при некотором  $m$ . Рассмотрим матрицу  $H = I + A^*A + \dots + (A^*)^{m-1} A^{m-1}$ . Легко проверить, что  $H - A^*HA = I - (A^*)^m A^m$ . Собственные значения последней матрицы равны  $1 - \alpha_i$ , где  $\alpha_i$  — собственные значения матрицы  $(A^*)^m A^m$ . Так как  $\alpha_i < 1$ , то  $1 - \alpha_i > 0$ .

2.4. Ясно, что

$$\begin{aligned} \operatorname{Re} \operatorname{tr} AB &= \operatorname{Re} \sum_{i,j=1}^n a_{ij} b_{ji} = \frac{1}{2} \sum_{i,j=1}^n (a_{ij} b_{ji} + \bar{a}_{ij} \bar{b}_{ji}), \\ \frac{1}{2} (\operatorname{tr} AA^* + \operatorname{tr} BB^*) &= \frac{1}{2} \sum_{i,j=1}^n (a_{ij} \bar{a}_{ij} + b_{ij} \bar{b}_{ij}) = \frac{1}{2} \sum_{i,j=1}^n (a_{ij} \bar{a}_{ij} + b_{ji} \bar{b}_{ji}) = \\ &= \frac{1}{2} \sum_{i,j=1}^n (|a_{ij}|^2 + |b_{ji}|^2). \end{aligned}$$

Остаётся заметить, что для произвольных комплексных чисел  $z$  и  $w$  выполняется неравенство  $z\bar{w} + \bar{z}w \leq |z|^2 + |w|^2$ . Действительно, если  $z = a + ib$  и  $w = c + id$ , то  $z\bar{w} + \bar{z}w = 2ac - 2bd \leq a^2 + c^2 + b^2 + d^2$ .

2.5. Матрица  $C = AB - BA$  косоеэрмитова, поэтому её собственные значения чисто мнимые, а значит,  $\operatorname{tr}(C^2) \leq 0$ . Из неравенства  $\operatorname{tr}(AB - BA) \leq 0$  получаем  $\operatorname{tr}(AB)^2 + \operatorname{tr}(BA)^2 \leq \operatorname{tr}(ABBA) + \operatorname{tr}(BAAB)$ . Легко проверить, что  $\operatorname{tr}(AB)^2 = \operatorname{tr}(BA)^2$  и  $\operatorname{tr}(ABBA) = \operatorname{tr}(BAAB) = \operatorname{tr}(A^2 B^2)$ .

2.6. Согласно теореме 4.1.6.1 собственные значения матрицы  $AB$  положительны. Согласно задаче 1.4 матрицы  $A^{-1}$  и  $B^{-1}$  положительно определённые, поэтому собственные значения матрицы  $(AB)^{-1} = B^{-1}A^{-1}$  тоже положительны. Согласно задаче 2.4

$$\begin{aligned} \operatorname{tr} AB &\leq \frac{1}{2} (\operatorname{tr} AA^* + \operatorname{tr} BB^*) = \frac{1}{2} (\operatorname{tr} A^2 + \operatorname{tr} B^2) = \\ &= \frac{1}{2} \sum_{i=1}^n (\alpha_i^2 + \beta_i^2), \\ \operatorname{tr} B^{-1} A^{-1} &\leq \frac{1}{2} (\operatorname{tr} B^{-1} B^{*-1} + \operatorname{tr} A^{-1} A^{*-1}) = \frac{1}{2} \sum_{i=1}^n \left( \frac{1}{\alpha_i^2} + \frac{1}{\beta_i^2} \right). \end{aligned}$$

Все собственные значения матриц  $AB$  и  $B^{-1}A^{-1}$  положительны, поэтому  $\lambda \leq \operatorname{tr} AB \leq \frac{n}{2} (\alpha_1^2 + \beta_1^2)$  и  $\lambda^{-1} \leq \operatorname{tr} B^{-1} A^{-1} \leq \frac{n}{2} \left( \frac{1}{\alpha_n^2} + \frac{1}{\beta_1^2} \right) = \frac{n}{2} \cdot \frac{\alpha_n^2 + \beta_n^2}{\alpha_n^2 \beta_n^2}$ .

2.7. Для матрицы  $A$  можно выбрать унитарную матрицу  $U$  так, что  $U^* A U = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \Lambda$ . Тогда  $\operatorname{tr}(AB) = \operatorname{tr}(U^* A U U^* B U) = \operatorname{tr}(\Lambda C)$ , где  $C = U^* B U$  — положительно определённая матрица. У матрицы  $\Lambda C$  на диагонали стоят числа  $\lambda_1 c_{11}, \dots, \lambda_n c_{nn}$ . Поэтому  $(\operatorname{tr} \Lambda + \operatorname{tr} C)^2 - 4 \operatorname{tr}(\Lambda C) = (\lambda_1 + c_{11})^2 + \dots + (\lambda_n + c_{nn})^2 + 2 \sum_{i \neq j} \lambda_i c_{jj} - 4\lambda_1 c_{11} - \dots - 4\lambda_n c_{nn} > (\lambda_1 - c_{11})^2 + \dots + (\lambda_n - c_{nn})^2 \geq 0$ .

2.8. Если матрица  $A$  невырожденная, то достаточно заметить, что  $A^* A = A^{-1} (A A^*) A$ . В случае вырожденной матрицы можно воспользоваться полярным разложением  $A = S U$  (теорема 3.6.1.1); здесь матрица  $S$  эрмитова, а  $U$  унитарная. Ясно, что  $A A^* = S U U^* S = S^2$  и  $A^* A = U^* S S U = U^{-1} S^2 U$ .

2.9. Собственные значения эрмитовой матрицы  $A^* A$  равны, поэтому  $A^* A = t I$  где  $t \in \mathbb{R}$ . Следовательно, матрица  $U = \sqrt{1/t} A$  унитарная.

2.10. Достаточно применить результат задачи 1.10 к матрице  $A^* A$ .

**2.11.** Достаточно заметить, что  $\begin{vmatrix} -\lambda I & A \\ A^* & -\lambda I \end{vmatrix} = |\lambda^2 I - A^* A| = \pm |A^* A - \lambda^2 I|$  (см. теорему 1.3.1.2).

**2.12.** Запишем матрицу  $A$  в блочном виде:  $A = \begin{pmatrix} 0 & I_{n-1} \\ a_0 & d \end{pmatrix}$ , где  $d = (a_1, \dots, a_{n-1})$ . Тогда

$$AA^* = \begin{pmatrix} 0 & I_{n-1} \\ a_0 & d \end{pmatrix} \begin{pmatrix} 0 & \overline{a_0} \\ I_{n-1} & d^* \end{pmatrix} = \begin{pmatrix} I_{n-1} & d^* \\ d & |a_0|^2 + \|d\|^2 \end{pmatrix}.$$

Поэтому согласно теореме 1.3.1.4 характеристический многочлен матрицы  $AA^*$  равен

$$(|a_0|^2 + \dots + |a_{n-1}|^2 - \lambda)(1 - \lambda)^{n-1} - d \operatorname{adj}((1 - \lambda)I_{n-1})d^*.$$

Согласно теореме 1.2.5.1 в)  $\operatorname{adj}((1 - \lambda)I_{n-1}) = (1 - \lambda)^{n-2}I_{n-1}$ . Деля характеристический многочлен на  $(1 - \lambda)^{n-2}$ , после несложных вычислений получаем требуемое.

**2.13.** Рассмотрим матрицу  $\Lambda^m C = \Lambda^m A \Lambda^m B$ . Здесь  $\Lambda^m C$  и  $\Lambda^m B$  — столбцы длиной  $\binom{n}{m}$ , а  $\Lambda^m A$  — квадратная матрица порядка  $\binom{n}{m}$ . Согласно теореме 6.2.4.1

$$\sigma_{\min} \|\Lambda^m B\| \leq \|\Lambda^m A \Lambda^m B\| \leq \sigma_{\max} \|\Lambda^m B\|,$$

где  $\sigma_{\max}$  и  $\sigma_{\min}$  — наибольшее и наименьшее сингулярные значения матрицы  $\Lambda^m A$ .

Сингулярные значения матриц  $\Lambda^m C$  и  $\Lambda^m B$  равны  $\gamma_1 \dots \gamma_m$  и  $\beta_1 \dots \beta_m$  (напомним, что для вектора—столбца сингулярное значение — это его длина). Сингулярные значения матрицы  $\Lambda^m A$  равны  $\alpha_{i_1} \dots \alpha_{i_m}$ , где числа  $i_1, \dots, i_m$  попарно различны. Поэтому  $\sigma_{\max} = \alpha_1 \dots \alpha_m$  и  $\sigma_{\min} = \alpha_{n-m+1} \dots \alpha_n$ .

**2.14.** Согласно теореме 4.1.4.1 существует положительно определённая матрица  $S$ , для которой  $S^2 = H$ . Применив результат задачи 2.13 к матрицам  $A = S$  и  $B = M$ , получаем требуемое.

**3.1.** Ясно, что

$$\begin{aligned} \|A + B\|_s &= \sup_{x \neq 0} \frac{\|Ax + Bx\|}{\|x\|} \leq \sup_{x \neq 0} \left( \frac{\|Ax\|}{\|x\|} + \frac{\|Bx\|}{\|x\|} \right) = \\ &= \|A\|_s + \|B\|_s. \end{aligned}$$

Пусть  $a$  и  $b$  — векторы в  $n^2$ -мерном пространстве с координатами  $a_{ij}$  и  $b_{ij}$ . Тогда  $\|A + B\|_e = \|a + b\| \leq \|a\| + \|b\| = \|A\|_e + \|B\|_e$ .

**3.2.** Предположим, что  $Ax = \lambda x$ , причём  $\lambda x \neq 0$ . Тогда  $A^{-1}x = \lambda^{-1}x$ . Поэтому  $\max_y \frac{\|Ay\|}{\|y\|} \geq \frac{\|Ax\|}{\|x\|} = \lambda$  и  $\left( \max_y \frac{\|A^{-1}y\|}{\|y\|} \right)^{-1} = \min_y \frac{\|y\|}{\|A^{-1}y\|} \leq \frac{\|x\|}{\|A^{-1}x\|} = \lambda$ .

**3.3.** Если  $\|AB\|_s \neq 0$ , то  $\|AB\|_s = \max_x \frac{\|ABx\|}{\|x\|} = \frac{\|ABx_0\|}{\|x_0\|}$ , где  $Bx_0 \neq 0$ . Пусть  $y = Bx_0$ . Тогда

$$\frac{\|ABx_0\|}{\|x_0\|} = \frac{\|Ay\|}{\|y\|} \cdot \frac{\|Bx_0\|}{\|x_0\|} \leq \|A\|_s \|B\|_s.$$

Согласно задаче 2.2  $\|AB\|_e^2 = \text{tr}(B^*A^*AB) = \text{tr}(BB^*A^*A) \leq \text{tr}(BB^*)\text{tr}(A^*A) = \|A\|_e^2\|B\|_e^2$ .

**3.4.** Равенство  $(I \pm (P_1 - P_2))x = x \pm (P_1 - P_2)x$  показывает, что

$$\|(I \pm (P_1 - P_2))x\| \geq \|x\| - \|(P_1 - P_2)x\| \geq (1 - \rho)\|x\|,$$

где  $\rho = \|P_1 - P_2\|_s < 1$ . Поэтому операторы  $I \pm (P_1 - P_2)$  невырожденные, а значит,

$$\text{rk } P_1 = \text{rk } P_1(I - P_1 + P_2) = \text{rk } P_1 P_2,$$

$$\text{rk } P_2 = \text{rk } P_2(I - P_2 + P_1) = \text{rk } P_2 P_1.$$

Остаётся заметить, что  $\text{rk } P_1 P_2 \leq \min\{\text{rk } P_1, \text{rk } P_2\}$  и  $\text{rk } P_2 P_1 \leq \min\{\text{rk } P_1, \text{rk } P_2\}$ .

**3.5.** Пусть  $\sigma_1, \dots, \sigma_n$  — сингулярные значения матрицы  $A$ . Тогда сингулярные значения матрицы  $\text{adj } A$  равны  $\prod_{i \neq 1} \sigma_i, \dots, \prod_{i \neq n} \sigma_i$ , (задача 2.10), поэтому  $\|A\|_e^2 = \sigma_1^2 + \dots + \sigma_n^2$  и  $\|\text{adj } A\|_e^2 = \prod_{i \neq 1} \sigma_i^2 + \dots + \prod_{i \neq n} \sigma_i^2$ . Предположим сначала, что матрица  $A$  невырожденная. Тогда  $\|\text{adj } A\|_e^2 = (\sigma_1^2 \dots \sigma_n^2) \left( \frac{1}{\sigma_1^2} + \dots + \frac{1}{\sigma_n^2} \right)$ . Перемножив неравенства  $\sigma_1^2 \dots \sigma_n^2 \leq n^{-n}(\sigma_1^2 + \dots + \sigma_n^2)^n$  и  $\left( \frac{1}{\sigma_1^2} + \dots + \frac{1}{\sigma_n^2} \right) (\sigma_1^2 + \dots + \sigma_n^2) \leq n^2$ , получим  $\|\text{adj } A\|_e^2 \leq n^{2-n} \|A\|_e^{2(n-1)}$ . Обе части этого неравенства непрерывно зависят от элементов матрицы  $A$ , поэтому неравенство верно и для вырожденных матриц.

Неравенство обращается в равенство, если  $\sigma_1 = \dots = \sigma_n$ , т.е. матрица  $A$  пропорциональна унитарной матрице (см. задачу 2.9).

**4.1.** Согласно теореме 6.4.1.4

$$|A + B| \geq |A| \left( 1 + \sum_{k=1}^{n-1} \frac{|B_k|}{|A_k|} \right) + |B| \left( 1 + \sum_{k=1}^{n-1} \frac{|A_k|}{|B_k|} \right).$$

Кроме того,  $|A_k|/|B_k| \geq 1$  (см. задачу 1.3).

**4.2.** Пусть  $B(\lambda) = \|b_{ij}\|_1^n$ , где  $b_{ii} = 1$  и  $b_{ij} = \lambda$  при  $i \neq j$ . Тогда  $B(\lambda) > 0$  при  $0 < \lambda < 1$  (см. задачу 1.7), а значит,  $C(\lambda) = A \circ B(\lambda) > 0$ . Кроме того,  $\lim_{\lambda \rightarrow 1} C(\lambda) = A$ . Поэтому существует такое число  $\lambda_0$ , что  $C(\lambda_0) > 0$  и  $\lambda_0 < 1$ . Тогда  $A = C(\lambda_0) \circ B(\lambda_0)^{-1}$  — искомое разложение.

**4.3.** Если  $B > 0$ , то можно воспользоваться теоремой Шура (теорема 6.4.2.1). Предположим теперь, что  $\text{rk } B = k$ , где  $0 < k < \text{rk } A$ . Тогда матрица  $B$  содержит положительно определённую главную подматрицу  $M(B)$  ранга  $k$  (см. задачу 1.2). Пусть  $M(A)$  — соответствующая подматрица  $A$ ; так как  $A > 0$ , то  $M(A) > 0$ . По теореме Шура подматрица  $M(A) \circ M(B)$  матрицы  $A \circ B$  невырождена.

**5.1.** Пусть  $A = \|a_{ij}\|_1^n$ ,  $A^k = B = \|b_{ij}\|_1^n$  и  $BA = C = \|c_{ij}\|_1^n$ . Тогда  $c_{pq} = b_{p1}a_{1q} + \dots + b_{pn}a_{nq}$ , причём  $b_{ij} > 0$  и  $a_{ij} \geq 0$ . Предположим, что  $c_{pq} = 0$ . Тогда



$a_{1q} = \dots = a_{nq} = 0$ , т.е.  $q$ -й столбец матрицы  $A$  нулевой. Следовательно,  $q$ -й столбец матрицы  $B = A^k$  тоже нулевой. Получено противоречие.

**5.2.** Предположим, что данный собственный вектор не является положительным. Можно считать, что он имеет вид  $\begin{pmatrix} x \\ 0 \end{pmatrix}$ , где  $x > 0$ . Тогда

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} Ax \\ Cx \end{pmatrix},$$

а значит,  $Cx = 0$ . А так как  $C \geq 0$ , то  $C = 0$ , поэтому данная матрица разложима. Получено противоречие.

**5.3.** Пусть  $y \geq 0$  — ненулевой собственный вектор матрицы  $B$ , соответствующий собственному значению  $\beta$ ;  $x = \begin{pmatrix} y \\ 0 \end{pmatrix}$ . Тогда  $Ax = \begin{pmatrix} B & C \\ D & E \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} By \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ Dy \end{pmatrix} = \beta x + z$ , где  $z = \begin{pmatrix} 0 \\ Dy \end{pmatrix} \geq 0$ . Равенство  $Ax = \beta x$  выполняться не может, так как собственный вектор неразложимой матрицы положителен (см. задачу 5.2). Кроме того,  $\sup\{t \geq 0 \mid Ax - tx \geq 0\} \geq \beta$ , причём если  $\beta = \alpha$ , то  $x$  — экстремальный вектор (см. теорему 6.5.2.1), а значит,  $Ax = \beta x$ . Поэтому  $\beta < \alpha$ .

**5.4.** Пусть  $f(\lambda) = |A - \lambda I|$ . Легко проверить, что  $f'(\lambda) = -\sum_{i=1}^n |A_i - \lambda I|$ , где  $A_i$  — матрица, полученная из матрицы  $A$  вычёркиванием  $i$ -й строки и  $i$ -го столбца (см. задачу 1.9). Если  $r$  и  $r_i$  — наибольшие собственные значения матриц  $A$  и  $A_i$ , то  $r > r_i$  (см. задачу 5.3). Следовательно, все числа  $|A_i - rI|$  положительны или отрицательны одновременно, так как старшие коэффициенты многочленов  $|A_i - \lambda I|$  совпадают. Поэтому  $f'(r) \neq 0$ .

**5.5.** Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения матрицы  $A$ ,  $\mu_1, \dots, \mu_n$  — собственные значения матрицы  $A^m$ . Тогда  $\mu_i = \lambda_i^m$ . Если  $|\mu_i| \geq \dots \geq |\mu_n|$ , то  $\mu_i > 0$  и  $\mu_1 > |\mu_i|$  при  $i \neq 1$  (см. следствие теоремы 6.5.3.1). Поэтому  $\lambda_1 = \sqrt[m]{\mu_1}$  и  $\lambda_1 > |\lambda_i|$  при  $i \neq 1$ .

**5.6.** Предположим, что матрица  $A$  не примитивна. Тогда для некоторой матрицы перестановки  $P$  матрица  $PAP^T$  имеет вид, указанный в условии теоремы 6.5.3.1. С другой стороны, диагональные элементы матрицы  $PAP^T$  получаются из диагональных элементов матрицы  $A$  перестановкой. Приходим к противоречию.

**5.7.** Да, может. Рассмотрим, например, неотрицательную матрицу  $A$ , соответствующую ориентированному графу  $1 \rightarrow (1, 2)$ ,  $2 \rightarrow (3, 4, 5)$ ,  $3 \rightarrow (6, 7, 8)$ ,  $4 \rightarrow (6, 7, 8)$ ,  $5 \rightarrow (6, 7, 8)$ ,  $6 \rightarrow (9)$ ,  $7 \rightarrow (9)$ ,  $8 \rightarrow (9)$ ,  $9 \rightarrow (1)$ . Легко проверить, что матрица  $A$  неразложима, а так как  $a_{11} > 0$ , то она примитивна (см. задачу 5.6). Матрице  $A^2$  соответствует ориентированный граф  $1 \rightarrow (1, 2, 3, 4, 5)$ ,  $2 \rightarrow (6, 7, 8)$ ,  $3 \rightarrow (9)$ ,  $4 \rightarrow (9)$ ,  $5 \rightarrow (9)$ ,  $6 \rightarrow (1)$ ,  $7 \rightarrow (1)$ ,  $8 \rightarrow (1)$ ,  $9 \rightarrow (1, 2)$ . У первого графа 18 рёбер, а у второго 16.

**6.1.** Существуют такие неотрицательные числа  $\xi_i$  и  $\eta_i$ , что  $x_r = \xi_r + \dots + \xi_n$

и  $y_r = \eta_r + \dots + \eta_n$ . Следовательно,

$$\begin{aligned} \sum_r x_r y_r - \sum_{r,s} a_{rs} x_r y_s &= \sum_{r,s} (\delta_{rs} - a_{rs}) x_r y_s = \\ &= \sum_{r,s} (\delta_{rs} - a_{rs}) \sum_{i \geq r} \xi_i \sum_{j \geq s} \eta_j = \sum_{i,j} \xi_i \eta_j \sum_{r \leq i} \sum_{s \leq j} (\delta_{rs} - a_{rs}). \end{aligned}$$

Достаточно проверить, что  $\sum_{r \leq i} \sum_{s \leq j} (\delta_{rs} - a_{rs}) \geq 0$ . Если  $i \leq j$ , то  $\sum_{r \leq i} \sum_{s \leq j} \delta_{rs} = \sum_{r \leq i} \sum_{s=1}^n \delta_{rs}$ , поэтому  $\sum_{r \leq i} \sum_{s \leq j} (\delta_{rs} - a_{rs}) \geq \sum_{r \leq i} \sum_{s=1}^n (\delta_{rs} - a_{rs}) = 0$ . Случай  $i \geq j$  разбирается аналогично.

**6.2.** Существует такая унитарная матрица  $U$ , что  $H = U \Lambda U^*$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Так как  $h_{ij} = \sum_k u_{ik} \bar{u}_{jk} \lambda_k$ , то  $h_{ii} = \sum_k x_{ik} \lambda_k$ , где  $x_{ik} = |u_{ik}|^2$ . Поэтому  $h = X \lambda$ , где  $\lambda$  — столбец  $(\lambda_1, \dots, \lambda_n)$  и  $X$  — дважды стохастическая матрица. Согласно теореме 6.6.2.1  $X = \sum_{\sigma} t_{\sigma} P_{\sigma}$ , где  $P_{\sigma}$  — матрица перестановки  $\sigma$ ,  $t_{\sigma} \geq 0$  и  $\sum_{\sigma} t_{\sigma} = 1$ . Следовательно,  $h = \sum_{\sigma} t_{\sigma} (P_{\sigma} \lambda)$ .

**6.3.** Пусть  $I + S + S^2 + \dots + S^n = A(n) = \|a_{ij}(n)\|_1^n$ . Матрица  $S^n$  неотрицательная, поэтому из равенства  $A(n) = A(n-1) + S^n$  следует, что  $a_{ij}(n-1) \leq a_{ij}(n)$ .

Докажем по индукции, что  $a_{ij}(n) \leq a_{jj}(n-1)$  при  $i \neq j$ . При  $n=0$  требуемое неравенство легко доказывается:  $a_{ij}(1) = s_{ij} \leq 1 = a_{jj}(0)$ . Чтобы сделать шаг индукции, воспользуемся равенством  $A(n+1) = I + SA(n)$ . Если  $i \neq j$ , то из этого равенства следует, что  $a_{ij}(n+1) = \sum_{k=1}^n s_{ik} a_{kj}(n)$ . Мы уже доказали неравенство  $a_{ij}(n-1) \leq a_{ij}(n)$ . Кроме того, по предположению индукции  $a_{ij}(n) \leq a_{jj}(n-1)$  при  $i \neq j$ . Значит,  $a_{ij}(n) \leq a_{jj}(n)$  (при  $i=j$  это неравенство очевидно). Поэтому  $a_{ij}(n+1) \leq \sum_{k=1}^n s_{ik} a_{jj}(n) = a_{jj}(n)$  при  $i \neq j$ . Таким образом, если  $i \neq j$ , то  $a_{ij}(n) \leq a_{jj}(n-1) \leq a_{jj}(n)$ .

## Глава 7.

# Коммутаторы

### 7.1. Коммутирующие операторы

#### 7.1.1. Матрицы, перестановочные с данной

Матрицы  $A$  и  $B$  одного порядка называют *коммутирующими* (или *перестановочными*), если  $AB = BA$ . Выясним, как устроены все матрицы  $X$ , коммутирующие с данной матрицей  $A$ . Можно рассмотреть даже более общее матричное уравнение  $AX = XB$ , где  $A$  и  $B$  — квадратные матрицы разного размера, так как решить это уравнение не сложнее, чем решить уравнение  $AX = XA$ .

Уравнение  $AX = XB$  эквивалентно уравнению

$$(PAP^{-1})(PXQ^{-1}) = (PXQ^{-1})(QBQ^{-1}),$$

т.е. уравнению  $A'X' = X'B'$ , где  $A' = PAP^{-1}$  и  $B' = QBQ^{-1}$  — жордановы формы матриц  $A$  и  $B$ ,  $X' = PXQ^{-1}$ . Поэтому можно считать, что  $A = \text{diag}(A_1, \dots, A_p)$  и  $B = \text{diag}(B_1, \dots, B_q)$ , где  $A_i$  и  $B_j$  — жордановы клетки. Разобьём строки матрицы  $X$  на  $p$  блоков в соответствии с разбиением матрицы  $A$ , а столбцы матрицы  $X$  разобьём на  $q$  блоков в соответствии с разбиением матрицы  $B$ ; пусть  $X_{ij}$  — соответствующие блоки матрицы  $X$ . Уравнение  $AX = XB$  эквивалентно системе уравнений  $A_i X_{ij} = X_{ij} B_j$  для блоков матрицы  $X$ .

Пусть  $X = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{m1} & \dots & x_{mn} \end{pmatrix}$ ,  $N_n$  и  $N_m$  — нильпотентные матрицы порядка  $n$  и  $m$  соответственно, у которых наддиагональные элементы вида  $a_{i,i+1}$  равны 1, а остальные элементы равны нулю. Рассмотрим

уравнение

$$(\lambda I_m + N_m)X = X(\mu I_n + N_n),$$

т.е.

$$N_m X - X N_n = (\mu - \lambda)X.$$

Легко проверить, что

$$N_m X = \begin{pmatrix} x_{21} & \dots & x_{2n} \\ \dots & \dots & \dots \\ x_{m1} & \dots & x_{mn} \\ 0 & \dots & 0 \end{pmatrix} \quad \text{и} \quad X N_n = \begin{pmatrix} 0 & x_{11} & \dots & x_{1,n-1} \\ \dots & \dots & \dots & \dots \\ 0 & x_{m1} & \dots & x_{m,n-1} \end{pmatrix}.$$

Если  $\lambda \neq \mu$ , то, рассмотрев первый столбец обеих частей уравнения, последовательно получим  $x_{m1} = 0, x_{m-1,1} = 0, \dots, x_{11} = 0$ . Продолжив далее такие же рассуждения для следующих столбцов, получим  $X = 0$ . Предположим теперь, что  $\lambda = \mu$ , т.е.  $N_m X = X N_n$ . Легко проверить, что в этом случае решение  $X$  имеет вид  $(0, Y)$  при  $n > m$ ,  $\begin{pmatrix} Y \\ 0 \end{pmatrix}$  при  $n < m$  и  $Y$  при  $n = m$ , где

$$Y = \begin{pmatrix} y_1 & y_2 & \dots & y_k \\ 0 & y_1 & \dots & y_{k-1} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & y_1 \end{pmatrix} \quad \text{и} \quad k = \min(m, n).$$

Размерность пространства решений равна  $\min(m, n)$ . В результате доказано следующее утверждение.

**Теорема 7.1.1.1.** Пусть  $\lambda$  — собственное значение матрицы  $A$  или матрицы  $B$ , причём этому собственному значению соответствуют жордановы клетки порядков  $a_1(\lambda), \dots, a_r(\lambda)$  и  $b_1(\lambda), \dots, b_s(\lambda)$ ; если  $\lambda$  — собственное значение лишь одной матрицы, то порядки жордановых клеток второй матрицы считаются нулевыми. Тогда размерность пространства решений уравнения  $AX = XB$  равна

$$\sum_{\lambda} \sum_{i,j} \min(a_i(\lambda), b_j(\lambda)).$$

**Следствие.** Размерность пространства решений уравнения  $AX = XA$  равна

$$\sum_{\lambda} \sum_{i,j} \min(a_i(\lambda), a_j(\lambda)).$$

**Теорема 7.1.1.2.** Пусть  $m$  — размерность пространства решений уравнения  $AX = XA$ , где  $A$  — квадратная матрица порядка  $n$ . Тогда следующие условия эквивалентны:

- а)  $m = n$ ;
- б) характеристический многочлен матрицы  $A$  совпадает с минимальным многочленом;
- в) любая матрица, коммутирующая с матрицей  $A$ , является многочленом от  $A$ .

**Доказательство.** Согласно следствию теоремы 7.1.1.1  $m = \sum_{\lambda} \sum_{i,j} \min(a_i(\lambda), a_j(\lambda)) \geq \sum_{\lambda} \sum_i a_i(\lambda) = n$ , причём равенство достигается тогда и только тогда, когда жордановы клетки матрицы  $A$  соответствуют попарно различным собственным значениям, т.е. характеристический многочлен совпадает с минимальным многочленом.

Если характеристический многочлен матрицы  $A$  совпадает с минимальным многочленом, то размерность пространства  $\langle I, A, \dots, A^{n-1} \rangle$  равна  $n$ , поэтому оно совпадает с пространством решений уравнения  $AX = XA$ , т.е. любая матрица, коммутирующая с  $A$ , является многочленом от  $A$ .

Если любая матрица, коммутирующая с  $A$ , является многочленом от  $A$ , то пространство решений уравнения  $AX = XA$  содержится в пространстве  $\langle I, A, \dots, A^{k-1} \rangle$ , причём  $k \leq n$  согласно теореме Гамильтона–Кэли. С другой стороны,  $k \geq m \geq n$ , поэтому  $m = n$ .  $\square$

### 7.1.2. Семейства коммутирующих операторов

**Теорема 7.1.2.1.** Пусть операторы  $A$  и  $B$  в пространстве  $V$  коммутируют и  $V_{\lambda}(k) = \{v \in V \mid (A - \lambda I)^k v = 0\}$ . Тогда  $BV_{\lambda}(k) \subset V_{\lambda}(k)$ .

**Доказательство.** Операторы  $(A - \lambda I)^k$  и  $B$  коммутируют, поэтому если  $v \in V_{\lambda}(k)$ , то  $(A - \lambda I)^k Bv = B(A - \lambda I)^k v = 0$ , т.е.  $Bv \in V_{\lambda}(k)$ .  $\square$

**Следствие.** Пусть операторы  $A$  и  $B$  в пространстве  $V$  коммутируют и  $V = V_1 \oplus V_2$ , причём  $AV_i \subset V_i$  и ограничения оператора  $A$  на подпространства  $V_1$  и  $V_2$  не имеют общих собственных значений. Тогда  $BV_i \subset V_i$ .

**Теорема 7.1.2.2.** Пусть  $A_1, \dots, A_n$  — коммутирующие операторы в пространстве  $V$  над  $\mathbb{C}$ , причём пространство  $V$  неприводимо относительно этого набора операторов, т.е. если подпространство  $W \subset V$  таково, что  $A_i W \subset W$  при  $i = 1, \dots, n$ , то  $W = 0$  или  $W = V$ . Тогда  $\dim V = 1$ .

**Доказательство.** Если  $A_i = \lambda_i I$  при всех  $i$ , то утверждение очевидно. Если же  $\lambda$  — собственное значение оператора  $A_i$  и  $A_i \neq \lambda I$ , то подпространство  $\{v \in V \mid (A_i - \lambda I)v = 0\}$  отлично от 0 и  $V$  и инвариантно относительно  $A_1, \dots, A_n$  (см. теорему 7.1.2.1).  $\square$

**Теорема 7.1.2.3.** Для любого семейства коммутирующих линейных операторов в пространстве  $V$  над  $\mathbb{C}$  существует такое разложение  $V = V_1 \oplus \dots \oplus V_k$ , что все подпространства  $V_i$  инвариантны относительно данных операторов и в каждом подпространстве  $V_i$  любой оператор имеет лишь одно собственное значение.

**Доказательство.** Для оператора  $A$  в пространстве  $W$  размерности  $m$  можно рассмотреть разложение в прямую сумму инвариантных подпространств  $W_\lambda = \{w \in W \mid (A - \lambda I)^m w = 0\}$ . Разложим пространство  $V$  в прямую сумму таких подпространств для первого оператора данного семейства операторов. Полученные подпространства инвариантны относительно всех операторов семейства (см. следствие теоремы 7.1.2.1). Если какой-либо оператор семейства имеет на одном из этих подпространств различные собственные значения, то для этого оператора и для полученного подпространства рассмотрим аналогичное разложение. Такое измельчение подпространств пространства  $V$  закончится на некотором шаге.  $\square$

**Следствие.** Для любого семейства диагонализируемых попарно коммутирующих операторов существует базис, в котором все эти операторы диагональны.

**Теорема 7.1.2.4.** Для любого семейства попарно коммутирующих матриц над полем  $\mathbb{C}$  можно выбрать базис, в котором матрицы этих операторов треугольные.

**Доказательство.** Требуемое утверждение достаточно доказать для конечного числа операторов. Действительно, можно рассмотреть линейное пространство, порождённое заданными операторами, и выбрать в нём базис.

Пусть  $AB = BA$  и  $Av = \lambda v$ . Тогда  $ABv = BAv = \lambda Bv$ , поэтому  $B \operatorname{Ker}(A - \lambda I) \subset \operatorname{Ker}(A - \lambda I)$ . В пространстве  $\operatorname{Ker}(A - \lambda I)$  оператор  $B$  имеет собственный вектор с некоторым собственным значением  $\mu$ . Оператор  $C$ , коммутирующий с  $A$  и  $B$ , переводит пространство  $\operatorname{Ker}(A - \lambda I) \cap \operatorname{Ker}(B - \mu I)$  в себя и т.д. В результате находим общий собственный вектор  $v$  семейства коммутирующих операторов.

Для оператора  $A : V \rightarrow V$  с собственным вектором  $v$  можно определить оператор  $\hat{A}$  на факторпространстве  $V/\langle v \rangle$ . При этом коммутирующие операторы переходят в коммутирующие. В пространстве  $V/\langle v \rangle$  тоже можно выбрать общий собственный вектор для семейства коммутирующих операторов и т.д. В результате в пространстве  $V$  мы выберем базис  $e_1, \dots, e_n$  так, что каждый оператор семейства переводит подпространство  $\langle e_1, \dots, e_k \rangle$ ,  $k = 1, \dots, n$ , в себя. В таком базисе матрица оператора треугольная.  $\square$

### 7.1.3. Перестановочность с $A$ влечёт перестановочность с $B$

**Теорема 7.1.3.1.** Пусть матрицы  $A$  и  $B$  таковы, что любая матрица, коммутирующая с матрицей  $A$ , коммутирует также и с матрицей  $B$ . Тогда  $B = g(A)$ , где  $g$  — некоторый многочлен.

**Доказательство.** Матрицы  $A$  и  $B$  можно рассматривать как линейные операторы в некотором пространстве  $V$ . Для оператора  $A$  существует циклическое разложение  $V = V_1 \oplus \dots \oplus V_k$ , обладающее следующими свойствами:  $AV_i \subset V_i$ ; ограничение  $A_i$  оператора  $A$  на подпространство  $V_i$  является циклической клеткой; характеристический многочлен оператора  $A_i$  равен  $p_i$ , причём  $p_i$  делится на  $p_{i+1}$  и  $p_1$  — минимальный многочлен оператора  $A$  (см. §3.4.1). Пусть  $e_i$  — вектор, порождающий пространство  $V_i$ , т.е.  $V_i = \langle e_i, Ae_i, A^2e_i, \dots \rangle$ ;  $P_i$  — проектор, отображающий пространство  $V$  на  $V_i$ , т.е.  $P_i(v_1 + \dots + v_k) = v_i$  для  $v_1 \in V_1, \dots, v_k \in V_k$ . Так как  $AV_j \subset V_j$  для всех  $j$ , то  $P_iA(v_1 + \dots + v_k) = Av_i = AP_i(v_1 + \dots + v_k)$ , т.е. операторы  $P_i$  и  $A$  коммутируют, а значит, операторы  $P_i$  и  $B$  тоже коммутируют. Следовательно,  $Be_i = BP_ie_i = P_iBe_i \in V_i$ , а значит,  $Be_i = g_i(A)e_i$ , где  $g_i$  — некоторый многочлен. Пусть  $f(A)e_i$  — произвольный вектор пространства  $V_i$ . Тогда  $Bf(A)e_i = f(A)Be_i = f(A)g_i(A)e_i = g_i(A)f(A)e_i$ , т.е.  $Bv_i = g_i(A)v_i$  для любого вектора  $v_i \in V_i$ .

Докажем, что  $B = g_1(A)$ . Пусть  $n_i = p_1/p_i$ . Для  $i = 1, \dots, k$  рассмотрим операторы  $X_i : V \rightarrow V$ , переводящие вектор  $f(A)e_i$  подпространства  $V_i$  в вектор  $f(A)n_i(A)e_1$ , а на остальных подпространствах действующие тождественно. Проверим сначала, что это определение корректно, т.е. если  $f(A)e_i = 0$ , то  $f(A)n_i(A)e_1 = 0$ . В самом деле, если  $f(A)e_i = 0$ , то  $f$  делится на  $p_i$ , а так как  $p_1 = n_i p_i$ , то  $n_i f$  делится на  $p_1$ , а значит,  $f(A)n_i(A)e_1 = 0$ . Так как  $X_i A f(A)e_i = A f(A)n_i(A)e_1 = A X_i f(A)e_i$ , то операторы  $X_i$  и  $A$  коммутируют, поэтому операторы  $X_i$  и  $B$  тоже коммутируют. С другой стороны,  $X_i B e_i = X_i g_i(A)e_i = g_i(A)n_i(A)e_1 = n_i(A)g_i(A)e_1$  и  $B X_i e_i = B n_i(A)e_1 = n_i(A)B e_1 = n_i(A)g_1(A)e_1$ , поэтому  $n_i(A)[g_i(A) - g_1(A)]e_1 = 0$ . Следовательно, многочлен  $n_i(g_i - g_1)$  делится на  $p_1 = n_i p_i$ ,

т.е.  $g_i - g_1$  делится на  $p_i$ , а значит,  $Bv_i = g_i(A)v_i = g_1(A)v_i$  для любого вектора  $v_i \in V_i$ . Поэтому  $B = g_1(A)$ .  $\square$

#### 7.1.4. Теорема Шура

Пусть  $L_n$  — линейное пространство, состоящее из матриц  $\|a_{ij}\|_1^n$ , для которых  $a_{ij} = 0$ , если  $i > [n/2]$  или  $j < [n/2] + 1$ . Тогда  $\dim L_n = (n - [n/2])[n/2] = [n^2/4]$  и произведение любых двух матриц из  $L_n$  представляет собой нулевую матрицу. Действительно, каждый элемент произведения двух таких матриц имеет вид  $\sum_k a_{ik}a'_{kj}$ ; если  $k > [n/2]$ , то  $a'_{kj} = 0$ , а если  $k \leq [n/2]$ , то  $a_{ik} = 0$ . Добавив к  $L_n$  единичную матрицу  $I_n$ , получим линейное пространство  $\tilde{L}_n$  размерности  $[n^2/4] + 1$ ; любые две матрицы из пространства  $\tilde{L}_n$  коммутируют. Шур [Sc4] доказал, что размерность любого линейного пространства попарно коммутирующих матриц порядка  $n$  не больше размерности пространства  $\tilde{L}_n$ .

**Теорема 7.1.4.1 (Шур [Sc4]).** Пусть  $V_n$  — линейное пространство над полем  $\mathbb{C}$ , состоящее из попарно коммутирующих матриц порядка  $n$ . Тогда  $\dim V_n \leq [n^2/4] + 1$ .

**Доказательство.** ([M1]) Применим индукцию по  $n$ . При  $n = 1$  утверждение очевидно. Предположим, что утверждение верно для матриц порядка  $n - 1$ , но существует линейное пространство  $V_n$  (состоящее из попарно коммутирующих матриц порядка  $n$ ), для которого  $\dim V_n \geq [n^2/4] + 2$ . Согласно теореме 7.1.2.4 существует такая матрица  $P$ , что все матрицы  $P^{-1}AP$ ,  $A \in V_n$ , являются верхними треугольными. Поэтому будем считать, что пространство  $V_n$  состоит из верхних треугольных матриц. Выберем в пространстве  $V_n$  линейно независимые матрицы  $A_1, A_2, \dots, A_{[n^2/4]+2}$ . Пусть  $\hat{A}_i$  — матрица порядка  $n - 1$ , полученная из матрицы  $A_i$  вычёркиванием первой строки и первого столбца. Для верхних треугольных матриц выполняется равенство  $\widehat{A_i A_j} = \hat{A}_i \hat{A}_j$ , поэтому матрицы  $\hat{A}_i$  попарно коммутируют. Согласно предположению индукции матрицы  $\hat{A}_1, \hat{A}_2, \dots$  порождают линейное пространство  $\hat{V}_n$  размерности  $k \leq [(n - 1)^2/4] + 1$ . Можно считать, что матрицы  $\hat{A}_1, \dots, \hat{A}_k$  образуют базис пространства  $\hat{V}_n$ .

Для  $i \geq k + 1$  представим матрицу  $\hat{A}_i$  в виде  $\hat{A}_i = \lambda_{i1}\hat{A}_1 + \dots + \lambda_{ik}\hat{A}_k$  и рассмотрим матрицу  $B_i = A_i - \lambda_{i1}A_1 - \dots - \lambda_{ik}A_k$ . Пусть  $b_i$  — первая строка матрицы  $B_i$  (остальные строки матрицы  $B_i$  нулевые). Матрицы  $B_i$ ,  $i = k + 1, \dots, [n^2/4] + 2$  линейно независимы, поэтому векторы  $b_i$ ,  $i = k + 1, \dots, [n^2/4] + 2$  тоже линейно независимы.



Те же самые рассуждения можно повторить, выбрасывая не первую строку и первый столбец, а последнюю строку и последний столбец. В результате получим линейно независимые матрицы  $B'_i$ ,  $i = k + 1, \dots, [n^2/4] + 2$ , у которых последние столбцы равны  $b'_i$ , а все остальные элементы нулевые.

По построению  $B_i, B'_j \in V_n$ , поэтому  $B_i B'_j = B'_j B_i$ ; в частности,  $(b_i, b'_j) = 0$ . Следовательно,

$$\dim \langle b_{k+1}, \dots, b_{[n^2/4]+1} \rangle + \dim \langle b'_{k+1}, \dots, b'_{[n^2/4]+1} \rangle \leq n,$$

т.е.  $2([n^2/4] - k + 2) \leq n$ . С другой стороны,  $2([n^2/4] - k + 2) \geq 2([n^2/4] - [(n-1)^2/4] + 1) \geq n + 1$  (чтобы доказать последнее неравенство, нужно рассмотреть два случая:  $n$  чётно и  $n$  нечётно). Приходим к противоречию.  $\square$

### Задачи

**1.1.** Пусть  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ , где числа  $\lambda_i$  попарно различны, и матрица  $X$  коммутирует с матрицей  $A$ .

а) Докажите, что матрица  $X$  диагональна.

б) Пусть, кроме того, числа  $\lambda_i$  ненулевые и матрица  $X$  коммутирует с матрицей  $NA$ , где  $N = \|\delta_{i+1,j}\|_1^n$ . Докажите, что  $X = \lambda I$ .

**1.2.** Докажите, что если матрица  $X$  коммутирует со всеми матрицами, то  $X = \lambda I$ .

**1.3.** Найдите все матрицы, коммутирующие с матрицей  $A = \|a_{ij}\|_1^n$ , где  $a_{ij} = 1$  при всех  $i, j$ .

**1.4.** Пусть  $P_\sigma$  — матрица, соответствующая перестановке  $\sigma$ . Докажите, что если  $AP_\sigma = P_\sigma A$  для всех  $\sigma$ , то  $A = \lambda I + \mu J$ , где  $J$  — матрица, все элементы которой равны 1.

**1.5.** Пусть  $A = P^{-1}JP$ , где  $J$  — жорданова матрица. Докажите, что  $A = P_1^{-1}JP_1$  тогда и только тогда, когда  $P_1 = XP$ , где  $X$  — невырожденная матрица, перестановочная с  $J$ .

**1.6.** Докажите, что для любой комплексной матрицы  $A$  существует такая матрица  $B$ , что  $AB = BA$  и характеристический многочлен матрицы  $B$  совпадает с минимальным многочленом.

**1.7.** а) Пусть  $A$  и  $B$  — коммутирующие нильпотентные матрицы. Докажите, что матрица  $A + B$  нильпотентна.

б) Пусть  $A$  и  $B$  — коммутирующие диагонализуемые матрицы. Докажите, что матрица  $A + B$  диагонализуема.

**1.8.** В пространстве размерности  $n$  даны (различные) попарно коммутирующие инволюции  $A_1, \dots, A_m$ . Докажите, что  $m \leq 2^n$ .

**1.9.** Диагонализируемые операторы  $A_1, \dots, A_n$  попарно коммутируют. Докажите, что все эти операторы полиномиально выражаются через некоторый диагонализируемый оператор.

**1.10.** Докажите, что если все матрицы, коммутирующие с матрицей  $A$ , являются многочленами от матрицы  $B$ , то они являются многочленами от матрицы  $A$ .

**1.11.** Матрицы  $A_{ij}$  попарно коммутируют. Докажите, что определитель матрицы  $A = \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \dots & \dots & \dots \\ A_{k1} & \dots & A_{kk} \end{pmatrix}$  равен

$$\det \left( \sum_{\sigma} (-1)^{\sigma} A_{1\sigma(1)} A_{2\sigma(2)} \dots A_{k\sigma(k)} \right).$$

**1.12.** [Та]Предположим, что характеристический многочлен матрицы  $A$  совпадает с минимальным многочленом. Докажите, что любая матрица  $X$ , для которой  $XA = A^T X$ , является симметрической.

## 7.2. Свойства коммутаторов

### 7.2.1. Определение и простейшие свойства

Пусть  $A$  и  $B$  — квадратные матрицы одного порядка. Матрицу  $AB - BA$  называют *коммутатором* матриц  $A$  и  $B$  и обозначают  $[A, B]$ . Равенство  $[A, B] = 0$  означает, что матрицы  $A$  и  $B$  коммутируют. Легко проверить, что  $\text{tr}[A, B] = 0$  для любых матриц  $A$  и  $B$  (см. §3.1.1).

Несложные вычисления показывают, что

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \quad (\text{тождество Якоби}).$$

Отображение  $\text{ad}_A: M_{n,n} \rightarrow M_{n,n}$ , определённое формулой  $\text{ad}_A(X) = [A, X]$ , является линейным оператором в пространстве матриц. Отображение, сопоставляющее каждой матрице  $A$  оператор  $\text{ad}_A$ , называют *присоединённым представлением* алгебры  $M_{n,n}$ . Присоединённое представление имеет важные приложения в теории алгебр Ли.

Легко проверить следующие свойства отображения  $\text{ad}_A$ :

1)  $\text{ad}_{[A, B]} = \text{ad}_A \text{ad}_B - \text{ad}_B \text{ad}_A$  (это равенство эквивалентно тождеству Якоби);

2) оператор  $D = \text{ad}_A$  является дифференцированием алгебры матриц, т.е.  $D(XY) = XD(Y) + (DX)Y$ ;

$$3) D^n(XY) = \sum_{k=0}^n \binom{n}{k} (D^k X)(D^{n-k} Y);$$

$$4) D(X^n) = \sum_{k=0}^{n-1} X^k (DX) X^{n-1-k}.$$

### 7.2.2. Представление матриц в виде коммутаторов

Если  $A = [X, Y]$ , то  $\text{tr } A = 0$ . Оказывается, верно и обратное: если  $\text{tr } A = 0$ , то существуют такие матрицы  $X$  и  $Y$ , что  $A = [X, Y]$ . Более того, на матрицы  $X$  и  $Y$  можно накладывать различные ограничения.

**Теорема 7.2.2.1.** Пусть  $\text{tr } A = 0$ . Тогда существуют такие матрицы  $X$  и  $Y$ , что  $X$  — эрмитова матрица,  $\text{tr } Y = 0$  и  $A = [X, Y]$ .

**Доказательство.** [Fe] Существует такая унитарная матрица  $U$ , что все диагональные элементы матрицы  $UAU^* = B = \|b_{ij}\|_1^n$  нулевые (см. §3.5.2). Рассмотрим матрицу  $D = \text{diag}(d_1, \dots, d_n)$ , где  $d_1, \dots, d_n$  — произвольные попарно различные действительные числа. Пусть  $Y_1 = \|y_{ij}\|_1^n$ , где  $y_{ii} = 0$  и  $y_{ij} = b_{ij}/(d_i - d_j)$  при  $i \neq j$ . Тогда  $DY_1 - Y_1D = \|(d_i - d_j)y_{ij}\|_1^n = \|b_{ij}\|_1^n = UAU^*$ . Следовательно,  $A = U^*DY_1U - U^*Y_1DU = XY - YX$ , где  $X = U^*DU$  и  $Y = U^*Y_1U$ . Ясно, что  $X$  — эрмитова матрица и  $\text{tr } Y = 0$ .  $\square$

**Замечание.** Если  $A$  — вещественная матрица, то матрицы  $X$  и  $Y$  можно выбрать вещественными.

**Теорема 7.2.2.2.** Пусть  $\text{tr } A = 0$  и  $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n$  — заданные комплексные числа, причём  $\lambda_i \neq \lambda_j$  при  $i \neq j$ . Тогда существуют комплексные матрицы  $X$  и  $Y$  с собственными значениями  $\lambda_1, \dots, \lambda_n$  и  $\mu_1, \dots, \mu_n$  соответственно, для которых  $A = [X, Y]$ .

**Доказательство.** [Gi] Существует такая матрица  $P$ , что все диагональные элементы матрицы  $PAP^{-1} = B = \|b_{ij}\|_1^n$  нулевые (см. §3.5.1). Пусть  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  и  $c_{ij} = b_{ij}/(\lambda_i - \lambda_j)$  при  $i \neq j$ . Диагональные элементы  $c_{ii}$  матрицы  $C$  можно подобрать так, чтобы она имела собственные значения  $\mu_1, \dots, \mu_n$  (теорема 8.10.2.2). Тогда  $DC - CD = \|(\lambda_i - \lambda_j)c_{ij}\|_1^n = B$ . Остаётся положить  $X = P^{-1}DP$  и  $Y = P^{-1}CP$ .  $\square$

### 7.2.3. Равенство $\text{ad}_A^s X = 0$ влечёт равенство $\text{ad}_X^s B = 0$

**Теорема 7.2.3.1.** Пусть матрицы  $A$  и  $B$  таковы, что для некоторого  $s > 0$  равенство  $\text{ad}_A^s X = 0$  влечёт равенство  $\text{ad}_X^s B = 0$ . Тогда матрица  $B$  полиномиально выражается через  $A$ .

**Доказательство.** [Sm] Случай  $s = 1$  разобран в §7.1.3; поэтому в дальнейшем будем считать, что  $s \geq 2$ . Отметим, что при  $s \geq 2$  из равенства  $\text{ad}_A^s X = 0$  не обязательно следует равенство  $\text{ad}_X^s A = 0$ .

Можно считать, что  $A = \text{diag}(J_1, \dots, J_t)$ , где  $J_i$  — жорданова клетка. Пусть  $X = \text{diag}(1, \dots, n)$ . Легко проверить, что  $\text{ad}_A^2 X = 0$  (см. задачу 2.1), а значит,  $\text{ad}_A^s X = 0$  и  $\text{ad}_X^s B = 0$ . Матрица  $X$  диагонализирована, поэтому  $\text{ad}_X B = 0$  (см. задачу 2.8). Следовательно,  $B$  — диагональная матрица (см. задачу 1.1, а)). В соответствии с блочной записью  $A = \text{diag}(J_1, \dots, J_t)$  запишем матрицы  $B$  и  $X$  в виде  $B = \text{diag}(B_1, \dots, B_t)$  и  $X = \text{diag}(X_1, \dots, X_t)$ . Пусть  $Y = \text{diag}((J_1 - \lambda_1 I)X_1, \dots, (J_t - \lambda_t I)X_t)$ , где  $\lambda_i$  — собственное значение жордановой клетки  $J_i$ . Тогда  $\text{ad}_A^2 Y = 0$  (см. задачу 2.1). Следовательно,  $\text{ad}_A^2(X + Y) = 0$ , а значит,  $\text{ad}_{X+Y}^2 B = 0$ . Матрица  $X + Y$  диагонализирована, так как её собственные значения равны  $1, \dots, n$ . Поэтому  $\text{ad}_{X+Y} B = 0$ , а значит,  $\text{ad}_Y B = 0$ .

Из равенств  $[X, B] = 0$  и  $[Y, B] = 0$  следует, что  $B_i = b_i I$  (см. задачу 1.1). Докажем, что если собственные значения матриц  $J_i$  и  $J_{i+1}$  рав-

ны, то  $b_i = b_{i+1}$ . Рассмотрим матрицу  $U = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 \end{pmatrix}$ , порядок

которой равен сумме порядков матриц  $J_i$  и  $J_{i+1}$ , и в соответствии с блочной записью  $A = \text{diag}(J_1, \dots, J_t)$  введём матрицу  $Z = \text{diag}(0, U, 0)$ . Легко проверить, что  $ZA = AZ = \lambda Z$ , где  $\lambda$  — собственное значение матриц  $J_i$  и  $J_{i+1}$ . Следовательно,  $\text{ad}_A(X + Z) = \text{ad}_A Z = 0$ ,  $\text{ad}_A^s(X + Z) = 0$  и  $\text{ad}_{X+Z}^s B = 0$ . А так как собственные значения матрицы  $X + Z$  равны  $1, \dots, n$ , то  $\text{ad}_{X+Z} B = 0$ . Поэтому  $[Z, B] = [X + Z, B] = 0$ , т.е.  $b_i = b_{i+1}$ .

Можно считать, что  $A = \text{diag}(M_1, \dots, M_q)$ , где  $M_i$  — объединение жордановых клеток с равными собственными значениями. Тогда  $B = \text{diag}(B'_1, \dots, B'_q)$ , где  $B'_i = b'_i I$ . Согласно §7.1.1 из равенства  $[W, A] = 0$  следует, что  $W = \text{diag}(W_1, \dots, W_q)$ , а значит,  $[W, B] = 0$ . Таким образом, случай  $s \geq 2$  сводится к случаю  $s = 1$ .  $\square$

### 7.2.4. Одновременная триангулируемость

Матрицы  $A_1, \dots, A_m$  называют *одновременно триангулируемыми*, если существует такая унитарная матрица  $P$ , что все матрицы  $P^* A_i P$  верхние треугольные.

**Теорема 7.2.4.1.** *Матрицы  $A_1, \dots, A_m$  одновременно триангулируемы тогда и только тогда, когда для любого многочлена  $p(x_1, \dots, x_m)$  от некоммутирующих переменных матрица  $p(A_1, \dots, A_m)[A_i, A_j]$  нильпотентна.*

**Доказательство.** [Dr2] Если матрицы  $A_1, \dots, A_m$  одновременно триангулируемы, то  $P^* p(A_1, \dots, A_m) P$  и  $P^* [A_i, A_j] P$  — верхние треугольные матрицы, причём на диагонали последней матрицы стоят нули. Поэтому матрица  $p(A_1, \dots, A_m)[A_i, A_j]$  нильпотентна.

Предположим теперь, что все матрицы вида  $p(A_1, \dots, A_m)[A_i, A_j]$  нильпотентны, и докажем, что тогда матрицы  $A_1, \dots, A_m$  одновременно триангулируемы. Докажем сначала, что для любого ненулевого вектора  $u$  существует такой многочлен  $h(x_1, \dots, x_m)$ , что  $h(A_1, \dots, A_m)u$  — ненулевой общий собственный вектор матриц  $A_1, \dots, A_m$ . Доказательство проведём индукцией по  $m$ . При  $m = 1$  существует такое число  $k$ , что векторы  $u, A_1 u, \dots, A_1^{k-1} u$  линейно независимы и  $A_1^k u = a_{k-1} A_1^{k-1} u + \dots + a_0 u$ .

Пусть  $g(x) = x^k - a_{k-1}x^{k-1} - \dots - a_0$  и  $g_0(x) = \frac{g(x)}{x - x_0}$ , где  $x_0$  — корень многочлена  $g$ . Тогда  $g_0(A_1)u \neq 0$  и  $(A_1 - x_0 I)g_0(A_1)u = g(A_1)u = 0$ , т.е.  $g_0(A_1)u$  — собственный вектор матрицы  $A_1$ . Предположим, что утверждение верно для любых  $m-1$  матриц  $A_1, \dots, A_{m-1}$ . Для данного ненулевого вектора  $u$  некоторый вектор  $v_1 = h(A_1, \dots, A_{m-1})u$  является общим собственным вектором матриц  $A_1, \dots, A_{m-1}$ . Возможны два случая.

1.  $[A_i, A_m]f(A_m)v_1 = 0$  для любого многочлена  $f$ . Тогда  $[A_i, A_m^k]v_1 = \sum_{s=0}^{k-1} A_m^s ([A_i, A_m]A_m^{k-s-1}v_1) = 0$ , т.е.  $A_i g(A_m)v_1 = g(A_m)A_i v_1$  для любого многочлена  $g$ . Для матрицы  $A_m$  существует такой многочлен  $g_1$ , что  $g_1(A_m)v_1$  — собственный вектор этой матрицы. А так как  $A_i g_1(A_m)v_1 = g_1(A_m)A_i v_1$  и  $v_1$  — собственный вектор матриц  $A_1, \dots, A_{m-1}$ , то  $g_1(A_m)v_1 = g_1(A_m)h(A_1, \dots, A_{m-1})u$  — собственный вектор матриц  $A_1, \dots, A_m$ .

2.  $[A_i, A_m]f(A_m)v_1 \neq 0$  для некоторого многочлена  $f_1$  и для некоторого  $i$ . Вектор  $C_1 f_1(A_m)v_1$ , где  $C_1 = [A_i, A_m]$ , ненулевой, поэтому у матриц  $A_1, \dots, A_{m-1}$  есть общий собственный вектор  $v_2 = g_1(A_1, \dots, A_{m-1})C_1 f_1(A_m)v_1$ . К вектору  $v_2$  можно применить такие же рассуждения и т.д. В результате получим последовательность  $v_1, v_2, v_3, \dots$ , где  $v_k$  — собственный вектор матриц  $A_1, \dots, A_{m-1}$  и  $v_{k+1} =$

$g_k(A_1, \dots, A_{m-1})C_k f_k(A_m)v_k$ , причём  $C_k = [A_s, A_m]$  для некоторого  $s$ . Эта последовательность обрывается на векторе  $v_p$ , если  $[A_i, A_m]f(A_m)v_p = 0$  для всех  $i$  и для всех многочленов  $f$ . Для матрицы  $A_m$  существует такой многочлен  $g_p(x)$ , что  $g_p(A_m)v_p$  — собственный вектор матрицы  $A_m$ . Как и в случае 1, получаем, что этот вектор является собственным вектором матриц  $A_1, \dots, A_m$  и  $g_p(A_m)v_p = g_p(A_m)g(A_1, \dots, A_m)h(A_1, \dots, A_{m-1})u$ .

Остаётся доказать, что последовательность  $v_1, v_2, \dots$  обрывается. Предположим, что эта последовательность не обрывается. Тогда существуют числа  $\lambda_1, \dots, \lambda_{n+1}$ , не все равные нулю, для которых  $\lambda_1 v_1 + \dots + \lambda_{n+1} v_{n+1} = 0$ , а значит, существует номер  $j$ , для которого  $\lambda_j \neq 0$  и  $-\lambda_j v_j = \lambda_{j+1} v_{j+1} + \dots + \lambda_{n+1} v_{n+1}$ . Ясно, что  $v_{j+1} = g_j(A_1, \dots, A_{m-1})C_j f_j(A_m)v_j$ ,  $v_{j+2} = u_{j+2}(A_1, \dots, A_m)C_j f_j(A_m)v_j$  и т.д. Поэтому  $-\lambda_j v_j = u(A_1, \dots, A_m)C_j f_j(A_m)v_j$ , а значит,  $f_j(A_m)u(A_1, \dots, A_m)C_j f_j(A_m)v_j = -\lambda_j f_j(A_m)v_j$ . Следовательно, ненулевой вектор  $f_j(A_m)v_j$  является собственным вектором оператора  $f_j(A_m)u(A_1, \dots, A_m)C_j$ , причём он соответствует ненулевому собственному значению  $-\lambda_j$ . Но по условию этот оператор нильпотентен, а значит, у него нет ненулевых собственных значений. Получено противоречие.

Перейдём теперь непосредственно к доказательству теоремы. Доказательство проведём индукцией по  $n$ . При  $n = 1$  утверждение очевидно. Как уже было доказано, операторы  $A_1, \dots, A_m$  имеют общий собственный вектор  $u$ , соответствующий некоторым собственным значениям  $\alpha_1, \dots, \alpha_m$ . Можно считать, что  $\|u\| = 1$ , т.е.  $u^*u = 1$ . Существует унитарная матрица  $Q$ , первым столбцом которой является вектор  $u$ . Ясно, что  $Q^* A_i Q = Q^*(\alpha_i u \dots) = \begin{pmatrix} \alpha_i & * \\ 0 & A_i' \end{pmatrix}$ , причём матрицы  $A_1', \dots, A_m'$  порядка  $n - 1$  удовлетворяют условию теоремы. По предположению индукции существует унитарная матрица  $P_1$  порядка  $n - 1$ , для которой матрицы  $P_1^* A_i' P_1$  верхние треугольные. Тогда  $P = Q \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix}$  — искомая матрица.  $\square$

**Следствие.** Множество наборов  $(A_1, \dots, A_m)$  одновременно триангулируемых матриц замкнуто.

**Доказательство.** Для фиксированных  $i$  и  $j$  и многочлена  $p$  множество наборов, для которых матрица  $X = p(A_1, \dots, A_m)[A_i, A_j]$  нильпотентна, задаётся уравнением  $X^n = 0$ ; это уравнение приводит к системе алгебраических уравнений для элементов матриц  $A_1, \dots, A_m$ .  $\square$

**Замечание.** Из доказательства теоремы 7.2.4.1 следует, что матрицы  $A_1,$

...,  $A_m$  одновременно триангулируемы тогда и только тогда, когда существует невырожденная матрица  $P$ , для которой все матрицы  $P^{-1}A_iP$  верхние треугольные.

### 7.2.5. Триангулируемость операторов с рангом коммутатора 1

**Теорема 7.2.5.1.** Если  $A$  и  $B$  — операторы в пространстве  $V$  над  $\mathbb{C}$ , причём  $\text{rk}[A, B] \leq 1$ , то  $A$  и  $B$  одновременно триангулируемы.

**Доказательство.** Достаточно доказать, что операторы  $A$  и  $B$  имеют общий собственный вектор  $v \in V$ . В самом деле, тогда операторы  $A$  и  $B$  индуцируют операторы  $A_1$  и  $B_1$  в пространстве  $V_1 = V/\langle v \rangle$ , причём  $\text{rk}[A_1, B_1] \leq 1$ . Следовательно, операторы  $A_1$  и  $B_1$  имеют общий собственный вектор в пространстве  $V_1$  и т.д. Кроме того, можно считать, что  $\text{Ker } A \neq 0$  (иначе вместо оператора  $A$  можно взять оператор  $A - \lambda I$ ).

Доказательство проведём индукцией по  $n = \dim V$ . Если  $n = 1$ , то утверждение очевидно. Пусть  $C = [A, B]$ . При доказательстве шага индукции рассмотрим два случая.

1.  $\text{Ker } A \subset \text{Ker } C$ . В этом случае  $B(\text{Ker } A) \subset \text{Ker } A$ , так как если  $Ax = 0$ , то  $Cx = 0$  и  $ABx = BAx + Cx = 0$ . Следовательно, можно рассмотреть ограничение оператора  $B$  на  $\text{Ker } A \neq 0$  и выбрать в  $\text{Ker } A$  собственный вектор  $v$  оператора  $B$ ; вектор  $v$  при этом будет также и собственным вектором оператора  $A$ .

2.  $\text{Ker } A \not\subset \text{Ker } C$ , т.е.  $Ax = 0$  и  $Cx \neq 0$  для некоторого вектора  $x$ . Так как  $\text{rk } C = 1$ , то  $\text{Im } C = \langle y \rangle$ , где  $y = Cx$ . Кроме того,  $y = Cx = ABx - BAx = ABx \in \text{Im } A$ . Следовательно,  $B(\text{Im } A) \subset \text{Im } A$ . В самом деле,  $BAz = ABz - Cz$ , где  $ABz \in \text{Im } A$  и  $Cz \in \text{Im } C \subset \text{Im } A$ . По предположению  $\text{Ker } A \neq 0$ , поэтому  $\dim \text{Im } A < n$ . Пусть  $A'$  и  $B'$  — ограничения операторов  $A$  и  $B$  на  $\text{Im } A$ . Тогда  $\text{rk}[A', B'] \leq 1$ , а значит, согласно предположению индукции у операторов  $A'$  и  $B'$  есть общий собственный вектор.  $\square$

### Задачи

2.1. Пусть  $J = N + \lambda I$  — жорданова клетка порядка  $n$ ,  $A = \text{diag}(1, 2, \dots, n)$  и  $B = NA$ . Докажите, что  $\text{ad}_J^2 A = \text{ad}_J^2 B = 0$ .

2.2. Докажите, что  $\text{tr}([A, B]C) = \text{tr}(A[B, C])$ .

2.3. Докажите, что если  $C = [A_1, B_1] + \dots + [A_n, B_n]$  и матрица  $C$  коммутирует с матрицами  $A_1, \dots, A_n$ , то матрица  $C$  нильпотентна.

2.4. Докажите, что  $\text{ad}_A^n(B) = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} A^i B A^{n-i}$ .

2.5. Докажите, что если оператор  $A$  нильпотентен, то оператор  $\text{ad}_A$  тоже нильпотентен.

2.6. [K11] Докажите, что если  $\text{ad}_A^2(B) = 0$ , то  $\text{ad}_A^n(B^n) = n!(\text{ad}_A(B))^n$ .

2.7. Докажите, что если  $[A, [A, B]] = 0$  и  $m \geq n > 0$ , то  $n[A^m, B] = m[A^n, B]A^{m-n}$ .

2.8. Докажите, что если  $A$  — диагонализируемая матрица и  $\text{ad}_A^n X = 0$ , то  $\text{ad}_A X = 0$ .

2.9. а) Докажите, что если  $\text{tr}(AXY) = \text{tr}(AYX)$  для любых матриц  $X$  и  $Y$ , то  $A = \lambda I$ .

б) Пусть  $f$  — линейная функция на пространстве матриц порядка  $n$ . Докажите, что если  $f(XY) = f(YX)$  для любых матриц  $X$  и  $Y$ , то  $f(X) = \lambda \text{tr } X$ .

## 7.3. Теория реплик

### 7.3.1. Определение и основные свойства

Любому линейному оператору  $A : V \rightarrow V$  можно сопоставить линейный оператор  $A_{p,q} : T_q^p(V) \rightarrow T_q^p(V)$ , заданный следующим образом:

$$\begin{aligned} A_{p,q}(v_1 \otimes \cdots \otimes v_p \otimes f_1 \otimes \cdots \otimes f_q) = \\ = A(v_1) \otimes v_2 \otimes \cdots \otimes v_p \otimes f_1 \otimes \cdots \otimes f_q + v_1 \otimes A(v_2) \otimes \cdots \otimes v_p \otimes f_1 \otimes \cdots \otimes f_q + \cdots - \\ - v_1 \otimes \cdots \otimes v_p \otimes A^*(f_1) \otimes \cdots \otimes f_q - \cdots - v_1 \otimes \cdots \otimes v_p \otimes f_1 \otimes \cdots \otimes A^*(f_q). \end{aligned}$$

Например, если мы отождествим  $T_1^1(V)$  с  $\text{Hom}(V, V)$ , то для  $X \in \text{Hom}(V, V)$  получаем  $A_{1,1}(X) = AX - XA = [A, X]$ .

Линейный оператор  $B : V \rightarrow V$  называют *репликой*  $A$ , если  $\text{Ker } A_{p,q} \subset \text{Ker } B_{p,q}$  для всех  $p, q$ . Это понятие ввёл Шевалле [Ch1]. То, что  $B$  — реплика  $A$ , мы будем обозначать  $A \twoheadrightarrow B$ .

**Теорема 7.3.1.1.** Если  $A \twoheadrightarrow B$ , то  $B$  можно представить в виде многочлена от  $A$  без свободного члена.

**Доказательство.** Условие  $\text{Ker } A_{1,1} \subset \text{Ker } B_{1,1}$  означает, что любой оператор  $X$ , коммутирующий с  $A$ , коммутирует также и с  $B$ . Поэтому согласно теореме 1.3  $B = g(A)$ , где  $g$  — некоторый многочлен. Остаётся доказать, что  $g$  можно выбрать так, что  $g(0) = 0$ . Рассмотрим для этого два случая.



**Случай 1.**  $\det A \neq 0$ .

В этом случае минимальный многочлен матрицы  $A$  имеет ненулевой свободный член, поэтому для любого  $\lambda$  матрицу  $\lambda I$  можно представить как многочлен от  $A$  без свободного члена. Этот многочлен можно подставить в  $g(A)$  вместо свободного члена.

**Случай 2.**  $\det A = 0$ .

В этом случае мы сразу получаем, что  $g(0) = 0$ . Действительно, существует ненулевой вектор  $v \in \text{Ker } A_{1,0} = \text{Ker } A$ . Тогда  $v \in \text{Ker } B_{1,0} = \text{Ker } B$ , поэтому  $g(A)v = 0$ . Но  $Av = 0$ , поэтому  $g(0)v = 0$ . А так как  $v \neq 0$ , получаем  $g(0) = 0$ .  $\square$

**Следствие.**  $A \rightarrow B$  тогда и только тогда, когда  $B_{p,q}$  можно представить в виде многочлена от  $A_{p,q}$  без свободного члена для всех  $p, q$ .

**Доказательство.** При отождествлении  $T_{q'}^{p'}(T_q^p(V))$  с  $T_{pq'+qp'}^{pp'+qq'}(V)$  (задача 1.9) получаем  $(A_{p,q})_{p',q'} = A_{pp'+qq',pq'+qp'}$ . Поэтому если  $A \rightarrow B$ , то  $A_{p,q} \rightarrow B_{p,q}$ . Следовательно,  $B_{p,q}$  можно представить в виде многочлена от  $A_{p,q}$  без свободного члена. Обратное утверждение очевидно.  $\square$

### 7.3.2. Полупростая и нильпотентная составляющая

**Теорема 7.3.2.1.** Если оператор  $A$  полупростой (нильпотентный), то оператор  $A_{p,q}$  полупростой (нильпотентный).

**Доказательство.** Если оператор  $A$  полупростой, то оператор  $A^*$  тоже полупростой. Ясно также, что тензорное произведение полупростого оператора и тождественного — полупростой оператор. Таким образом,  $A_{p,q}$  — сумма попарно коммутирующих полупростых операторов. Следовательно,  $A_{p,q}$  — полупростой оператор (задача 1.7). Для нильпотентного оператора доказательство аналогично.  $\square$

Оператору  $A$  можно сопоставить его полупростую составляющую  $A_s$  и нильпотентную составляющую  $A_n$  (см. §3.2.5).

**Теорема 7.3.2.2.** Для всех  $p, q$

$$(A_{p,q})_s = (A_s)_{p,q} \quad \text{и} \quad (A_{p,q})_n = (A_n)_{p,q}.$$

**Доказательство.** Согласно теореме 7.3.2.1 оператор  $(A_s)_{p,q}$  полупростой, а оператор  $(A_n)_{p,q}$  нильпотентный. Кроме того, тождество  $[X, Y]_{p,q} = [X_{p,q}, Y_{p,q}]$  показывает, что  $(A_s)_{p,q}$  и  $(A_n)_{p,q}$  коммутируют, а тождество  $(X+Y)_{p,q} = X_{p,q} + Y_{p,q}$  показывает, что  $A_{p,q} = (A_s)_{p,q} + (A_n)_{p,q}$ . Полупростая и нильпотентная составляющие оператора определены однозначно, поэтому получаем требуемое.  $\square$

**Следствие.**  $A \twoheadrightarrow A_s$  и  $A \twoheadrightarrow A_n$ .

**Доказательство.** Согласно следствию теоремы 3.2.1.1  $\text{Ker } A_{p,q} \subset \text{Ker}(A_{p,q})_s = \text{Ker}(A_s)_{p,q}$ . Поэтому  $A \twoheadrightarrow A_s$ . Для нильпотентной составляющей доказательство аналогично.  $\square$

### 7.3.3. Комплексификация

**Теорема 7.3.3.1.** Пусть  $A$  и  $B$  — вещественные операторы, причём  $A \twoheadrightarrow B$ . Тогда  $A^{\mathbb{C}} \twoheadrightarrow B^{\mathbb{C}}$ . Кроме того, любая реплика оператора  $A^{\mathbb{C}}$  представляется в виде суммы реплик оператора  $A$  с комплексными коэффициентами.

**Доказательство.** Согласно следствию теоремы 7.3.1.1  $B_{p,q} = g(A_{p,q})$ , где  $g$  — многочлен с вещественными коэффициентами без свободного члена. Ясно, что  $B_{p,q}^{\mathbb{C}} = g(A_{p,q}^{\mathbb{C}})$ , поэтому следствие теоремы 7.3.1.1 показывает, что  $A^{\mathbb{C}} \twoheadrightarrow B^{\mathbb{C}}$ .

Докажем теперь, что любая реплика оператора  $A^{\mathbb{C}}$  представляется в виде суммы реплик оператора  $A$  с комплексными коэффициентами. Пусть  $\tilde{B}$  — реплика оператора  $A^{\mathbb{C}}$ . Запишем оператор  $\tilde{B}$  в виде  $\tilde{B} = C^{\mathbb{C}} + iD^{\mathbb{C}}$ , где  $C$  и  $D$  — вещественные операторы. Если  $x \in \text{Ker } A_{p,q}$ , то  $x \in \text{Ker } A_{p,q}^{\mathbb{C}} \subset \text{Ker } \tilde{B}_{p,q}$ , поэтому  $x \in \text{Ker } C_{p,q}$  и  $x \in \text{Ker } D_{p,q}$ . Следовательно,  $C$  и  $D$  — реплики  $A$ .  $\square$

**Теорема 7.3.3.2.** Пусть  $A$  — полупростой оператор,  $\lambda_1, \dots, \lambda_n$  — его собственные значения,  $g$  — многочлен (с вещественными или комплексными коэффициентами). Оператор  $g(A)$  является репликой  $A$  тогда и только тогда, когда из любого соотношения  $\sum n_i \lambda_i = 0$  с целыми  $n_i$  следует соотношение  $\sum n_i g(\lambda_i) = 0$ .

**Доказательство.** Воспользовавшись теоремой 7.3.3.1, можно рассмотреть комплексификацию оператора  $A$  и считать, что он диагональный:  $Ae_i = \lambda_i e_i$ . Тогда матрица оператора  $A_{p,q}$  относительно соответствующего базиса тоже диагональная, причём её диагональные элементы имеют вид  $\sum_{k=1}^p \lambda_{i_k} - \sum_{s=1}^q \lambda_{j_s}$ . Для этой диагональной матрицы собственные векторы с нулевым собственным значением образуют подпространство,

порождённое базисными векторами, соответствующими соотношениям  $\sum_{k=1}^p \lambda_{i_k} - \sum_{s=1}^q \lambda_{j_s} = 0$ . Матрица оператора  $g(A)_{p,q}$  тоже диагональная; её диагональные элементы имеют вид  $\sum_{k=1}^p g(\lambda_{i_k}) - \sum_{s=1}^q g(\lambda_{j_s})$ . Поэтому ядро оператора  $A_{p,q}$  содержится в ядре оператора  $g(A)_{p,q}$  тогда и только тогда, когда из соотношения  $\sum_{k=1}^p \lambda_{i_k} - \sum_{s=1}^q \lambda_{j_s} = 0$  вытекает соотношение  $\sum_{k=1}^p g(\lambda_{i_k}) - \sum_{s=1}^q g(\lambda_{j_s}) = 0$ .  $\square$

### 7.3.4. Критерий нильпотентности

Одним из наиболее важных приложений теории реплик является следующий критерий нильпотентности.

**Теорема 7.3.4.1.** *Оператор  $A$  нильпотентен тогда и только тогда, когда  $\text{tr}(AB) = 0$  для любой реплики  $B$  оператора  $A$ .*

**Доказательство.** Предположим сначала, что оператор  $A$  нильпотентен. Пусть  $B$  — его реплика. Согласно теореме 7.3.1.1  $B$  полиномиально выражается через  $A$ , а значит,  $B$  коммутирует с  $A$ . Но тогда оператор  $AB$  нильпотентен, а значит,  $\text{tr}(AB) = 0$ .

Предположим теперь, что  $\text{tr}(AB) = 0$  для любой реплики  $B$  оператора  $A$ . Пусть  $A = A_s + A_n$  — разложение оператора  $A$  на полупростую и нильпотентную составляющие. Возьмём произвольную реплику  $C$  оператора  $A_s$ . Согласно следствию теоремы 7.3.2.2  $A \rightarrow A_s$ , поэтому  $A \rightarrow C$ , а значит,  $C$  коммутирует с  $A$  и с  $A_n$ . Следовательно, оператор  $CA_n$  нильпотентен и  $\text{tr}(CA_n) = 0$ . Кроме того,  $\text{tr}(CA) = 0$ , а значит,  $\text{tr}(CA_s) = 0$ . Остается доказать, что если  $\text{tr}(CA_s) = 0$  для любой реплики  $C$  полупростого оператора  $A_s$ , то  $A_s = 0$ .

Пусть  $\lambda_1, \dots, \lambda_n$  — собственные значения оператора  $A_s$ ,  $\Gamma$  — пространство над полем рациональных чисел  $\mathbb{Q}$ , порождённое  $\lambda_1, \dots, \lambda_n$ . Требуется доказать, что  $\Gamma = \{0\}$ , т.е.  $\Gamma^* = \{0\}$ . Пусть  $\varphi \in \Gamma$  и  $\varphi(\lambda_i) = \mu_i$ . Существует многочлен  $g$ , для которого  $g(\lambda_i) = \mu_i$ . Покажем, что многочлен  $g$  удовлетворяет условию теоремы 7.3.3.2, т.е. оператор  $C = g(A_s)$  является репликой  $A_s$ . Пусть  $\sum n_i \lambda_i = 0$ , где числа  $n_i$  целые. Отображение  $\varphi$  линейно над  $\mathbb{Q}$ , поэтому  $\varphi(n_i) = n_i$ , а значит,  $0 = \sum n_i \varphi(\lambda_i) = \sum n_i \mu_i = \sum n_i g(\lambda_i)$ , что и требовалось. Таким образом, оператор  $C = g(A_s)$  является репликой  $A_s$ . Поэтому по предположению  $0 = \text{tr}(CA_s) = \sum g(\lambda_i) \lambda_i = \sum \varphi(\lambda_i) \lambda_i$ . Линейная над  $\mathbb{Q}$  функция  $\varphi$  принимает значения в  $\mathbb{Q}$ , поэтому  $\varphi(\varphi(\lambda_i)) = \varphi(\lambda_i)$ . Следовательно,  $0 = \varphi(\sum \varphi(\lambda_i) \lambda_i) = \sum (\varphi(\lambda_i))^2$ . Таким образом,  $\varphi(\lambda_i) = 0$ , т.е.  $\varphi = 0$ .  $\square$

### Задачи

**3.1.** Оператор  $A$  нильпотентный. Докажите, что  $A \rightarrow B$  тогда и только тогда, когда  $B = \lambda A$ .

## 7.4. Элементы теории алгебр Ли

*Алгебра Ли*  $\mathfrak{g}$  — это векторное пространство (над  $\mathbb{R}$  или над  $\mathbb{C}$ ), в котором для любых двух элементов  $X, Y \in \mathfrak{g}$  задан *коммутатор*  $[X, Y] \in \mathfrak{g}$ , обладающий следующими свойствами:

- 1)  $[\lambda_1 X_1 + \lambda_2 X_2, Y] = \lambda_1 [X_1, Y] + \lambda_2 [X_2, Y]$ ;
- 2)  $[X, Y] = -[Y, X]$ ;
- 3)  $[X, [Y, Z]] + [Y, [X, Z]] + [Z, [X, Y]] = 0$  (*тождество Якоби*).

Простейшим примером алгебры Ли является *абелева* (*коммутативная*) алгебра Ли  $\mathfrak{g}$ , для которой  $[A, B] = 0$  для всех  $A, B \in \mathfrak{g}$ .

Ещё одним примером является алгебра Ли всех линейных операторов в пространстве  $V$ , где коммутатор элементов алгебры Ли — обычный коммутатор линейных операторов. Эту алгебру мы будем обозначать  $\mathfrak{gl}(V)$ . Ту же самую алгебру часто обозначают  $\mathfrak{gl}(n)$ , где  $n = \dim V$ . Если нужно уточнить, над каким именно полем рассматривается эта алгебра Ли, используются обозначения  $\mathfrak{gl}(n, \mathbb{R})$  и  $\mathfrak{gl}(n, \mathbb{C})$ .

Можно также рассмотреть алгебру Ли матриц с нулевым следом  $\mathfrak{sl}(n)$ , алгебру Ли кососимметрических матриц  $\mathfrak{so}(n)$  и алгебру Ли косоэрмитовых матриц  $\mathfrak{su}(n)$ .

Отображение  $f: \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$  называют *гомоморфизмом* алгебр Ли, если  $f$  — линейное отображение  $f([X, Y]) = [f(X), f(Y)]$  для всех  $X, Y \in \mathfrak{g}_1$ . Если при этом  $f$  — изоморфизм линейных пространств, то  $f$  называют *изоморфизмом*.

Алгебру Ли  $\mathfrak{g}$  называют *матричной алгеброй Ли*, если она является подалгеброй Ли в  $\mathfrak{gl}(V)$  для некоторого пространства  $V$ .

Линейное подпространство  $\mathfrak{a}$  в алгебре Ли  $\mathfrak{g}$  называют *идеалом*, если  $[A, X] \in \mathfrak{a}$  для любых матриц  $A \in \mathfrak{a}$  и  $X \in \mathfrak{g}$ . Для двух идеалов  $\mathfrak{a}_1$  и  $\mathfrak{a}_2$  в алгебре Ли  $\mathfrak{g}$  можно рассмотреть множество

$$[\mathfrak{a}_1, \mathfrak{a}_2] = \left\{ \sum [X_i, Y_i] \mid X_i \in \mathfrak{a}_1, Y_i \in \mathfrak{a}_2 \right\}.$$

Тождество  $[Z, [X, Y]] = [[Z, X], Y] + [X, [Z, Y]]$ , которое следует из тождества Якоби, показывает, что  $[\mathfrak{a}_1, \mathfrak{a}_2]$  — идеал.

Для любого идеала  $\mathfrak{a} \subset \mathfrak{g}$  можно определить *факторалгебру* Ли  $\mathfrak{g}/\mathfrak{a}$ . Коммутатор классов эквивалентности определяется следующим образом:  $[X + \mathfrak{a}, Y + \mathfrak{a}] = [X, Y] + \mathfrak{a}$ . Это определение корректно, потому что если  $A_1, A_2 \in \mathfrak{a}$ , то  $[X + A_1, Y + A_2] = [X, Y] + [A_1, Y] + [X, A_2] + [A_1, A_2]$  и  $[A_1, Y] + [X, A_2] + [A_1, A_2] \in \mathfrak{a}$ .

*Центром* алгебры Ли  $\mathfrak{g}$  называют множество

$$Z(\mathfrak{g}) = \{Z \in \mathfrak{g} \mid [X, Z] = 0 \quad \forall X \in \mathfrak{g}\}.$$

Если  $Z_1, Z_2 \in Z(\mathfrak{g})$ , то  $[X, [Z_1, Z_2]] = -[Z_1, [Z_2, X]] - [Z_2, [X, Z_1]] = 0$  для всех  $X \in \mathfrak{g}$ , поэтому центр алгебры Ли является идеалом.

Для каждого целого неотрицательного числа  $m$  можно определить по индукции идеалы  $\mathfrak{g}^{(m)}$  и  $\mathfrak{g}_m$ :  $\mathfrak{g}^{(0)} = \mathfrak{g}_0 = \mathfrak{g}$ ,  $\mathfrak{g}^{(m)} = [\mathfrak{g}^{(m-1)}, \mathfrak{g}^{(m-1)}]$ ,  $\mathfrak{g}_m = [\mathfrak{g}, \mathfrak{g}_{m-1}]$ . Ясно, что при этом  $\mathfrak{g}^{(m)} \subset \mathfrak{g}_m$ .

Алгебру Ли  $\mathfrak{g}$  называют *нильпотентной*, если  $\mathfrak{g}_m = \{0\}$  для некоторого  $m$ .

Алгебру Ли  $\mathfrak{g}$  называют *разрешимой*, если  $\mathfrak{g}^{(m)} = \{0\}$  для некоторого  $m$ .

**Пример 7.4.0.1.** Алгебра Ли  $\mathfrak{n}$  верхних треугольных матриц с нулями на главной диагонали нильпотентна.

**Доказательство.** Пусть  $H_m$  — пространство матриц  $\|a_{ij}\|$ , для которых  $a_{ij} = 0$  при  $i < j + m + 1$ , т.е. если  $a_{ij} \neq 0$ , то  $i \geq j + m + 1$ . Покажем, что  $\mathfrak{n}_m \subset H_m$ . Ясно, что  $\mathfrak{n}_0 = H_0$ . Пусть  $\|a_{ij}\| \in H_m$  и  $\|b_{ij}\| \in H_n$ . Рассмотрим матрицу  $\|c_{ij}\| = \|a_{ij}\| \cdot \|b_{ij}\|$ . Предположим, что  $c_{ij} \neq 0$ . Тогда  $a_{ip}b_{pj} \neq 0$  для некоторого  $p$ . Следовательно,  $i \geq p + m + 1$  и  $p \geq j + n + 1$ , поэтому  $i \geq j + m + n + 2$ . Таким образом, произведение матриц из  $H_m$  и  $H_n$  лежит в  $H_{n+m+1}$ . В частности, произведение матриц из  $H_0$  и  $H_{m-1}$  лежит в  $H_m$ . Тем более, коммутатор этих матриц лежит в  $H_m$ . Остаётся заметить, что если  $m$  достаточно велико, то  $H_m = 0$ .  $\square$

**Пример 7.4.0.2.** Алгебра Ли  $\mathfrak{t}$  верхних треугольных матриц разрешима.

**Доказательство.** Пусть  $A$  и  $B$  — верхние треугольные матрицы с диагональными элементами  $a_1, \dots, a_n$  и  $b_1, \dots, b_n$ . Тогда  $AB$  и  $BA$  — верхние треугольные матрицы с диагональными элементами  $a_1b_1, \dots, a_nb_n$ . Поэтому  $[A, B]$  — верхняя треугольная матрица с нулями на диагонали. Таким образом, алгебра  $\mathfrak{t}_1 \subset \mathfrak{n}$  нильпотентна, а значит, алгебра  $\mathfrak{t}$  разрешима.  $\square$

### 7.4.1. Нильпотентные алгебры Ли

Нильпотентность алгебры Ли часто можно проверить, не вычисляя идеалы  $\mathfrak{g}_m$ . Это позволяет сделать теорема Энгеля. Для доказательства теоремы Энгеля нам понадобится следующее утверждение.

**Теорема 7.4.1.1.** Пусть все операторы алгебры Ли  $\mathfrak{g} \subset \mathfrak{gl}(V)$ , где  $V \neq 0$ , нильпотентны. Тогда существует ненулевой вектор  $v \in V$ , для которого  $Xv = 0$  для всех  $X \in \mathfrak{g}$ .

**Доказательство.** Применим индукцию по  $\dim \mathfrak{g}$ . При  $\dim \mathfrak{g} = 1$  утверждение очевидно: у любого нильпотентного оператора есть собственный вектор с нулевым собственным значением. Предположим теперь, что  $\dim \mathfrak{g} > 1$ . Пусть  $\mathfrak{h} \subset \mathfrak{g}$  — некоторая собственная (т.е. отличная от  $\mathfrak{g}$ ) подалгебра Ли. (Отметим, что одномерное подпространство в алгебре Ли всегда является подалгеброй Ли). Если оператор  $X$  нильпотентный, то оператор  $\text{ad}_X$  тоже нильпотентный (задача 2.5). Поэтому если  $X \in \mathfrak{h}$ , то  $\text{ad}_X$  — нильпотентный оператор на  $\mathfrak{g}$ , переводящий  $\mathfrak{h}$  в  $\mathfrak{h}$ . В таком случае  $\text{ad}_X$  — нильпотентный оператор на  $\mathfrak{g}/\mathfrak{h}$ . Такие операторы образуют алгебру Ли, поскольку  $[\text{ad}_X, \text{ad}_Y] = \text{ad}_{[X, Y]}$ . Размерность этой алгебры Ли не превосходит размерности  $\mathfrak{h}$ , поэтому к ней можно применить предположение индукции. В результате получим, что можно выбрать  $X \in \mathfrak{g}$  так, что класс  $X + \mathfrak{h} \neq \mathfrak{h}$  в  $\mathfrak{g}/\mathfrak{h}$  аннулируется всеми операторами  $\text{ad}_Y$ ,  $Y \in \mathfrak{h}$ . Это означает, что  $[Y, X] \in \mathfrak{h}$  для всех  $Y \in \mathfrak{h}$ . Таким образом, нормализатор  $N(\mathfrak{h}) = \{X \in \mathfrak{g} \mid [X, \mathfrak{h}] \in \mathfrak{h}\}$  отличен от  $\mathfrak{h}$ .

Можно считать, что  $\mathfrak{h}$  — максимальная собственная подалгебра в  $\mathfrak{g}$  (т.е. она не содержится ни в какой другой собственной подалгебре в  $\mathfrak{g}$ ). Тожество Якоби показывает, что нормализатор  $N(\mathfrak{h})$  является подалгеброй в  $\mathfrak{g}$ , поэтому из максимальной  $\mathfrak{h}$  следует, что  $N(\mathfrak{h}) = \mathfrak{g}$ , т.е.  $\mathfrak{h}$  — идеал  $\mathfrak{g}$ . Покажем, что коразмерность  $\mathfrak{h}$  в  $\mathfrak{g}$  равна 1. Предположим, что  $\dim \mathfrak{g}/\mathfrak{h} > 1$ . Тогда прообраз в  $\mathfrak{g}$  одномерной подалгебры в  $\mathfrak{g}/\mathfrak{h}$  является собственной подалгеброй, строго содержащей  $\mathfrak{h}$ , а это противоречит минимальности  $\mathfrak{h}$ .

Итак,  $\mathfrak{g}$  как линейное пространство порождено  $\mathfrak{h}$  и некоторым оператором  $Z$ . По предположению индукции пространство  $W = \{v \in V \mid \mathfrak{h}v = 0\}$  ненулевое. Кроме того, как было доказано,  $\mathfrak{h}$  — идеал в  $\mathfrak{g}$ , поэтому пространство  $W$  инвариантно относительно действия  $\mathfrak{g}$ . Действительно, если  $X \in \mathfrak{g}$ ,  $Y \in \mathfrak{h}$  и  $w \in W$ , то  $YXw = XYw - [X, Y]w = 0$ , поскольку  $[X, Y] \in \mathfrak{h}$ . Ограничение нильпотентного оператора  $Z$  на подпространство  $W$  нильпотентно, поэтому существует ненулевой вектор  $v \in W$ , для которого  $Zv = 0$ . В итоге получаем, что все операторы из  $\mathfrak{g}$  аннулируют  $v$ , что и требовалось.  $\square$

**Теорема 7.4.1.2 (Энгель).** *Алгебра Ли  $\mathfrak{g}$  нильпотентна тогда и только тогда, когда для любого элемента  $X \in \mathfrak{g}$  оператор  $\text{ad}_X : \mathfrak{g} \rightarrow \mathfrak{g}$  нильпотентен.*

**Доказательство.** Если  $X \in \mathfrak{g}$ , то  $\text{ad}_X \mathfrak{g} \subset \mathfrak{g}_1$ ,  $\text{ad}_X \mathfrak{g}_1 \subset \mathfrak{g}_2$ , ...,  $\text{ad}_X \mathfrak{g}_{m-1} \subset \mathfrak{g}_m$ . Поэтому если алгебра Ли  $\mathfrak{g}$  нильпотентна, то оператор  $\text{ad}_X : \mathfrak{g} \rightarrow \mathfrak{g}$  нильпотентен.

Обратное утверждение докажем индукцией по размерности алгебры Ли  $\mathfrak{g}$ . Если  $\dim \mathfrak{g} = 1$ , то алгебра  $\mathfrak{g}$  нильпотентна. Операторы  $\text{ad}_X$ ,  $X \in \mathfrak{g}$ , образуют подалгебру Ли в  $\mathfrak{gl}(\mathfrak{g})$ , поскольку  $[\text{ad}_X, \text{ad}_Y] = \text{ad}_{[X, Y]}$ . Все операторы в этой алгебре Ли нильпотентны, поэтому можно применить теорему 7.4.1.1. (Случай  $\mathfrak{g} = 0$  очевиден, поэтому мы будем предполагать, что  $\mathfrak{g} \neq 0$ .) Таким образом, в алгебре Ли  $\mathfrak{g}$  существует элемент  $X \neq 0$ , для которого  $[\mathfrak{g}, X] = 0$ . Следовательно,  $Z(\mathfrak{g}) \neq 0$ . Алгебра Ли  $\mathfrak{g}/Z(\mathfrak{g})$  тоже обладает тем свойством, что на ней все операторы  $\text{ad}$  нильпотентны, причём её размерность строго меньше размерности  $\mathfrak{g}$ . По предположению индукции алгебра Ли  $\mathfrak{g}/Z(\mathfrak{g})$  нильпотентна, т.е.  $\mathfrak{g}_m \subset Z(\mathfrak{g})$  для некоторого  $m$ . Следовательно,  $\mathfrak{g}_{m+1} = [\mathfrak{g}, \mathfrak{g}_m] \subset [\mathfrak{g}, Z(\mathfrak{g})] = 0$ , поэтому алгебра Ли  $\mathfrak{g}$  нильпотентна.  $\square$

**Следствие.** *Подалгебра Ли в  $\mathfrak{gl}(V)$ , состоящая из нильпотентных операторов, нильпотентна.*

**Доказательство.** Если оператор  $A : V \rightarrow V$  нильпотентен, то оператор  $\text{ad}_A : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$  тоже нильпотентен (задача 2.5). Поэтому к рассматриваемой алгебре Ли можно применить теорему Энгеля.  $\square$

### 7.4.2. Разрешимые алгебры Ли

Для разрешимых алгебр Ли имеет место теорема Ли, аналогичная теореме 7.4.1.1. Но теорема Ли, в отличие от теоремы 7.4.1.1, верна только над полем комплексных чисел. Дело в том, что нильпотентный оператор имеет собственное значение 0, которое вещественно. Поэтому нильпотентный оператор над полем вещественных чисел имеет собственный вектор. А произвольный оператор над полем вещественных чисел может и не иметь собственных векторов.

**Теорема 7.4.2.1 (Ли).** *Пусть  $V \neq 0$  — векторное пространство над  $\mathbb{C}$  и  $\mathfrak{g} \subset \mathfrak{gl}(V)$  — разрешимая алгебра Ли. Тогда в  $V$  есть ненулевой вектор, который является собственным для всех операторов из  $\mathfrak{g}$ .*

**Доказательство.** Применим индукцию по  $\dim \mathfrak{g}$ . При  $\dim \mathfrak{g} = 1$  утверждение очевидно: у любого оператора над  $\mathbb{C}$  есть собственный вектор.

Предположим теперь, что  $\dim \mathfrak{g} > 1$ . Как и при доказательстве теоремы 7.4.1.1, будем доказывать, что в  $\mathfrak{g}$  есть идеал  $\mathfrak{h}$  коразмерности 1. Алгебра Ли  $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$  абелева, поэтому любое подпространство в ней является идеалом. Выберем в  $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$  подпространство коразмерности 1. Пробраз этого подпространства при естественной проекции  $\mathfrak{g} \rightarrow \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$  — это искомый идеал  $\mathfrak{h}$  коразмерности 1.

Идеал  $\mathfrak{h}$  разреши, поэтому согласно предположению индукции в пространстве  $V$  можно выбрать вектор  $v \neq 0$  так, что  $Yv = \lambda(Y)v$  для всех  $Y \in \mathfrak{h}$ . Здесь  $\lambda$  — некоторая линейная функция на  $\mathfrak{h}$ . Рассмотрим в  $V$  подпространство

$$W = \{w \in V \mid Xw = \lambda(X)w \quad \forall X \in \mathfrak{h}\}.$$

Это подпространство ненулевое, потому что оно содержит вектор  $v$ .

Покажем, что подпространство  $W$  инвариантно относительно действия  $\mathfrak{g}$ , т.е. если  $X \in \mathfrak{g}$  и  $w \in W$ , то  $Xw \in W$ . Пусть  $Y \in \mathfrak{h}$ . Тогда  $[X, Y] \in \mathfrak{h}$ , а потому  $YXw = XYw - [X, Y]w = \lambda(Y)Xw - \lambda([X, Y])w$ . Мы хотим доказать, что  $YXw = \lambda(Y)Xw$ , т.е.  $\lambda([X, Y]) = 0$ . Фиксируем  $X \in \mathfrak{g}$  и  $w \in W$ . Выберем наименьшее натуральное число  $n$ , для которого векторы  $w, Xw, \dots, X^n w$  линейно зависимы, и рассмотрим пространства  $W_k = \langle w, Xw, \dots, X^{k-1}w \rangle$  (мы полагаем  $W_0 = 0$ ). Ясно, что  $\dim W_n = n$ ,  $W_{n+i} = W_n$  для всех натуральных  $i$  и  $XW_n \subset W_n$ . Докажем индукцией по  $k$ , что  $YX^k w - \lambda(Y)X^k w \in W_k$  для всех  $Y \in \mathfrak{h}$ . При  $k = 0$  это следует непосредственно из определения пространства  $W$ . Если  $k > 0$ , то  $YX^k w = YXX^{k-1}w = XYX^{k-1}w - [X, Y]X^{k-1}w$ . По предположению индукции  $YX^{k-1}w = \lambda(Y)X^{k-1}w + w'$ , где  $w' \in W_{k-1}$ . Ясно также, что  $Xw' \in W_k$  и  $[X, Y]X^{k-1}w \in W_k$ , поскольку  $[X, Y] \in \mathfrak{h}$ . Таким образом, ограничение оператора  $Y$  на пространство  $W_n$  в базисе  $w, Xw, \dots, X^{n-1}w$  записывается верхней треугольной матрицей с элементами  $\lambda(Y)$  на диагонали. В частности,  $\operatorname{tr} Y|_{W_n} = n\lambda(Y)$  для любого  $Y \in \mathfrak{h}$ . Это равенство должно выполняться и для оператора  $[X, Y]$ , поскольку он тоже принадлежит  $\mathfrak{h}$ . Оба оператора  $X$  и  $Y$  сохраняют  $W_n$ , поэтому ограничение оператора  $[X, Y]$  на  $W_n$  является коммутатором двух операторов, а значит, его след равен нулю. Теперь из равенства  $\operatorname{tr}[X, Y]|_{W_n} = n\lambda([X, Y])$  следует требуемое равенство  $\lambda([X, Y]) = 0$ .

Алгебра  $\mathfrak{g}$  как линейное пространство порождена подпространством  $\mathfrak{h}$  коразмерности 1 и некоторым оператором  $Z$ . Как мы только что доказали,  $Z$  сохраняет подпространство  $W$ , поэтому можно рассмотреть ограничение  $Z$  на  $W$  и выбрать собственный вектор  $w_0$  (здесь мы пользуемся тем, что  $Z$  — оператор над полем  $\mathbb{C}$ ). Ясно, что этот вектор  $w_0$  является собственным вектором любого оператора из  $\mathfrak{g}$ .  $\square$



### 7.4.3. Представления алгебры Ли $\mathfrak{sl}(2, \mathbb{C})$

*Представлением* алгебры Ли  $\mathfrak{g}$  в пространстве  $V$  называют гомоморфизм алгебр Ли  $f: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ . Представление называют *неприводимым*, если единственными подпространствами в  $V$ , инвариантными относительно всех элементов  $\mathfrak{g}$ , являются  $V$  и  $\{0\}$ . Представление называют *точным*, если гомоморфизм  $f: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  имеет нулевое ядро.

Для любой алгебры Ли  $\mathfrak{g}$  можно рассмотреть *присоединённое представление*  $f: \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ , которое сопоставляет каждому элементу  $X \in \mathfrak{g}$  оператор  $\text{ad}_X$  в пространстве  $\mathfrak{g}$ . Присоединённое представление является точным тогда и только тогда, когда центр алгебры Ли равен нулю.

Два представления  $f_1: \mathfrak{g} \rightarrow \mathfrak{gl}(V_1)$  и  $f_2: \mathfrak{g} \rightarrow \mathfrak{gl}(V_2)$  называют *эквивалентными*, если существует такой изоморфизм линейных пространств  $\varphi: V_1 \rightarrow V_2$ , что  $\varphi f_1(X)v = f_2(X)\varphi v$  для всех  $X \in \mathfrak{g}$  и  $v \in V_1$ , т.е.  $\varphi f_1(X)\varphi^{-1} = f_2(X)$ .

Здесь мы опишем все представления алгебры Ли  $\mathfrak{sl}(2, \mathbb{C})$  с точностью до эквивалентности.

В пространстве  $\mathfrak{sl}(2, \mathbb{C})$  можно выбрать базис  $X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ,  $Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ . Несложные вычисления показывают, что

$$[X, Y] = H, \quad [H, X] = 2X \quad \text{и} \quad [H, Y] = -2Y.$$

Пусть дано представление алгебры Ли  $\mathfrak{sl}(2, \mathbb{C})$  в пространстве  $V$ . Для каждого  $\lambda \in \mathbb{C}$  рассмотрим пространство собственных векторов оператора  $H$  (который мы рассматриваем как оператор в  $V$ ):

$$V^\lambda = \{v \in V \mid Hv = \lambda v\}.$$

Ясно, что  $V = \bigoplus_\lambda V^\lambda$ .

Легко проверить, что если  $v \in V^\lambda$ , то  $Xv \in V^{\lambda+2}$  и  $Yv \in V^{\lambda-2}$ . Действительно,  $HXv = [H, X]v + XHv = 2Xv + \lambda Xv$  и  $HYv = [H, Y]v + YHv = -2Yv + \lambda Yv$ .

Выберем в пространстве  $V$  собственный вектор  $v$  оператора  $H$  и рассмотрим векторы  $v, Xv, X^2v, X^3v, \dots$ . Эти векторы лежат в пространствах  $V^\lambda, V^{\lambda+2}, V^{\lambda+4}, V^{\lambda+6}, \dots$ , среди которых есть лишь конечное число ненулевых. Пусть  $e$  — последний ненулевой вектор в рассматриваемой последовательности векторов. Он является собственным вектором оператора  $H$  с некоторым собственным значением  $\lambda$ , и при этом  $Xe = 0$ .

Положим  $e_n = \frac{1}{n!} Y^n e$  при  $n \geq 0$  и  $e_{-1} = 0$ . Непосредственно из определения видно, что  $He_n = (\lambda - 2n)e_n$  и  $Ye_n = (n+1)e_{n+1}$ . Докажем индукцией по  $n$ , что  $Xe_n = (\lambda - n + 1)e_{n-1}$ . При  $n = 0$  это равенство очевидно.

Пусть теперь  $n > 0$ . Тогда

$$\begin{aligned} nXe_n &= XYe_{n-1} = [X, Y]e_{n-1} + YXe_{n-1} = \\ &= He_{n-1} + (\lambda - n + 2)Ye_{n-2} = \\ &= ((\lambda - 2n - 2) + (\lambda - n + 2)(n - 1))e_{n-1} = \\ &= n(\lambda - n + 1)e_{n-1}. \end{aligned}$$

Пространство  $V$  конечномерное и все ненулевые векторы  $e_n$  линейно независимы, поскольку они являются собственными векторами оператора  $H$ , соответствующими разным собственным значениям. Поэтому можно выбрать целое неотрицательное число  $m$  так, что  $e_m \neq 0$  и  $e_{m+1} = 0$ . Если рассматриваемое представление неприводимо, то  $V = \langle e_0, e_1, \dots, e_m \rangle$ . Напомним, что операторы  $H$ ,  $X$  и  $Y$  действуют в этом пространстве следующим образом:

$$He_n = (m - 2n)e_n, \quad Xe_n = (m - n + 1)e_{n-1} \quad \text{и} \quad Ye_n = (n + 1)e_{n+1}. \quad (7.1)$$

Мы доказали, что любое неприводимое представление алгебры Ли  $\mathfrak{sl}(2, \mathbb{C})$  в пространстве  $W_m = \langle e_0, e_1, \dots, e_m \rangle$  размерности  $m + 1$  эквивалентно представлению, для которого операторы  $H$ ,  $X$  и  $Y$  заданы формулами (7.1). Легко проверить, что для любого целого неотрицательного  $m$  эти формулы задают некоторое представление алгебры Ли  $\mathfrak{sl}(2, \mathbb{C})$ . Действительно,

$$\begin{aligned} HXe_n - XHe_n &= 2Xe_n, \\ HYe_n - YHe_n &= -2Ye_n, \\ XYe_n - YXe_n &= He_n. \end{aligned}$$

Таким образом, в пространстве каждой размерности с точностью до эквивалентности есть ровно одно неприводимое представление алгебры Ли  $\mathfrak{sl}(2, \mathbb{C})$ .

#### 7.4.4. Критерий Картана

Подалгебру Ли  $\mathfrak{g} \subset \mathfrak{gl}(V)$  называют *алгебраической*, если вместе с каждым оператором  $X \in \mathfrak{g}$  она содержит любую его реплику  $X'$ .

Ясно, что алгебра Ли  $\mathfrak{gl}(V)$  алгебраична.

Пересечение всех алгебраических алгебр Ли, содержащих данную алгебру Ли  $\mathfrak{g} \subset \mathfrak{gl}(V)$ , — это некоторая алгебраическая алгебра  $\bar{\mathfrak{g}}$ ; её называют *алгебраическим замыканием* алгебры Ли  $\mathfrak{g}$ .

**Теорема 7.4.4.1.** Пусть в пространстве  $V$  заданы подпространства  $W_1$  и  $W_2$ , причём  $W_1 \supset W_2$ . Тогда множество

$$\mathfrak{g} = \{X \in \mathfrak{gl}(V) \mid \text{ad}_X W_1 \subset W_2\}$$

является алгебраической алгеброй Ли.

**Доказательство.** Прежде всего проверим, что  $\mathfrak{g}$  — алгебра Ли. Пусть  $[X, W_1] \subset W_2$  и  $[Y, W_1] \subset W_2$ . Тогда

$$[[X, Y], W_1] = -[X, [Y, W_1]] + [Y, [X, W_1]].$$

Здесь  $[Y, W_1] \subset W_2 \subset W_1$  и  $[X, W_1] \subset W_2 \subset W_1$ , поэтому  $[[X, Y], W_1] \subset [X, W_1] + [Y, W_1] \subset W_2$ .

Проверим теперь, что алгебра Ли  $\mathfrak{g}$  алгебраическая. Напомним, что  $\text{ad } A = A_{1,1}$ . Если  $B$  — реплика  $A$ , то согласно следствию теоремы 7.3.1.1 оператор  $B_{1,1} = \text{ad}_B$  можно представить в виде многочлена от  $\text{ad } A$  без свободного члена. Поэтому из включения  $\text{ad}_A W_1 \subset W_2$  следует включение  $\text{ad}_B W_1 \subset W_2$ .  $\square$

**Следствие.** Коммутант  $[\bar{\mathfrak{g}}, \bar{\mathfrak{g}}]$  совпадает с  $[\mathfrak{g}, \mathfrak{g}]$ .

**Доказательство.** Достаточно проверить, что  $[\bar{\mathfrak{g}}, \bar{\mathfrak{g}}] \subset [\mathfrak{g}, \mathfrak{g}]$ . Воспользуемся дважды теоремой 7.4.4.1. Сначала положим  $W_1 = \mathfrak{g}$  и  $W_2 = [\mathfrak{g}, \mathfrak{g}]$ . Из включения  $[\mathfrak{g}, W_1] = [\mathfrak{g}, \mathfrak{g}] \subset W_2$  следует включение  $[\bar{\mathfrak{g}}, W_1] = [\bar{\mathfrak{g}}, \mathfrak{g}] \subset W_2$ . Затем положим  $W_1 = \bar{\mathfrak{g}}$  и  $W_2 = [\mathfrak{g}, \mathfrak{g}]$ . Из включения  $[\mathfrak{g}, W_1] = [\mathfrak{g}, \bar{\mathfrak{g}}] \subset W_2$  следует включение  $[\bar{\mathfrak{g}}, \bar{\mathfrak{g}}] = [\bar{\mathfrak{g}}, W_1] \subset W_2$ .  $\square$

**Теорема 7.4.4.2.** Пусть  $\mathfrak{g} \subset \mathfrak{gl}(V)$  — подалгебра Ли и  $\mathfrak{n}$  — множество тех операторов  $N \in \mathfrak{g}$ , для которых  $\text{tr}(NX) = 0$  для всех  $X \in \mathfrak{g}$ . Тогда:

- а)  $\mathfrak{n}$  — идеал в  $\mathfrak{g}$ ;
- б) если алгебра Ли  $\mathfrak{g}$  алгебраична, то любой элемент из  $\mathfrak{n}$  нильпотентен;
- в) любой элемент из  $[\mathfrak{n}, \mathfrak{g}]$  нильпотентен.

**Доказательство.** а) Пусть  $X, Y \in \mathfrak{g}$  и  $N \in \mathfrak{n}$ . Тогда  $\text{tr}([X, Y]N) = 0$ , поэтому тождество  $\text{tr}([X, Y]N) = \text{tr}(X[Y, N])$  (задача 2.2) показывает, что  $\text{tr}(X[Y, N]) = 0$ . Следовательно,  $[Y, N] \in \mathfrak{n}$ .

б) Если алгебра  $\mathfrak{g}$  алгебраична, то для любого оператора  $N \in \mathfrak{n}$  и любой его реплики  $N'$  выполняется равенство  $\text{tr}(N'N) = 0$ . Поэтому согласно теореме 7.3.4.1 оператор  $N$  нильпотентен.

в) Для алгебры Ли  $\mathfrak{g}$  рассмотрим её алгебраическое замыкание  $\bar{\mathfrak{g}}$ . Пусть  $\mathfrak{n}' = \{N \in \bar{\mathfrak{g}} \mid \text{tr}(NX) = 0 \ \forall X \in \bar{\mathfrak{g}}\}$ . Достаточно проверить, что

$[\mathfrak{n}, \mathfrak{g}] \subset \mathfrak{n}'$ . Пусть  $X \in \bar{\mathfrak{g}}$ ,  $X_i \in \mathfrak{g}$  и  $N_i \in \mathfrak{n}$ . Тогда

$$\operatorname{tr} \left( \sum_i [N_i, X_i] X \right) = \operatorname{tr} \left( \sum_i N_i [X_i, X] \right),$$

причём  $[X_i, X] \in [\mathfrak{g}, \bar{\mathfrak{g}}] \subset [\bar{\mathfrak{g}}, \bar{\mathfrak{g}}] = [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$  (мы воспользовались следствием теоремы 7.4.4.1). Поэтому  $\operatorname{tr}(N_i [X_i, X]) = 0$ , так как  $N_i \in \mathfrak{n}$ .  $\square$

**Следствие.** Если билинейная форма  $\operatorname{tr}(XY)$  на матричной алгебре Ли  $\mathfrak{g}$  тождественно равна нулю, то эта алгебра Ли разрешима.

**Доказательство.** Билинейная форма  $\operatorname{tr}(XY)$  тождественно равна нулю, поэтому  $\mathfrak{g} = \mathfrak{n}$ . Следовательно, алгебра Ли  $[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{n}]$  нильпотентна, а значит, алгебра Ли  $\mathfrak{g}$  разрешима.  $\square$

Алгебру Ли  $\mathfrak{g}$  называют *полупростой*, если она не содержит ненулевых разрешимых идеалов.

Алгебру Ли  $\mathfrak{g}$  называют *простой*, если её размерность больше 1 и она не содержит нетривиальных идеалов.

Любая простая алгебра Ли является полупростой. Действительно, если простая алгебра Ли  $\mathfrak{g}$  содержит ненулевой разрешимый идеал, то тогда она сама должна быть разрешима. Поэтому она содержит идеал ко-размерности 1, чего не может быть.

**Теорема 7.4.4.3.** Пусть  $\mathfrak{g}$  — полупростая алгебра Ли и  $f: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  — её точное представление. Тогда билинейная форма  $\operatorname{tr}(f(X)f(Y))$  невырождена.

**Доказательство.** Рассмотрим множество  $\mathfrak{n} = \{N \in \mathfrak{g} \mid \operatorname{tr}(f(N)f(X)) = 0 \quad \forall X \in \mathfrak{g}\}$ . Согласно теореме 7.4.4.2 множество  $f(\mathfrak{n})$  является идеалом в алгебре Ли  $f(\mathfrak{g}) \subset \mathfrak{gl}(V)$ . Множество  $\{N \in f(\mathfrak{n}) \mid \operatorname{tr}(NX) = 0 \quad \forall X \in f(\mathfrak{n})\}$  совпадает с  $f(\mathfrak{n})$ , поэтому согласно следствию теоремы 7.4.4.2 алгебра Ли  $f(\mathfrak{n})$  разрешима, т.е.  $f(\mathfrak{n})$  — разрешимый идеал в алгебре Ли  $f(\mathfrak{g})$ . Представление  $f$  точное, поэтому  $\mathfrak{n}$  — разрешимый идеал в алгебре Ли  $\mathfrak{g}$ . Но по условию алгебра Ли  $\mathfrak{g}$  полупростая, поэтому  $\mathfrak{n} = 0$ . Это означает, что билинейная форма  $\operatorname{tr}(f(X)f(Y))$  невырождена.  $\square$

*Формой Киллинга* называют билинейную форму  $B(X, Y) = \operatorname{tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y)$  на алгебре Ли  $\mathfrak{g}$ .

**Теорема 7.4.4.4 (Картан).** Алгебра Ли  $\mathfrak{g}$  полупроста тогда и только тогда, когда форма Киллинга на ней невырождена.

**Доказательство.** Предположим сначала, что алгебра Ли  $\mathfrak{g}$  полупростая. Центр алгебры Ли является абелевой алгеброй Ли, а любая абелева алгебра Ли разрешима. Поэтому алгебра Ли  $\mathfrak{g}$  имеет нулевой центр, а значит, её присоединённое представление точно. Поэтому согласно теореме 7.4.4.3 форма Киллинга невырождена.

Предположим теперь, что алгебра Ли  $\mathfrak{g}$  не полупростая. Пусть  $\mathfrak{r}$  — ненулевой разрешимый идеал в  $\mathfrak{g}$ . Выберем целое число  $m$  так, что  $\mathfrak{r}^{(m-1)} \neq 0$  и  $\mathfrak{r}^{(m)} = 0$ . Положим  $\mathfrak{a} = \mathfrak{r}^{(m-1)}$ . Тогда  $\mathfrak{a}$  — ненулевой идеал в  $\mathfrak{g}$ , причём  $[\mathfrak{a}, \mathfrak{a}] = 0$ . Для любых элементов  $A \in \mathfrak{a}$  и  $X \in \mathfrak{g}$  мы имеем

$$\mathrm{ad}_A \mathrm{ad}_X \mathrm{ad}_A(\mathfrak{g}) = [A, [X, [A, \mathfrak{g}]]] \subset [A, [X, \mathfrak{a}]] \subset [A, \mathfrak{a}] = 0.$$

Поэтому  $(\mathrm{ad}_A \mathrm{ad}_X)^2 = 0$ , т.е. оператор  $\mathrm{ad}_A \mathrm{ad}_X$  нильпотентен и  $\mathrm{tr}(\mathrm{ad}_A \mathrm{ad}_X) = 0$ . Таким образом,  $B(A, X) = 0$ . Идеал  $\mathfrak{a}$  ненулевой, поэтому элемент  $A \in \mathfrak{a}$  можно считать ненулевым. Для этого элемента равенство  $B(A, X) = 0$  выполняется для всех  $X \in \mathfrak{g}$ , т.е. форма Киллинга  $B$  вырожденная.  $\square$

## Решения

**1.1. а)** Ясно, что  $AX = \|\lambda_i x_{ij}\|_1^n$  и  $XA = \|\lambda_j x_{ij}\|_1^n$ , а значит,  $\lambda_i x_{ij} = \lambda_j x_{ij}$ . Поэтому  $x_{ij} = 0$  при  $i \neq j$ .

**б)** Согласно задаче а)  $X = \mathrm{diag}(x_1, \dots, x_n)$ . Легко проверить, что  $(NAX)_{i,i+1} = \lambda_{i+1} x_{i+1}$  и  $(XNA)_{i,i+1} = \lambda_{i+1} x_i$ . Поэтому  $x_i = x_{i+1}$  при  $i = 1, 2, \dots, n-1$ .

**1.2.** Достаточно воспользоваться результатом задачи 1.1.

**1.3.** Пусть  $p_1, \dots, p_n$  — суммы элементов строк матрицы  $X$ ,  $q_1, \dots, q_n$  — суммы элементов её столбцов. Тогда

$$AX = \begin{pmatrix} q_1 & \dots & q_n \\ \dots & \dots & \dots \\ q_1 & \dots & q_n \end{pmatrix} \quad \text{и} \quad XA = \begin{pmatrix} p_1 & \dots & p_1 \\ \dots & \dots & \dots \\ p_n & \dots & p_n \end{pmatrix}.$$

Поэтому  $AX = XA$  тогда и только тогда, когда  $q_1 = \dots = q_n = p_1 = \dots = p_n$ .

**1.4.** Равенство  $AP_\sigma = P_\sigma A$  можно переписать в виде  $A = P_\sigma^{-1} A P_\sigma$ . Если  $P_\sigma^{-1} A P_\sigma = \|b_{ij}\|_1^n$ , то  $b_{ij} = a_{\sigma(i), \sigma(j)}$  (см. задачу 2.3). Для любых чисел  $p$  и  $q$  существует такая перестановка  $\sigma$ , что  $p = \sigma(q)$ . Поэтому  $a_{qq} = b_{qq} = a_{\sigma(q), \sigma(q)} = a_{pp}$ , т.е. все диагональные элементы матрицы  $A$  равны. Если  $i \neq j$  и  $p \neq q$ , то существует такая перестановка  $\sigma$ , что  $i = \sigma(p)$  и  $j = \sigma(q)$ . Поэтому  $a_{pq} = b_{pq} = a_{\sigma(p), \sigma(q)} = a_{ij}$ , т.е. все внедиагональные элементы матрицы  $A$  равны. Следовательно,  $A = \alpha I + \beta(J - I) = (\alpha - \beta)I + \beta J$ .

**1.5.** Равенство  $P^{-1}JP = P_1^{-1}JP_1$  эквивалентно равенству  $P_1 P^{-1}J = JP_1 P^{-1}$ . Положим  $X = P_1 P^{-1}$ . Матрица  $X$  перестановочна с  $J$  и  $P_1 = XP$ .

**1.6.** Можно считать, что  $A = \text{diag}(A_1, \dots, A_k)$ , где  $A_i$  — жорданова клетка. Пусть  $\mu_1, \dots, \mu_k$  — попарно различные числа и  $B_i$  — жорданова клетка, соответствующая собственному значению  $\mu_i$  и имеющая такой же размер, как и  $A_i$ . Тогда в качестве  $B$  можно взять матрицу  $\text{diag}(B_1, \dots, B_k)$ .

**1.7.** а) Для коммутирующих матриц  $A$  и  $B$  справедлива формула  $(A + B)^n = \sum \binom{n}{k} A^k B^{n-k}$ . Пусть  $A^m = B^m = 0$ . Если  $n = 2m - 1$ , то либо  $k \geq m$ , либо  $n - k \geq m$ , поэтому  $(A + B)^n = 0$ .

б) Это утверждение следует из теоремы 7.1.2.3.

**1.8.** Инволюции являются диагонализруемыми операторами, причём они приводятся к диагональному виду с элементами  $\pm 1$  на диагонали (см. §4.8.1). Значит, согласно следствию теоремы 7.1.2.3 существует базис, в котором все матрицы  $A_i$  имеют вид  $\text{diag}(\pm 1, \dots, \pm 1)$ . Количество различных матриц такого вида равно  $2^n$ .

**1.9.** Разложим пространство  $V$  в такую прямую сумму инвариантных подпространств  $V_i$ , что каждый оператор  $A_j$  имеет на любом подпространстве  $V_i$  лишь одно собственное значение  $\lambda_{ji}$  (см. теорему 7.1.2.3). Рассмотрим диагональный оператор  $D$ , ограничение которого на подпространство  $V_i$  имеет вид  $\mu_i I$ , причём все числа  $\mu_i$  попарно различны. Для каждого  $j$  существует интерполяционный многочлен Лагранжа  $f_j$ , для которого  $f_j(\mu_i) = \lambda_{ji}$  при всех  $i$ . Ясно, что  $f_j(D) = A_j$ .

**1.10.** Пусть  $V$  — пространство решений уравнения  $AX = XA$ ,  $n$  — порядок матриц  $A$  и  $B$ . По условию  $V \subset \langle I, B, \dots, B^{k-1} \rangle$ , причём  $k \leq n$  согласно теореме Гамильтона—Кэли. С другой стороны,  $\dim V \geq n$  (см. доказательство теоремы 7.1.1.2). Следовательно,  $\dim V = n$ , а значит, минимальный многочлен матрицы  $A$  совпадает с характеристическим многочленом, т.е.  $\dim \langle I, A, \dots, A^{n-1} \rangle = n$ . Остаётся заметить, что  $\langle I, A, \dots, A^{n-1} \rangle \subset V$ .

**1.11.** Применим индукцию по  $k$ . При  $k = 1$  утверждение очевидно. Предположим, что требуемое утверждение доказано для блочной матрицы порядка  $k - 1$ . Перестановочность матриц  $A_{ij}$  приводит к матричному тождеству вида

$$\begin{pmatrix} I & 0 & \dots & 0 \\ -A_{21} & I & \dots & 0 \\ \dots & \dots & \dots & \dots \\ -A_{k1} & 0 & \dots & I \end{pmatrix} \begin{pmatrix} I & 0 & \dots & 0 \\ 0 & A_{11} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{11} \end{pmatrix} A = \begin{pmatrix} A_{11} & * & \dots & * \\ 0 & B_{11} & \dots & B_{1,n-1} \\ \dots & \dots & \dots & \dots \\ 0 & B_{n-1,1} & \dots & B_{n-1,n-1} \end{pmatrix}, \quad (1)$$

где  $B_{pq} = -A_{p+1,1}A_{1,q+1} + A_{11}A_{p+1,q+1}$ . Из перестановочности матриц  $A_{ij}$  следует перестановочность матриц  $B_{pq}$ , поэтому для блочной матрицы  $B = \|B_{pq}\|_1^{k-1}$  справедливо предположение индукции. Выражение определителя матрицы  $B$  через матрицы  $B_{pq}$  (и соответственно через матрицы  $A_{ij}$ ) выглядит почти так же как соответствующее выражение определителя матрицы порядка  $k - 1$  с элементами  $b_{pq}$ ; нужно только от этого выражения взять определитель. Но ясно, что если бы вместо матриц  $A_{ij}$  мы взяли числа  $a_{ij}$ , то получили бы тождество

$$a_{11} \det B = (a_{11})^{k-1} \left( \sum_{\sigma} (-1)^{\sigma} a_{1\sigma(1)} \dots a_{k\sigma(k)} \right),$$

т.е.

$$\det B = (a_{11})^{k-2} \left( \sum_{\sigma} (-1)^{\sigma} a_{1\sigma(1)} \cdots a_{k\sigma(k)} \right).$$

(сокращение на  $a_{11}$  возможно, потому что здесь это — независимая переменная). Следовательно, для блочной матрицы получаем тождество

$$\det B = (\det A_{11})^{k-2} \det \left( \sum_{\sigma} (-1)^{\sigma} A_{1\sigma(1)} \cdots A_{k\sigma(k)} \right).$$

Возвращаясь к матричному тождеству (1), получаем

$$\begin{aligned} (\det A_{11})^{k-1} \det A &= \det A_{11} \det B = \\ &= (\det A_{11})^{k-1} \det \left( \sum_{\sigma} (-1)^{\sigma} A_{1\sigma(1)} \cdots A_{k\sigma(k)} \right). \end{aligned}$$

Если  $\det A_{11} \neq 0$ , то из этого следует требуемое равенство. Теперь уже легко получить требуемое тождество и в случае, когда  $\det A_{11} = 0$ , воспользовавшись теми же самыми рассуждениями, что и при доказательстве теоремы 1.3.1.3.

**1.12.** Рассмотрим систему матричных уравнений  $XA = A^T X$ ,  $X = X^T$ . Она эквивалентна системе уравнений  $XA = A^T X^T$ ,  $X = X^T$ . Эта система уравнений означает, что матрицы  $X$  и  $XA$  симметрические. Условие симметричности матрицы порядка  $n$  записывается как система из  $(n^2 - n)/2$  линейных уравнений, поэтому всего мы получаем  $n^2 - n$  уравнений для  $n^2$  элементов матрицы  $X$ . Таким образом, рассматриваемая система уравнений имеет по крайней мере  $n$  линейно независимых решений.

Матрицы  $A$  и  $A^T$  имеют одну и ту же жорданову нормальную форму, поэтому существует невырожденная матрица  $P$ , для которой  $PAP^{-1} = A^T$ . Заметим теперь, что если  $XA = A^T X$ , то  $P^{-1}XA = P^{-1}A^T X = P^{-1}(PAP^{-1})X = AP^{-1}X$ . И наоборот, если  $P^{-1}XA = AP^{-1}X$ , то  $XA = PAP^{-1}X = A^T X$ . Поэтому отображение  $X \mapsto P^{-1}X$  является изоморфизмом пространства решений уравнения  $XA = A^T X$  на пространство матриц, коммутирующих с  $A$ . Если характеристический многочлен матрицы  $A$  совпадает с минимальным многочленом, то согласно теореме 7.1.1.2 размерность пространства матриц, коммутирующих с  $A$ , равна  $n$ . Поэтому размерность пространства решений матричного уравнения  $XA = A^T X$  тоже равна  $n$ . С другой стороны, размерность пространства решений системы уравнений  $XA = A^T X$ ,  $X = X^T$  не меньше  $n$ . Следовательно, любое решение уравнения  $XA = A^T X$  является также и решением уравнения  $X = X^T$ .

**2.1.** Легко проверить, что  $[N, A] = N$ . Поэтому  $\text{ad}_J A = [J, A] = [N, A] = N = J - \lambda I$ . Ясно также, что  $\text{ad}_J(J - \lambda I) = 0$ .

Для любых матриц  $X$  и  $Y$  справедливо равенство  $\text{ad}_Y((Y - \lambda I)X) = (Y - \lambda I)\text{ad}_Y X$ , поэтому  $\text{ad}_Y^2((Y - \lambda I)X) = (Y - \lambda I)\text{ad}_Y^2 X$ . Положив  $Y = J$  и  $X = A$ , получим  $\text{ad}_J^2(NA) = N\text{ad}_J^2 A = 0$ .

**2.2.** По определению  $[A, B]C = ABC - BAC$  и  $A[B, C] = ABC - ACB$ . Ясно также, что  $\text{tr } B(AC) = \text{tr}(AC)B$ .

**2.3.** Так как  $C^n = C^{n-1} \sum [A_i, B_i] = \sum C^{n-1} A_i B_i - \sum C^{n-1} B_i A_i = \sum A_i (C^{n-1} B_i) - \sum (C^{n-1} B_i) A_i = \sum [A_i, C^{n-1} B_i]$ , то  $\text{tr } C^n = 0$  при  $n \geq 1$ . Следовательно, матрица  $C$  нильпотентна (см. теорему 4.6.2.1).

**2.4.** При  $n = 1$  утверждение верно. Ясно также, что если утверждение верно для  $n$ , то

$$\begin{aligned} \text{ad}_A^{n+1}(B) &= \\ &= \sum_{i=0}^n (-1)^{n-1} \binom{n}{i} A^{i+1} B A^{n-i} - \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} A^i B A^{n-i+1} = \\ &= \sum_{i=1}^{n+1} (-1)^{n-i+1} \binom{n}{i-1} A^i B A^{n-i+1} + \sum_{i=0}^n (-1)^{n-i+1} \binom{n}{i} A^i B A^{n-i+1} = \\ &= \sum_{i=0}^{n+1} (-1)^{n+1-i} \binom{n+1}{i} A^i B A^{n+1-i}. \end{aligned}$$

**2.5.** Пусть  $A^k = 0$  и  $n = 2k - 1$ . Согласно задаче 2.4  $\text{ad}_A^n(B) = \sum_{i=0}^n (-1)^i \binom{n}{i} A^i B A^{n-i}$ . Одно из чисел  $i$  и  $n - i$  не меньше  $k$ , поэтому  $A^i B A^{n-i} = 0$ . Следовательно,  $\text{ad}_A^n(B) = 0$ .

**2.6.** Отображение  $D = \text{ad}_A: M_{n,n} \rightarrow M_{n,n}$  является дифференцированием. Нужно доказать, что если  $D^2 B = 0$ , то  $D^n(B^n) = n!(DB)^n$ . Для  $n = 1$  утверждение очевидно. Предположим, что утверждение верно для некоторого  $n$ . Тогда

$$\begin{aligned} D^{n+1}(B^n) &= D[D^n(B^n)] = n! D[(DB)^n] = \\ &= n! \sum_{i=0}^{n-1} (DB)^i (D^2 B) (DB)^{n-1-i} = 0. \end{aligned}$$

Ясно, что

$$D^{n+1}(B^{n+1}) = D^{n+1}(B \cdot B^n) = \sum_{i=0}^{n+1} \binom{n+1}{i} (D^i B) (D^{n+1-i}(B^n)).$$

А так как  $D^i B = 0$  при  $i \geq 2$ , то  $D^{n+1}(B^{n+1}) = B \cdot D^{n+1}(B^n) + (n+1)(DB)(D^n(B^n)) = (n+1)(DB)(D^n(B^n)) = (n+1)!(DB)^{n+1}$ .

**2.7.** Докажем сначала требуемое утверждение при  $n = 1$ . При  $m = 1$  утверждение очевидно. Ясно также, что если утверждение верно для некоторого  $m$ , то  $[A^{m+1}, B] = A(A^m B - B A^m) + (AB - BA)A^m = mA[A, B]A^{m-1} + [A, B]A^m = (m+1)[A, B]A^m$ .

Пусть теперь  $m > n > 0$ . Домножив равенство  $[A^n, B] = n[A, B]A^{n-1}$  справа на  $mA^{m-n}$ , получим  $m[A^n, B]A^{m-n} = mn[A, B]A^{m-1} = n[A^m, B]$ .

**2.8.** Оператору  $\text{ad}_A$  в пространстве  $\text{Hom}(V, V)$  соответствует оператор  $L = I \otimes A - A^T \otimes I$  в пространстве  $V^* \otimes V$  (см. §1.8). Если оператор  $A$  диагонален в



базисе  $e_1, \dots, e_n$ , то оператор  $L$  диагонален в базисе  $e_i^* \otimes e_j$ . Поэтому  $\text{Ker } L^n = \text{Ker } L$ .

**2.9.** а) Если  $\text{tr } Z = 0$ , то  $Z = [X, Y]$  (см. §7.2.2), поэтому  $\text{tr}(AZ) = \text{tr}(AXY) - \text{tr}(AYX) = 0$ . Следовательно,  $A = \lambda I$  (см. задачу 1.1).

б) Для любой линейной функции  $f$  на пространстве матриц существует такая матрица  $A$ , что  $f(X) = \text{tr}(AX)$ . А так как  $f(XY) = f(YX)$ , то  $\text{tr}(AXY) = \text{tr}(AYX)$ . Следовательно,  $A = \lambda I$ .

**3.1.** Ясно, что оператор  $B = \lambda A$  является репликой  $A$ . Докажем, что для нильпотентного оператора  $A$  любая реплика имеет вид  $\lambda A$ . Выберем натуральное число  $r$  так, что  $A^{r-1} \neq 0$  и  $A^r = 0$ . Тогда матрицы  $A^k$ ,  $0 \leq k \leq r-1$ , линейно независимы и матрицы  $A^i \otimes A^j$ ,  $0 \leq i, j \leq r-1$ , тоже линейно независимы.

Поскольку  $A_{2,0} = A \otimes I + I \otimes A$  и  $A^r = 0$ , получаем

$$A_{2,0}^m = \sum_{i=0}^m \binom{m}{i} A^i \otimes A^{m-i} = \begin{cases} \sum_{i=0}^m \binom{m}{i} A^i \otimes A^{m-i} & \text{при } m < r; \\ \sum_{i=m-r+1}^{r-1} \binom{m}{i} A^i \otimes A^{m-i} & \text{при } r \leq m \leq 2r-2; \\ 0 & \text{при } m \geq 2r-1. \end{cases}$$

Если  $A \rightarrow B$ , то  $A_{2,0} \rightarrow B_{2,0}$ , поэтому согласно теореме 7.3.1.1

$$B = \alpha_1 A + \alpha_2 A^2 + \dots + \alpha_{r-1} A^{r-1}, \quad (7.1)$$

$$B_{2,0} = \beta_1 A_{2,0} + \beta_2 A_{2,0}^2 + \dots + \beta_{2r-2} A_{2,0}^{2r-2}. \quad (7.2)$$

Из равенства (7.1) следует, что

$$B_{2,0} = B \otimes I + I \otimes B = \sum_{i=1}^{r-1} \alpha_i (A^i \otimes I + I \otimes A^i).$$

С другой стороны, если мы подставим выражение для  $A_{2,0}^m$  в (7.2), то получим

$$B_{2,0} = \sum_{m=1}^{r-1} \beta_m \left( \sum_{i=1}^m \binom{m}{i} A^i \otimes A^{m-i} \right) + \sum_{m=r}^{2r-2} \beta_m \left( \sum_{i=m-r+1}^{r-1} \binom{m}{i} A^i \otimes A^{m-i} \right).$$

Сравнивая эти два выражения для  $B_{2,0}$ , получаем

$$\alpha_1 = \beta_1, \dots, \alpha_{r-1} = \beta_{r-1}, \beta_r = 0, \dots, \beta_{2r-2} = 0.$$

Более того, если  $\beta_m \neq 0$  для некоторого  $m \geq 2$ , то мы получаем равенства  $\binom{m}{1} = \dots = \binom{m}{m-1} = 0$ , чего не может быть, так как  $\binom{m}{1} = m$ . Следовательно,  $\alpha_m = \beta_m = 0$  при  $m \geq 2$ , т.е.  $B = \alpha_1 A$ .

## Глава 8.

# Матрицы в алгебре и анализе

### 8.1. Кватернионы и числа Кэли. Алгебры Клиффорда

#### 8.1.1. Удвоение алгебры

Пусть  $\mathcal{A}$  — алгебра с единицей над  $\mathbb{R}$ , в которой задана операция сопряжения  $a \mapsto \bar{a}$ , обладающая следующими свойствами:  $\bar{\bar{a}} = a$  и  $\overline{ab} = \bar{b}\bar{a}$ .

Рассмотрим пространство  $\mathcal{A} \oplus \mathcal{A} = \{(a, b) \mid a, b \in \mathcal{A}\}$  и зададим в нём умножение по формуле  $(a, b)(u, v) = (au - \bar{v}b, b\bar{u} + va)$ . Полученную алгебра называют *удвоением алгебры*  $\mathcal{A}$ . Эта конструкция интересна тем, что алгебра комплексных чисел  $\mathbb{C}$  является удвоением алгебры  $\mathbb{R}$ , алгебра кватернионов  $\mathbb{H}$  является удвоением алгебры  $\mathbb{C}$ , а алгебра Кэли  $\mathbb{O}$  является удвоением алгебры  $\mathbb{H}$ .

Легко проверить, что элемент  $(1, 0)$  является двусторонней единицей. Пусть  $e = (0, 1)$ . Тогда  $(b, 0)e = (0, b)$ , поэтому, отождествив элемент  $x$  алгебры  $\mathcal{A}$  с элементом  $(x, 0)$  удвоения алгебры  $\mathcal{A}$ , получим разложение  $(a, b) = a + be$ .

В удвоении алгебры  $\mathcal{A}$  операцию сопряжения можно задать формулой  $\overline{(a, b)} = (\bar{a}, -b)$  т.е.  $\overline{a + be} = \bar{a} - be$ . Если  $x = a + be$  и  $y = u + ve$ , то  $\overline{xy} = \overline{au + (be)u + a(ve) + (be)(ve)} = \bar{u}\bar{a} - \bar{u}(be) - (ve)\bar{a} + (ve)(be) = \bar{y}\bar{x}$ .

Легко проверить, что  $ea = \bar{a}e$  и  $a(be) = (ba)e$  для  $a, b \in \mathcal{A}$ . Поэтому удвоение алгебры  $\mathcal{A}$  некоммутативно, если сопряжение в  $\mathcal{A}$  не тождественно, и неассоциативно, если алгебра  $\mathcal{A}$  некоммутативна. Если же

алгебра  $\mathcal{A}$  коммутативна и ассоциативна, то её удвоение ассоциативно.

### 8.1.2. Кватернионы

Так как  $(0, 1)(0, 1) = (-1, 0)$ , то  $e^2 = -1$ , поэтому удвоением алгебры  $\mathbb{R}$  с тождественным сопряжением является алгебра  $\mathbb{C}$ . Рассмотрим удвоение алгебры  $\mathbb{C}$  со стандартным сопряжением. Любой элемент полученной алгебры можно записать в виде  $q = a + be$ , где  $a = a_0 + a_1i$ ,  $b = a_2 + a_3i$  и  $a_0, \dots, a_3 \in \mathbb{R}$ . Введя обозначения  $j = e$  и  $k = ie$ , получим обычную запись кватерниона  $q = a_0 + a_1i + a_2j + a_3k$ . Число  $a_0$  называют *вещественной частью* кватерниона  $q$ , а кватернион  $a_1i + a_2j + a_3k$  называют его *мнимой частью*. Кватернион называют *вещественным*, если  $a_1 = a_2 = a_3 = 0$ , и *чисто мнимым*, если  $a_0 = 0$ .

Умножение в алгебре кватернионов  $\mathbb{H}$  задаётся следующими равенствами:  $i^2 = j^2 = k^2 = -1$ ,  $ij = -ji = k$ ,  $jk = -kj = i$ ,  $ki = -ik = j$ .

Алгебра кватернионов является удвоением ассоциативной и коммутативной алгебры, поэтому она ассоциативна.

Сопряжённый с кватернионом  $q = a + be$  кватернион  $\bar{q}$  равен  $\bar{a} - be = a_0 - a_1i - a_2j - a_3k$ . В §8.1.1 было показано, что  $\overline{\bar{q}_1 \bar{q}_2} = \bar{q}_2 \bar{q}_1$ .

**Теорема 8.1.2.1.** *Скалярное произведение  $(q, r)$  кватернионов  $q$  и  $r$  равно  $(\bar{q}r + r\bar{q})/2$ , в частности,  $|q|^2 = (q, q) = \bar{q}q$ .*

**Доказательство.** Функция  $B(q, r) = (\bar{q}r + r\bar{q})/2$  симметрична и билинейна. Поэтому достаточно проверить, что  $B(q, r) = (q, r)$  для элементов базиса. Легко проверить, что  $B(1, i) = 0$ ,  $B(i, i) = 1$  и  $B(i, j) = 0$ ; остальные равенства проверяются аналогично.  $\square$

**Следствие.** *Элемент  $\bar{q}/|q|^2$  является двусторонним обратным для элемента  $q$ .*

В самом деле,  $\bar{q}q = |q|^2 = q\bar{q}$ .

**Теорема 8.1.2.2.**  $|qr| = |q| \cdot |r|$ .

**Доказательство.** Ясно, что  $|qr|^2 = qr\overline{(qr)} = qr\bar{r}\bar{q} = q|r|^2\bar{q} = |q|^2|r|^2$ .  $\square$

**Следствие.** *Если  $q \neq 0$  и  $r \neq 0$ , то  $qr \neq 0$ .*

### 8.1.3. Кватернионы и простейшие алгебры Ли

Любому кватерниону  $q = \alpha + xi + yj + zk$  можно сопоставить матрицу  $C(q) = \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$ , где  $u = \alpha + ix$  и  $v = y + iz$ . При этом  $C(qr) = C(q)C(r)$  (см. задачу 1.5).

Мнимому кватерниону  $q = xi + yj + zk$  можно сопоставить матрицу  $R(q) = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$ . Произведение мнимых кватернионов может иметь ненулевую вещественную часть, поэтому матрица  $R(qr)$  определена не для всех  $q$  и  $r$ . Но легко проверить, что  $R(qr - rq) = R(q)R(r) - R(r)R(q)$ , т.е. *векторное произведение*  $q \times r = (qr - rq)/2$  соответствует коммутатору трёхмерных кососимметрических матриц.

Линейное подпространство в пространстве матриц называют *матричной алгеброй Ли*, если оно вместе с любыми двумя матрицами  $A$  и  $B$  содержит и их коммутатор  $[A, B]$ . Легко проверить, что матричными алгебрами Ли являются множество вещественных кососимметрических матриц и множество комплексных косоэрмитовых матриц; эти алгебры Ли обозначают  $\mathfrak{so}(n, \mathbb{R})$  и  $\mathfrak{su}(n)$  соответственно.

**Теорема 8.1.3.1.** *Алгебры Ли  $\mathfrak{so}(3, \mathbb{R})$  и  $\mathfrak{su}(2)$  изоморфны.*

**Доказательство.** Обе эти алгебры, как было показано выше, изоморфны алгебре чисто мнимых кватернионов с коммутатором  $[q, r] = (qr - rq)/2$ .  $\square$

**Теорема 8.1.3.2.** *Алгебры Ли  $\mathfrak{so}(4, \mathbb{R})$  и  $\mathfrak{so}(3, \mathbb{R}) + \mathfrak{so}(3, \mathbb{R})$  изоморфны.*

**Доказательство.** Алгебру Ли  $\mathfrak{so}(3, \mathbb{R})$  можно отождествить с алгеброй Ли чисто мнимых кватернионов. Сопоставим кватерниону  $q \in \mathfrak{so}(3, \mathbb{R})$  преобразование  $P(q): u \mapsto qu$  пространства  $\mathbb{R}^4 = \mathbb{H}$ . Легко проверить, что

$$P(xi + yj + zk) = \begin{pmatrix} 0 & -x & -y & -z \\ x & 0 & -z & y \\ y & z & 0 & -x \\ z & -y & x & 0 \end{pmatrix} \in \mathfrak{so}(4, \mathbb{R}).$$

Аналогично отображение  $Q(q): u \mapsto u\bar{q}$  принадлежит  $\mathfrak{so}(4, \mathbb{R})$ . Легко проверить, что отображения  $q \mapsto P(q)$  и  $q \mapsto Q(q)$  являются гомоморфизмами алгебр Ли, т.е.  $P(qr - rq) = P(q)P(r) - P(r)P(q)$  и  $Q(qr - rq) = Q(q)Q(r) - Q(r)Q(q)$ . Поэтому отображение  $(q, r) \mapsto P(q) + Q(r)$  является гомоморфизмом алгебры Ли  $\mathfrak{so}(3, \mathbb{R}) + \mathfrak{so}(3, \mathbb{R})$  в  $\mathfrak{so}(4, \mathbb{R})$ . Размерности

этих алгебр совпадают, поэтому достаточно проверить, что рассматриваемое отображение мономорфно. Равенство  $P(q) + Q(r) = 0$  означает, что  $qx + x\bar{r} = 0$  для всех  $x$ . При  $x = 1$  получаем  $q = -\bar{r}$ , а значит,  $qx - xq = 0$  для всех  $x$ . Следовательно,  $q$  — вещественный кватернион (см. задачу 1.4); с другой стороны, по определению  $q$  — мнимый кватернион. Поэтому  $q = r = 0$ .  $\square$

#### 8.1.4. Кватернионы и движения

С помощью кватернионов можно интерпретировать движения 3-мерного и 4-мерного пространства. При этом мы рассматриваем 3-мерное пространство как пространство чисто мнимых кватернионов, а 4-мерное пространство как пространство всех кватернионов.

Для каждого ненулевого кватерниона  $s$  можно рассмотреть отображения  $f_s(q) = -sq s^{-1}$  и  $h_s(q) = sq s^{-1}$ . Легко проверить, что оба отображения  $f_s$  и  $h_s$  переводят в себя пространство чисто мнимых кватернионов. Действительно, кватернион  $q$  чисто мнимый тогда и только тогда, когда  $q^2 < 0$  (задача 1.2). Поэтому достаточно проверить, что  $(sq s^{-1})^2 < 0$ . Ясно, что  $sq s^{-1}sq s^{-1} = sq^2 s^{-1} = q^2 ss^{-1} = q^2 < 0$ . Легко также проверить, что оба отображения  $f_s$  и  $h_s$  являются ортогональными преобразованиями пространства чисто мнимых кватернионов. Действительно,  $\|sq s^{-1}\| = \|s\| \cdot \|q\| \cdot \|s\|^{-1} = \|q\|$ .

**Теорема 8.1.4.1.** *Для чисто мнимого кватерниона  $s$  отображение  $f_s$  (пространства чисто мнимых кватернионов) является симметрией относительно плоскости, перпендикулярной  $s$ , а отображение  $h_s$  является симметрией относительно оси, содержащей  $s$ .*

**Доказательство.** Ясно, что  $f_s(s) = -sss^{-1} = -s$  и  $h_s(s) = sss^{-1} = s$ . Поэтому остаётся проверить, что если  $q \perp s$ , то  $f_s(q) = q$  и  $h_s(q) = -q$ . Равенство  $(q, s) = \frac{1}{2}(\bar{q}s + \bar{s}q) = -\frac{1}{2}(qs + sq)$  показывает, что если  $(q, s) = 0$ , то  $sq = -qs$ . Поэтому в таком случае  $f_s(q) = -sq s^{-1} = qss^{-1} = q$  и  $h_s(q) = sq s^{-1} = -qss^{-1} = -q$ .  $\square$

**Следствие.** *Для любых двух чисто мнимых кватернионов  $q_1$  и  $q_2$ , имеющих одинаковую длину, можно выбрать чисто мнимый кватернион  $s$  так, что  $q_2 = sq_1 s^{-1}$ .*

**Доказательство.** Пусть  $s$  — вектор, идущий из начала координат в середину отрезка, соединяющего концы векторов  $q_1$  и  $q_2$ . Тогда при симметрии относительно оси, содержащей вектор  $s$ , вектор  $q_1$  переходит в  $q_2$ .  $\square$

**Теорема 8.1.4.2.** Для кватерниона  $s = a + r$ , где число  $a$  вещественное, а  $r \neq 0$  — чисто мнимый кватернион, отображение  $h_s$  пространства чисто мнимых кватернионов является поворотом вокруг оси, содержащей  $r$ , причём угол поворота  $\varphi$  находится из соотношения  $\operatorname{tg}(\varphi/2) = \|r/a\|$  (если  $a = 0$ , то  $\varphi = \pi$ ).

**Доказательство.** Кватернион  $r$  коммутирует с  $s$ , поэтому  $f_s(r) = r$ . Определитель отображения  $f_s$  непрерывно зависит от  $s$ , и топологическое пространство кватернионов, отличных от нуля, линейно связно (т.е. любые два ненулевых кватерниона можно соединить путём, состоящим из ненулевых кватернионов). Поэтому  $\det(f_s) = 1$ , а значит, отображение  $f_s$  является поворотом вокруг оси, содержащей  $r$ . Остаётся вычислить угол этого поворота.

Для любого ненулевого кватерниона  $q$  отображения  $f_s$  и  $f_{sq^{-1}} = f_q f_s (f_q)^{-1}$  являются поворотами на один и тот же угол, поэтому вместо  $s$  можно взять любой кватернион  $sq^{-1} = a + qrq^{-1}$ . Согласно следствию теоремы 8.1.4.1 кватернион  $r$  сопряжён с кватернионом  $bi$ , где  $b = \|r\|$ . Поэтому вместо кватерниона  $s$  можно взять кватернион  $a + bi$ . Таким образом, остаётся вычислить угол  $\varphi$  между векторами  $j$  и

$$\begin{aligned} f_{a+bi}(j) &= (a + bi)j(a + bi)^{-1} = (aj + bk) \frac{a - bi}{a^2 + b^2} = \\ &= \frac{a^2 j + 2abk - b^2 j}{a^2 + b^2} = \frac{a^2 - b^2}{a^2 + b^2} j + \frac{2ab}{a^2 + b^2} k. \end{aligned}$$

Ясно, что  $\operatorname{tg} \varphi = \frac{2ab}{a^2 - b^2}$ . Если мы будем считать, что  $0 < \varphi < \pi$ , то  $\operatorname{tg}(\varphi/2) = |b/a| = \|r/a\|$ .  $\square$

Для кватернионов  $s$  и  $r$ , имеющих единичную длину, можно рассмотреть отображение  $f_{s,r}(q) = sq\bar{r}$  пространства всех кватернионов в себя. Легко проверить, что это отображение является ортогональным преобразованием. Кроме того, сопоставляя паре  $(s, r) \in S^3 \times S^3$  ортогональное преобразование  $f_{s,r}$ , мы получаем гомоморфизм групп  $f: S^3 \times S^3 \rightarrow \operatorname{SO}(4)$ ; здесь  $S^3$  — группа кватернионов единичной длины, а  $\operatorname{SO}(4)$  — группа ортогональных матриц с определителем 1. Действительно,  $s_1 s_2 q \overline{r_1 r_2} = s_1 (s_2 q \bar{r}_2) \bar{r}_1$ .

**Теорема 8.1.4.3.** Гомоморфизм  $f$  является эпиморфизмом, причём его ядро состоит из элементов  $(1, 1)$  и  $(-1, -1)$ .

**Доказательство.** Ядро отображения  $f$  состоит из таких пар  $(s, r)$ , что  $sx\bar{r} = x$  для всех  $x$ . При  $x = 1$  получаем  $s\bar{r} = 1$ , т.е.  $\bar{r} = s^{-1}$ . Таким

образом, для всех  $x$  выполняется равенство  $sxs^{-1} = x$ , т.е.  $sx = xs$ . Поэтому согласно задаче 1.4 кватернион  $s$  вещественный. По условию его длина равна 1, т.е.  $s = \pm 1$ .

Докажем теперь, что любое ортогональное преобразование  $A \in \text{SO}(4)$  можно представить в виде  $A(x) = sx\bar{r}$  для некоторых единичных кватернионов  $s$  и  $r$ . Выберем сначала кватернион  $q$  так, что  $qA(1) = 1$ , т.е.  $q = A(1)^{-1}$ . Отображение  $x \mapsto qA(x)$  является ортогональным преобразованием с определителем 1 пространства чисто мнимых кватернионов, поэтому его можно представить в виде  $x \mapsto rxr^{-1} = rx\bar{r}$ , где  $r$  — кватернион единичной длины. Итак,  $qA(x) = rx\bar{r}$ , т.е.  $A(x) = (q^{-1}r)x\bar{r}$ .  $\square$

**Замечание.** Несложная проверка показывает, что  $f(S^3 \times 1)$  и  $f(1 \times S^3)$  — нормальные подгруппы в  $\text{SO}(4)$ .

### 8.1.5. Изоморфизм $\mathbb{H} \otimes \mathbb{H}$ и $M_4(\mathbb{R})$

Рассмотрим алгебру кватернионов  $\mathbb{H}$  как линейное пространство над  $\mathbb{R}$ . В пространстве  $\mathbb{H} \otimes \mathbb{H}$  можно ввести структуру алгебры, определив произведение элементов  $x_1 \otimes x_2$  и  $y_1 \otimes y_2$  как  $x_1y_1 \otimes x_2y_2$ . Отождествим пространство  $\mathbb{R}^4$  с  $\mathbb{H}$ . Легко проверить, что отображение  $w: \mathbb{H} \otimes \mathbb{H} \rightarrow M_4(\mathbb{R})$ , определённое формулой  $[w(x_1 \otimes x_2)]x = x_1x\bar{x}_2$ , является гомоморфизмом алгебр, т.е.  $w(uv) = w(u)w(v)$ .

**Теорема 8.1.5.1.** *Отображение  $w$  является изоморфизмом алгебр  $\mathbb{H} \otimes \mathbb{H}$  и  $M_4(\mathbb{R})$ .*

**Доказательство.** Размерности алгебр  $\mathbb{H} \otimes \mathbb{H}$  и  $M_4(\mathbb{R})$  равны, но вычисление ядра отображения  $w$  совсем не так просто, как вычисление ядра отображения  $(q, r) \mapsto P(q) + Q(r)$ , потому что пространство  $\mathbb{H} \otimes \mathbb{H}$  содержит не только элементы вида  $x \otimes y$ . Вместо вычисления ядра мы докажем, что образ отображения  $w$  совпадает с  $M_4(\mathbb{R})$ . Матрицы

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \varepsilon = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, a = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

образуют базис пространства матриц порядка 2. Образы элементов  $x \otimes y$ , где  $x, y \in \{1, i, j, k\}$ , при отображении  $w$  приведены в таблице 1. Из этой таблицы видно, что среди линейных комбинаций пар образов этих элементов встречаются все матрицы, три блока которых нулевые, а четвёртым блоком является одна из матриц  $e, \varepsilon, a$  и  $b$ . Среди линейных комбинаций четвёрок этих матриц встречаются все матрицы, содержащие

ровно один ненулевой элемент, причём этот элемент равен 1. Такие матрицы образуют базис пространства  $M_4(\mathbb{R})$ .  $\square$

Таблица 1

$y$	1	$i$	$j$	$k$
$x$				
1	$\begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}$	$\begin{pmatrix} b & 0 \\ 0 & -b \end{pmatrix}$	$\begin{pmatrix} 0 & e \\ -e & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}$
$i$	$\begin{pmatrix} -b & 0 \\ 0 & -b \end{pmatrix}$	$\begin{pmatrix} e & 0 \\ 0 & -e \end{pmatrix}$	$\begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & e \\ e & 0 \end{pmatrix}$
$j$	$\begin{pmatrix} 0 & -\varepsilon \\ \varepsilon & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}$	$\begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$	$\begin{pmatrix} -a & 0 \\ 0 & a \end{pmatrix}$
$k$	$\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -\varepsilon \\ -\varepsilon & 0 \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$	$\begin{pmatrix} \varepsilon & 0 \\ 0 & -\varepsilon \end{pmatrix}$

### 8.1.6. Числа Кэли

Удвоением алгебры кватернионов с естественной операцией сопряжения является *алгебра Кэли* (*алгебра октав*). Октавы открыл Джон Грейвз в декабре 1843 г., после того как в октябре он получил письмо от Гамильтона, сообщившего о своём открытии кватернионов. Но Грейвз опубликовал свою работу лишь в 1848 г. За это время октавы переоткрыл Кэли в 1845 г.

Базисом алгебры Кэли как пространства над  $\mathbb{R}$  служат элементы  $1, i, j, k, e, f = ie, g = je$  и  $h = ke$ . Таблицу умножения базисных элементов удобно представить с помощью рис. 8.1. Произведением двух элементов, лежащих на одной прямой или одной окружности, является третий элемент, лежащий на той же прямой или окружности. Выбор знака определяется ориентацией; например,  $ie = f, if = -e$ .

Пусть  $\xi = a + be$ , где  $a$  и  $b$  — кватернионы. Сопряжение в алгебре Кэли задаётся формулой  $\overline{(a, b)} = (\bar{a}, -b)$ , т.е.  $\overline{a + be} = \bar{a} - be$ . Ясно, что  $\xi \bar{\xi} = (a, b)(\bar{a}, -b) = (a, b)(\bar{a}, -b) = (\bar{a}a + \bar{b}b, ba - ab) = \bar{a}a + \bar{b}b = a\bar{a} + b\bar{b}$ , т.е.  $\xi \bar{\xi}$  — сумма квадратов координат вектора  $\xi$ . Поэтому  $|\xi| = \sqrt{\xi \bar{\xi}} = \sqrt{\bar{\xi} \xi}$  — длина вектора  $\xi$ .

**Теорема 8.1.6.1.**  $|\xi\eta| = |\xi| \cdot |\eta|$ .

**Доказательство.** Пусть  $\xi = a + be$  и  $\eta = u + ve$ , где  $a, b, u, v$  — кватернионы. Так как  $(be)u = (0, b)(u, 0) = (0, b\bar{u}) = (b\bar{u})e$  и  $(be)(ve) = (0, b)(0, v) =$



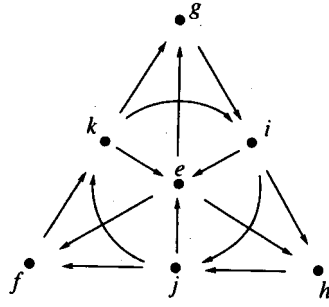


Рис. 8.1.

$(-\bar{v}b, 0) = -\bar{v}b$ , то

$$|\xi\eta|^2 = (au - \bar{v}b)(\bar{u}\bar{a} - \bar{b}\bar{v}) + (b\bar{u} + va)(u\bar{b} + \bar{a}\bar{v}).$$

Запишем кватернион  $v$  в виде  $v = \lambda + v_1$ , где  $\lambda$  — действительное число и  $\bar{v}_1 = -v_1$ . Тогда

$$\begin{aligned} |\xi\eta|^2 &= (au - \lambda b + v_1 b)(\bar{u}\bar{a} - \lambda\bar{b} - \bar{b}v_1) + \\ &+ (b\bar{u} + \lambda a + v_1 a)(u\bar{b} + \lambda\bar{a} - \bar{a}v_1). \end{aligned}$$

Кроме того,  $|\xi|^2|\eta|^2 = (a\bar{a} + b\bar{b})(u\bar{u} + \lambda^2 - v_1v_1)$ . Так как числа  $u\bar{u}$  и  $b\bar{b}$  действительные, то  $au\bar{u}\bar{a} = a\bar{a}u\bar{u}$  и  $b\bar{b}v_1 = v_1b\bar{b}$ . Используя аналогичные равенства, получаем  $|\xi\eta|^2 - |\xi|^2|\eta|^2 = v_1(b\bar{u}\bar{a} + au\bar{b}) - (au\bar{b} + b\bar{u}\bar{a})v_1 = 0$ , так как  $b\bar{u}\bar{a} + au\bar{b}$  — действительное число.  $\square$

**Следствие 1.** Если  $\xi \neq 0$ , то  $\bar{\xi}/|\xi|^2$  — двусторонний обратный элемент для  $\xi$ .

**Следствие 2.** Если  $\xi \neq 0$  и  $\eta \neq 0$ , то  $\xi\eta \neq 0$ .

Алгебра кватернионов некоммутативна, поэтому алгебра Кэли неассоциативна. Но для элементов алгебры Кэли справедливы равенства  $x(xy) = (xx)y$ ,  $y(xx) = (yx)x$  и  $(xy)x = x(yx)$  (см. задачу 1.8). Можно доказать, что любая подалгебра алгебры Кэли, порождённая двумя элементами, ассоциативна.

### 8.1.7. Векторное произведение чисел Кэли

По аналогии с векторным произведением в пространстве чисто мнимых кватернионов можно определить векторное произведение в семи-мерном пространстве чисто мнимых чисел Кэли. Пусть  $x$  и  $y$  — чисто мнимые числа Кэли. Их *векторным произведением* называют мнимую часть  $xy$ ; его обозначают  $x \times y$ . Ясно, что  $x \times y = (xy - \overline{xy})/2 = (xy - yx)/2$ .

Можно проверить, что скалярное произведение чисел Кэли  $x$  и  $y$  равно  $(\overline{x}y + y\overline{x})/2$ ; для чисто мнимых чисел Кэли  $x$  и  $y$  получаем  $(x, y) = -(xy + yx)/2$ .

**Теорема 8.1.7.1.** *Векторное произведение чисто мнимых чисел Кэли обладает следующими свойствами:*

- а)  $x \times y \perp x$ ,  $x \times y \perp y$ ;
- б)  $|x \times y|^2 = |x|^2|y|^2 - |(x, y)|^2$ .

**Доказательство.** а) Нужно доказать, что  $x(xy - yx) + (xy - yx)x = 0$ . Так как  $x(yx) = (xy)x$  (см. задачу 1.8, б), приходим к равенству  $x(xy) = (yx)x$ . Согласно задаче 1.8, а)  $x(xy) = (xx)y$  и  $(yx)x = y(xx)$ . Остаётся заметить, что  $xx = -x\overline{x} = -(x, x)$  — действительное число.

б) Нужно доказать, что  $-(xy - yx)(xy - yx) = 4|x|^2|y|^2 - (xy + yx)(xy + yx)$ , т.е.  $2|x|^2|y|^2 = (xy)(yx) + (yx)(xy)$ . Пусть  $a = xy$ . Тогда  $\bar{a} = yx$  и  $2|x|^2|y|^2 = 2(a, a) = a\bar{a} + \bar{a}a = (xy)(yx) + (yx)(xy)$ .  $\square$

### 8.1.8. Антicomмутирующие комплексные структуры

Оставшаяся часть этого параграфа будет посвящена решению следующей задачи: «Чему равно наибольшее количество ортогональных операторов  $A_1, \dots, A_m$  в пространстве  $\mathbb{R}^n$ , удовлетворяющих соотношениям  $A_i^2 = -I$  и  $A_i A_j + A_j A_i = 0$  при  $i \neq j$ ?» Эта задача может показаться весьма неестественной. К ней, однако, сводятся многие важные задачи, так или иначе связанные с алгеброй кватернионов и алгеброй Кэли (заметим, что операторы умножения на  $i, j, \dots, h$  удовлетворяют требуемым соотношениям). Мы сначала сформулируем ответ, а затем расскажем, какие задачи сводятся к нашей задаче.

**Теорема 8.1.8.1.** *Запишем число  $n$  в виде  $n = (2a + 1)2^b$ , где  $b = c + 4d$  и  $0 \leq c \leq 3$ . Пусть  $\rho(n) = 2^c + 8d$ . Тогда наибольшее количество требуемых операторов в  $\mathbb{R}^n$  равно  $\rho(n) - 1$ .*

**Произведение квадратичных форм**

Пусть  $a = x_1 + ix_2$  и  $b = y_1 + iy_2$ . Тогда равенство  $|a|^2|b|^2 = |ab|^2$  можно переписать в виде

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = z_1^2 + z_2^2,$$

где  $z_1 = x_1y_1 - x_2y_2$  и  $z_2 = x_1y_2 + x_2y_1$ . Аналогичные тождества можно записать для алгебры кватернионов и для алгебры Кэли.

**Теорема 8.1.8.2.** *Тождество вида*

$$(x_1^2 + \dots + x_m^2)(y_1^2 + \dots + y_n^2) = z_1^2 + \dots + z_n^2,$$

где  $z_i$  — вещественные билинейные функции от  $x$  и  $y$ , существует тогда и только тогда, когда  $m \leq n$ .

**Доказательство.** Пусть  $z_i = \sum_j b_{ij}(x)y_j$ , где  $b_{ij}(x)$  — линейные функции.

Тогда

$$z_i^2 = \sum_j b_{ij}^2(x)y_j^2 + 2 \sum_{j < k} b_{ij}(x)b_{ik}(x)y_jy_k.$$

Поэтому  $\sum_i b_{ij} = x_1^2 + \dots + x_m^2$  и  $\sum_{j < k} b_{ij}(x)b_{ik}(x) = 0$ . Пусть  $B(x) = \|b_{ij}(x)\|_1^n$ . Тогда  $B^T(x)B(x) = (x_1^2 + \dots + x_m^2)I$ . Матрицу  $B(x)$  можно записать в виде  $B(x) = x_1B_1 + \dots + x_mB_m$ . Поэтому  $B^T(x)B(x) = x_1^2B_1^TB_1 + \dots + x_m^2B_m^TB_m + \sum_{i < j} (B_i^TB_j + B_j^TB_i)x_ix_j$ , а значит,  $B_i^TB_i = I$  и  $B_i^TB_j + B_j^TB_i = 0$ .

Операторы  $B_i$  ортогональны и  $B_i^{-1}B_j = -B_j^{-1}B_i$  при  $i \neq j$ .

Рассмотрим ортогональные операторы  $A_1, \dots, A_{m-1}$ , где  $A_i = B_m^{-1}B_i$ . Тогда  $B_m^{-1}B_i = -B_i^{-1}B_m$ , а значит,  $A_i = -A_i^{-1}$ , т.е.  $A_i^2 = -I$ . Кроме того,  $B_i^{-1}B_j = -B_j^{-1}B_i$  при  $i \neq j$ , а значит,  $A_iA_j = B_m^{-1}B_iB_m^{-1}B_j = -B_i^{-1}B_mB_m^{-1}B_j = B_j^{-1}B_i = -A_jA_i$ .

Легко проверить также, что если ортогональные операторы  $A_1, \dots, A_{m-1}$  таковы, что  $A_i^2 = I$  и  $A_iA_j + A_jA_i = 0$ , то операторы  $B_1 = A_1, \dots, B_{m-1} = A_{m-1}, B_m = I$  обладают требуемыми свойствами. Для завершения доказательства теоремы 8.1.8.2 остаётся воспользоваться теоремой 8.1.8.1.  $\square$

**Нормированные алгебры**

**Теорема 8.1.8.3.** *Пусть алгебра  $\mathcal{A}$  над  $\mathbb{R}$  является евклидовым пространством, причём  $\|xy\| = \|x\| \cdot \|y\|$  для любых  $x, y \in \mathcal{A}$ . Тогда размерность  $\mathcal{A}$  равна 1, 2, 4 или 8.*

**Доказательство.** Пусть  $e_1, \dots, e_n$  — ортонормированный базис алгебры  $A$ . Тогда

$$(x_1 e_1 + \dots + x_n e_n)(y_1 e_1 + \dots + y_n e_n) = z_1 e_1 + \dots + z_n e_n,$$

где  $z_1, \dots, z_n$  — билинейные функции от  $x$  и  $y$ . Из равенства  $\|z\|^2 = \|x\|^2 \|y\|^2$  следует, что

$$(x_1^2 + \dots + x_n^2)(y_1^2 + \dots + y_n^2) = z_1^2 + \dots + z_n^2.$$

Остаётся воспользоваться теоремой 8.1.8.2 и заметить, что  $\rho(n) = n$  тогда и только тогда, когда  $n = 1, 2, 4$  или  $8$ .  $\square$

### Векторные произведения

**Теорема 8.1.8.4.** Пусть в пространстве  $\mathbb{R}^n$ , где  $n \geq 3$ , определена билинейная операция  $f(v, w) = v \times w \in \mathbb{R}^n$ , обладающая следующими свойствами: векторы  $v$  и  $w$  перпендикулярны вектору  $v \times w$  и  $\|v \times w\|^2 = \|v\|^2 \|w\|^2 - (v, w)^2$ . Тогда  $n = 3$  или  $7$ .

**Доказательство.** [Ма5] Рассмотрим пространство  $\mathbb{R}^{n+1} = \mathbb{R} \oplus \mathbb{R}^n$  и зададим в нём умножение с помощью формулы

$$(a, v)(b, w) = (ab - (v, w), aw + bv + v \times w),$$

где  $(v, w)$  — скалярное произведение в  $\mathbb{R}^n$ . Легко проверить, что в полученной алгебре размерности  $n + 1$  справедливо тождество  $\|xy\|^2 = \|x\|^2 \|y\|^2$ . Остаётся воспользоваться теоремой 8.1.8.3.  $\square$

**Замечание.** В пространствах размерности 3 и 7 билинейная операция, обладающая указанными свойствами, существует (см. §8.1.7).

### Векторные поля на сферах

Векторным полем на сфере  $S^n = \{v \in \mathbb{R}^{n+1} \mid \|v\| = 1\}$  называют отображение, сопоставляющее каждой точке  $v \in S^n$  вектор  $F(v)$ , лежащий в касательном пространстве к  $S^n$  в точке  $v$ . Касательное пространство к  $S^n$  в точке  $v$  состоит из векторов, перпендикулярных  $v$ , поэтому  $F(v) \perp v$ . Векторное поле называют *линейным*, если  $F(v) = Av$  для некоторого линейного оператора  $A$ . Легко проверить, что  $Av \perp v$  при всех  $v$  тогда и только тогда, когда  $A$  — кососимметрический оператор (теорема 4.3.1.2). Поэтому на сфере  $S^{2n}$  любое линейное векторное поле обращается в нуль в некоторой точке.

Чтобы исключить векторные поля, обращающиеся в нуль, будем рассматривать только ортогональные операторы; в этом случае  $\|Av\| = 1$ . Легко проверить, что ортогональный оператор  $A$  кососимметричен тогда и только тогда, когда  $A^2 = -I$ . Напомним, что оператор, квадрат которого равен  $-I$ , называют комплексной структурой (см. §7.6).

Векторные поля  $F_1, \dots, F_m$  называют линейно независимыми, если векторы  $F_1(v), \dots, F_m(v)$  линейно независимы в каждой точке  $v$ . В частности, линейно независимы векторные поля, соответствующие ортогональным операторам  $A_1, \dots, A_m$ , для которых  $A_i v \perp A_j v$  при всех  $i \neq j$ . Равенство  $(A_i v, A_j v) = 0$  означает, что  $(v, A_i^T A_j v) = 0$ . Поэтому  $A_i^T A_j + (A_i^T A_j)^T = 0$ , т.е.  $A_i A_j + A_j A_i = 0$ .

Итак, для построения  $m$  линейно независимых векторных полей на сфере  $S^n$  достаточно указать ортогональные операторы  $A_1, \dots, A_m$  в  $(n+1)$ -мерном пространстве, удовлетворяющие соотношениям  $A_i^2 = -I$  и  $A_i A_j + A_j A_i = 0$  при  $i \neq j$ . Таким образом, доказано следующее утверждение.

**Теорема 8.1.8.5.** *На сфере  $S^{n-1}$  существует  $\rho(n) - 1$  линейно независимых векторных полей.*

**Замечание.** Гораздо сложнее доказывается теорема о том, что на сфере  $S^{n-1}$  не существует  $\rho(n)$  непрерывных линейно независимых векторных полей (см. [Ad]).

#### Линейные подпространства в пространстве матриц

**Теорема 8.1.8.6.** *В пространстве матриц порядка  $n$  над полем действительных чисел есть подпространство размерности  $\rho(n)$ , все ненулевые матрицы которого невырождены.*

**Доказательство.** Добавим оператор  $A_0 = I$  к операторам  $A_1, \dots, A_{\rho(n)-1}$ . При доказательстве теоремы 8.1.8.5 было показано, что  $(A_i v, A_j v) = 0$  при  $i \neq j$ . Поэтому  $(\sum \lambda_i A_i v, \sum \lambda_j A_j v) = \sum \lambda_i^2 \|v\|^2$ . Следовательно, оператор  $\sum \lambda_i A_i$  невырожден, если хотя бы одно из чисел  $\lambda_i$  отлично от нуля.  $\square$

#### 8.1.9. Доказательство теоремы о комплексных структурах

Перейдём теперь к доказательству теоремы 8.1.8.1. Рассмотрим алгебру  $C_m$ , заданную образующими  $e_1, \dots, e_m$  и соотношениями  $e_i^2 = -1$  и  $e_i e_j + e_j e_i = 0$  при  $i \neq j$ . Любому набору ортогональных матриц  $A_1,$

...,  $A_m$ , удовлетворяющих соотношениям  $A_i^2 = -I$  и  $A_i A_j + A_j A_i = 0$  при  $i \neq j$ , соответствует представление (см. §8.2.1) алгебры  $C_m$ , при котором элементы  $e_1, \dots, e_m$  переходят в ортогональные матрицы. Чтобы изучить строение алгебры  $C_m$ , введём вспомогательную алгебру  $C'_m$ , заданную образующими  $\varepsilon_1, \dots, \varepsilon_m$  и соотношениями  $\varepsilon_i^2 = 1$  и  $\varepsilon_i \varepsilon_j + \varepsilon_j \varepsilon_i = 0$  при  $i \neq j$ .

**Замечание.** Алгебры  $C_m$  и  $C'_m$  называют алгебрами Клиффорда.

**Лемма 1.**  $C_1 \cong \mathbb{C}$ ,  $C_2 \cong \mathbb{H}$ ,  $C'_1 \cong \mathbb{R} \oplus \mathbb{R}$  и  $C'_2 \cong M_2(\mathbb{R})$ .

**Доказательство.** Последний изоморфизм задаётся следующим образом: элементы  $\varepsilon_1$  и  $\varepsilon_2$  переходят в  $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  и  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Остальные изоморфизмы очевидны.  $\square$

**Следствие.**  $\mathbb{C} \otimes \mathbb{H} \cong M_2(\mathbb{C})$ .

В самом деле, комплексификации алгебр  $C_2$  и  $C'_2$  изоморфны.

**Лемма 2.**  $C_{k+2} \cong C'_k \otimes C_2$  и  $C'_{k+2} \cong C_k \otimes C'_2$ .

**Доказательство.** Первый изоморфизм задаётся формулами  $f(e_i) = 1 \otimes e_i$  при  $i = 1, 2$  и  $f(e_i) = \varepsilon_{i-2} \otimes e_1 e_2$  при  $i \geq 3$ . Второй изоморфизм задаётся формулами  $g(\varepsilon_i) = 1 \otimes \varepsilon_i$  при  $i = 1, 2$  и  $g(\varepsilon_i) = e_{i-2} \otimes \varepsilon_1 \varepsilon_2$  при  $i \geq 3$ .  $\square$

**Лемма 3.**  $C_{k+4} \cong C_k \otimes M_2(\mathbb{H})$  и  $C'_{k+4} \cong C'_k \otimes M_2(\mathbb{H})$ .

**Доказательство.** Согласно лемме 2  $C_{k+4} = C'_{k+2} \otimes C_2 = C_k \otimes C'_2 \otimes C_2$ . А так как  $C'_2 \otimes C_2 = \mathbb{H} \otimes M_2(\mathbb{R}) = M_2(\mathbb{H})$ , то  $C_{k+4} = C_k \otimes M_2(\mathbb{H})$ . Аналогично  $C'_{k+4} = C'_k \otimes M_2(\mathbb{H})$ .  $\square$

**Лемма 4.**  $C_{k+8} \cong C_k \otimes M_{16}(\mathbb{R})$ .

**Доказательство.** Согласно лемме 3  $C_{k+8} = C_{k+4} \otimes M_2(\mathbb{H}) = C_k \otimes M_2(\mathbb{H}) \otimes M_2(\mathbb{H})$ . А так как  $\mathbb{H} \otimes \mathbb{H} \cong M_4(\mathbb{R})$  (см. §8.1.5), то  $M_2(\mathbb{H}) \otimes M_2(\mathbb{H}) = M_2(M_4(\mathbb{R})) = M_{16}(\mathbb{R})$ .  $\square$

Леммы 1–3 позволяют вычислить  $C_k$  при  $1 \leq k \leq 8$ . Например,  $C_5 = 1 \otimes M_2(\mathbb{H}) = \mathbb{C} \otimes M_2(\mathbb{H}) = M_2(\mathbb{C} \otimes \mathbb{H}) = M_2(M_2(\mathbb{C})) = M_4(\mathbb{C})$ ;  $C_6 = C_2 \otimes M_2(\mathbb{H}) = M_2(\mathbb{H} \otimes \mathbb{H}) = M_8(\mathbb{R})$  и т.д. Результаты вычислений записаны в таблице 2.

Таблица 2

$k$	1	2	3	4
$C_k$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{H} \oplus \mathbb{H}$	$M_2(\mathbb{H})$
$C'_k$	$\mathbb{R} \oplus \mathbb{R}$	$M_2(\mathbb{R})$	$M_2(\mathbb{C})$	$M_2(\mathbb{H})$
$k$	5	6	7	8
$C_k$	$M_4(\mathbb{C})$	$M_8(\mathbb{R})$	$M_8(\mathbb{R}) \oplus M_8(\mathbb{R})$	$M_{16}(\mathbb{R})$
$C'_k$	$M_2(\mathbb{H}) \oplus M_2(\mathbb{H})$	$M_4(\mathbb{H})$	$M_8(\mathbb{C})$	$M_{16}(\mathbb{R})$

Лемма 4 позволяет теперь вычислить  $C_k$  для любого  $k$ . Алгебры  $C_1, \dots, C_8$  имеют естественные представления в пространствах  $\mathbb{C}, \mathbb{H}, \mathbb{H}, \mathbb{H}^2, \mathbb{C}^4, \mathbb{R}^8, \mathbb{R}^8$  и  $\mathbb{R}^{16}$ ; размерности этих пространств над  $\mathbb{R}$  равны 2, 4, 4, 8, 8, 8, 8 и 16. Кроме того, при переходе от алгебры  $C_k$  к алгебре  $C_{k+8}$  размерность пространства естественного представления увеличивается в 16 раз. Простой перебор показывает, что при  $n = 2^k$  наибольшее  $m$ , для которого алгебра  $C_m$  имеет естественное представление в  $\mathbb{R}^n$ , равно  $\rho(n) - 1$ .

Пусть  $f: C_m \rightarrow M_n(\mathbb{R})$  — представление алгебры  $C_m$ , при котором элементы  $e_1, \dots, e_m$  переходят в ортогональные матрицы. Тогда  $f(1 \cdot e_i) = f(1)f(e_i)$ , причём матрица  $f(e_i)$  обратима. Поэтому  $f(1) = f(1 \cdot e_i)f(e_i)^{-1} = I$  — единичная матрица. Алгебра  $C_m$  имеет вид  $M_p(F)$  или  $M_p(F) \otimes M_p(F)$ , где  $F = \mathbb{R}, \mathbb{C}$  или  $\mathbb{H}$ . Поэтому если  $f$  — такое представление алгебры  $C_m$ , что  $f(1) = I$ , то представление  $f$  вполне приводимо и его неприводимые компоненты изоморфны  $F^p$  (см. §8.2.1); значит, его размерность делится на  $p$ . Следовательно, для любого числа  $n$  наибольшее число  $m$ , для которого алгебра  $C_m$  имеет такое представление в  $\mathbb{R}^n$ , что  $f(1) = I$ , равно  $\rho(n) - 1$ .

Непосредственной проверкой можно убедиться, что при изоморфизмах, указанных в леммах 1–4, элементы  $e_1, \dots, e_m$  переходят в ортогональные матрицы.

## Задачи

**1.1.** Докажите, что вещественная часть произведения кватернионов  $x_1i + y_1j + z_1k$  и  $x_2i + y_2j + z_2k$  равна скалярному произведению векторов  $(x_1, y_1, z_1)$  и  $(x_2, y_2, z_2)$ , взятому со знаком минус, а мнимая часть равна векторному произведению этих векторов.

**1.2.** а) Докажите, что кватернион  $q$  чисто мнимый тогда и только тогда, когда  $q^2 \leq 0$ .

б) Докажите, что кватернион  $q$  вещественный тогда и только тогда, когда  $q^2 \geq 0$ .

**1.3.** Найдите все решения уравнения  $q^2 = -1$  в алгебре кватернионов.

**1.4.** Докажите, что кватернион, коммутирующий со всеми чисто мнимыми кватернионами, веществен.

**1.5.** Матрицу  $A$  с кватернионными элементами можно представить в виде  $A = Z_1 + Z_2j$ , где  $Z_1$  и  $Z_2$  — комплексные матрицы. Сопоставим матрице  $A$  матрицу  $A_c = \begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1 \end{pmatrix}$ . Докажите, что  $(AB)_c = A_c B_c$ .

**1.6.** Рассмотрим отображение пространства  $\mathbb{R}^4 = \mathbb{H}$ , переводящее кватернион  $x$  в  $qx$ , где  $q$  — некоторый фиксированный кватернион.

а) Докажите, что это отображение ортогональные векторы переводит в ортогональные.

б) Докажите, что определитель этого отображения равен  $|q|^4$ .

**1.7.** Дан тетраэдр  $ABCD$ . Докажите с помощью кватернионов, что  $AB \cdot CD + BC \cdot AD \geq AC \cdot BD$ .

**1.8.** Пусть  $x$  и  $y$  — элементы алгебры Кэли. Докажите, что

а)  $x(yu) = (xy)u$  и  $x(xu) = (xx)u$ ;

б)  $(yx)y = y(xy)$ .

## 8.2. Представления матричных алгебр

### 8.2.1. Представления матричных алгебр

Пусть  $\mathcal{A}$  — некоторая ассоциативная алгебра,  $\mathfrak{gl}(V)$  — алгебра линейных преобразований пространства  $V$ . Гомоморфизм  $f: \mathcal{A} \rightarrow \mathfrak{gl}(V)$  называют *представлением* алгебры  $\mathcal{A}$ . Если задан гомоморфизм  $f$ , то элементы алгебры  $\mathcal{A}$  действуют на пространстве  $V$  согласно формуле  $av = f(a)v$ ; при этом  $(ab)v = a(bv)$ . Таким образом, пространство  $V$  является  $\mathcal{A}$ -модулем.

Подпространство  $W \subset V$  называют *инвариантным* подпространством представления  $f$ , если  $\mathcal{A}W \subset W$ , т.е.  $W$  — подмодуль  $\mathcal{A}$ -модуля  $V$ . Представление называют *неприводимым*, если любое его ненулевое инвариантное подпространство совпадает со всем пространством  $V$ . Представление  $f: \mathcal{A} \rightarrow \mathfrak{gl}(V)$  называют *вполне приводимым*, если пространство  $V$  является прямой суммой таких инвариантных подпространств, что ограничения на них представления  $f$  неприводимы.

**Теорема 8.2.1.1.** Пусть  $\mathcal{A} = \mathfrak{gl}(V)$  и  $f: \mathcal{A} \rightarrow \mathfrak{gl}(W)$  — такое представление, что  $f(1) = I$ . Тогда  $W = W_1 \oplus \dots \oplus W_k$ , где  $W_i$  — инвариантные подпространства, изоморфные  $V$ .



**Доказательство.** Пусть  $e_1, \dots, e_m$  — базис пространства  $W$ . Так как  $f(1)e_i = e_i$ , то  $W \subset \langle Ae_1, \dots, Ae_m \rangle$ . Пространство  $\mathcal{A}$  можно представить в виде прямой суммы подпространств  $F_i$ , состоящих из матриц, у которых столбцы с номерами  $1, \dots, i-1, i+1, \dots, n$  нулевые. Ясно, что  $\mathcal{A}F_i = F_i$  и если  $a$  — ненулевой элемент  $F_i$ , то  $\mathcal{A}a = F_i$ .

Пространство  $W$  является суммой пространств  $F_i e_j$ . Эти пространства инвариантны, так как  $\mathcal{A}F_i = F_i$ . Если  $x = ae_j$ , где  $a \in F_i$  и  $x \neq 0$ , то  $\mathcal{A}x = \mathcal{A}ae_j = F_i e_j$ . Поэтому любые два пространства вида  $F_i e_j$  либо не имеют общих ненулевых элементов, либо совпадают. Пространство  $W$  можно представить в виде прямой суммы некоторых ненулевых подпространств  $F_i e_j$ ; для этого нужно на каждом шаге добавлять подпространства, не содержащиеся в прямой сумме ранее выбранных подпространств. Остаётся доказать, что любое ненулевое пространство  $F_i e_j$  изоморфно  $V$ . Рассмотрим отображение  $h: F_i \rightarrow F_i e_j$ , для которого  $h(a) = ae_j$ . Ясно, что  $\mathcal{A} \operatorname{Ker} h \subset \operatorname{Ker} h$ . Если  $a$  — ненулевой элемент  $\operatorname{Ker} h$ , то  $\mathcal{A}a = F_i$ . Поэтому  $h$  либо изоморфизм, либо нулевое отображение.  $\square$

**Замечание 1.** Доказательство остаётся справедливым и для алгебры матриц над алгеброй кватернионов.

**Замечание 2.** Аналогичное утверждение можно доказать для алгебры  $\mathcal{A} = \mathfrak{gl}(V) \oplus \mathfrak{gl}(V)$ .

### 8.2.2. Теорема Бернсайда

**Теорема 8.2.2.1 (Бернсайд).** Пусть  $\mathcal{A}$  — подалгебра алгебры линейных преобразований пространства  $V$  над  $\mathbb{C}$ , причём пространство  $V$  неприводимо относительно действия  $\mathcal{A}$  и  $\dim V > 1$ . Тогда  $\mathcal{A} = \mathfrak{gl}(V)$ .

**Доказательство.** Если  $x \in V$  и  $x \neq 0$ , то  $\mathcal{A}x$  — инвариантное подпространство, поэтому  $\mathcal{A}x = V$  или  $\mathcal{A}x = 0$ ; равенство  $\mathcal{A}x = 0$  невозможно, так как в этом случае  $\langle x \rangle$  — инвариантное подпространство. Покажем сначала, что алгебра  $\mathcal{A}$  содержит некоторый оператор ранга 1. Предположим, что  $P_0$  — ненулевой оператор из  $\mathcal{A}$  минимального ранга и  $\operatorname{rk} P_0 = d > 1$ . Тогда векторы  $P_0 x_1$  и  $P_0 x_2$  линейно независимы для некоторых  $x_1$  и  $x_2$ . Пусть  $B$  — такой оператор из  $\mathcal{A}$ , что  $B P_0 x_1 = x_2$ . Тогда векторы  $P_0 B P_0 x_1 = P_0 x_2$  и  $P_0 x_1$  линейно независимы, а значит,  $P_0 B P_0 - \lambda P_0 \neq 0$  при всех  $\lambda$ . Оператор  $P_0 B - \lambda I$  можно ограничить на  $\operatorname{Im} P_0$ , а число  $\lambda \in \mathbb{C}$  можно выбрать так, чтобы полученный оператор был вырожденным. Тогда  $0 < \operatorname{rk}(P_0 B - \lambda I) P_0 < d$ . Получено противоречие, поэтому  $d = 1$ .

Оператор  $P_0$  ранга 1 можно записать в виде  $P_0x = f_0(x)y_0$ , где  $f_0$  — линейная функция. Докажем, что множество  $M = \{f \mid Bx = f(x)y_0 \text{ для } B \in \mathcal{A}\}$  содержит все линейные функции. Это множество является линейным пространством, поэтому достаточно проверить, что нет ненулевого вектора  $x$ , аннулируемого всеми  $f \in M$ . Пусть  $x \neq 0$  и  $f_0(x_1) \neq 0$ . Ясно, что  $P_0Bx = f_0(Bx)y_0$  и в  $\mathcal{A}$  существует такой оператор  $B$ , что  $Bx = x_1$ . Тогда оператору  $P_0B$  соответствует линейная функция  $f$ , для которой  $f(x) = f_0(Bx) = f_0(x_1) \neq 0$ . Так как  $CP_0x = f_0(x)Cy_0$ , то алгебра  $\mathcal{A}$  содержит все операторы ранга 1. Ясно также, что любой линейный оператор можно представить в виде суммы линейных операторов ранга 1.  $\square$

## Задачи

**2.1.** Матрицы  $A_1, \dots, A_n$  образуют группу относительно умножения и  $S = A_1 + \dots + A_n$ . Докажите, что если  $\text{tr } S = 0$ , то  $S = 0$ .

## 8.3. Конечные поля

### 8.3.1. Линейные пространства над конечными полями

Конечное поле, состоящее из  $n$  элементов, мы будем обозначать  $\mathbb{F}_n$ . По поводу конечных полей мы будем предполагать известными следующие факты:

- 1) для любого  $n = p^m$ , где  $p$  — простое число, существует единственное (с точностью до изоморфизма) поле из  $n$  элементов;
- 2) поле из  $n$  элементов существует только в том случае, когда  $n = p^m$ , где  $p$  — простое число.

Для поля из  $n = p^m$  элементов, где  $p$  — простое число, это число  $p$  называют *характеристикой*.

Линейные пространства над конечными полями определяются точно так же, как линейные пространства над  $\mathbb{R}$  или над  $\mathbb{C}$ , но у них, конечно, есть некоторые специфические свойства.

Прежде всего отметим, что линейное пространство размерности  $k$  над полем  $\mathbb{F}_n$  состоит из  $n^k$  векторов. Действительно, любой вектор этого пространства имеет вид  $(x_1, \dots, x_k)$ , где каждая из координат  $x_i$  принимает любое из  $n$  возможных значений.

Если поле  $\mathbb{F}_n$  содержит  $\mathbb{F}_m$  в качестве подполя, то  $\mathbb{F}_n$  является линейным пространством над  $\mathbb{F}_m$  некоторой размерности  $k$ , поэтому  $n = m^k$ . В частности, поле  $\mathbb{F}_8$  не содержит подполя  $\mathbb{F}_4$ . И вообще, если поле  $\mathbb{F}_{p^n}$  содержит подполе  $\mathbb{F}_{p^m}$ , то  $n = mk$  для некоторого натурального числа  $k$ .

### 8.3.2. Комбинаторика подпространств

**Теорема 8.3.2.1.** *Количество упорядоченных наборов  $k$  линейно независимых векторов в  $n$ -мерном пространстве над полем  $\mathbb{F}_q$  равно*

$$(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{k-1}).$$

**Доказательство.** Первый вектор можно выбрать  $q^n - 1$  способами (единственное ограничение заключается в том, что он должен быть отличен от нулевого вектора). Второй вектор должен быть отличен от  $q$  векторов, пропорциональных первому вектору, поэтому его можно выбрать  $q^n - q$  способами. Третий вектор не должен принадлежать плоскости, порождённой двумя первыми векторами, поэтому его можно выбрать  $q^n - q^2$  способами, и т.д. Последний вектор не должен принадлежать пространству, порождённым первыми  $k - 1$  векторами; количество векторов в этом пространстве равно  $q^{k-1}$ .  $\square$

**Следствие.** *Количество базисов в  $n$ -мерном пространстве над полем  $\mathbb{F}_q$  равно*

$$(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{n-1}).$$

**Теорема 8.3.2.2.** *Количество  $k$ -мерных подпространств в  $n$ -мерном пространстве над полем  $\mathbb{F}_q$  равно*

$$\frac{(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{k-1})}{(q^k - 1)(q^k - q)(q^k - q^2) \dots (q^k - q^{k-1})}.$$

**Доказательство.** В теореме 8.3.2.1 вычислено количество упорядоченных наборов  $k$  линейно независимых векторов в  $n$ -мерном пространстве над полем  $\mathbb{F}_q$ . Это число нужно поделить на количество базисов в  $k$ -мерном пространстве.  $\square$

### 8.3.3. Минимальный многочлен вектора

Минимальный многочлен оператора  $A$  и минимальный многочлен вектора  $v$  (относительно оператора  $A$ ) для полей конечной характеристики определяются точно так же, как и в случае поля  $\mathbb{R}$  (см. §3.3.1). Но в

доказательстве теоремы 3.3.1.2 о существовании вектора, минимальный многочлен которого совпадает с минимальным многочленом оператора, используется тот факт, что векторное пространство нельзя представить в виде конечного объединения его подпространств (отличных от самого пространства). Для векторных пространств над конечными полями этот факт неверен, но сама теорема 3.3.1.2 остаётся верной, нужно лишь доказывать её по-другому. Здесь мы приведём такое доказательство. Оно основано на следующей лемме.

**Лемма.** Пусть  $p_v$  и  $p_w$  — минимальные многочлены векторов  $v$  и  $w$ , причём  $p_v$  и  $p_w$  взаимно просты. Тогда минимальный многочлен вектора  $v + w$  равен  $p_v p_w$ .

**Доказательство.** Пусть  $p_{v+w}$  — минимальный многочлен вектора  $v + w$ . Ясно, что  $p_v(A)p_w(A)(v + w) = 0$ , поэтому  $p_v p_w$  делится на  $p_{v+w}$ .

Докажем теперь, что  $p_{v+w}$  делится на  $p_v$ . Рассмотрим для этого многочлен  $f = p_{v+w} p_w$ . Ясно, что  $f(A)(v + w) = 0$  и  $f(A)w = 0$ , поэтому  $f(A)v = 0$ , а значит,  $f$  делится на  $p_v$ . Многочлены  $p_v$  и  $p_w$  взаимно просты, поэтому  $p_{v+w}$  делится на  $p_v$ . Аналогично  $p_{v+w}$  делится на  $p_w$ . Ещё раз воспользовавшись взаимной простотой многочленов  $p_v$  и  $p_w$ , получаем требуемое.  $\square$

Пусть  $p$  — минимальный многочлен оператора  $A$ . Возьмём произвольный вектор  $v \neq 0$ . Если  $p_v \neq p$ , то найдётся вектор  $u$ , для которого минимальный многочлен  $p_u$  не делит  $p_v$ . Пусть  $d$  — наибольший общий делитель многочленов  $p_v$  и  $p_u$ . Рассмотрим многочлен  $f = p_u/d$ . Он является минимальным многочленом вектора  $w = d(A)u$ . Поэтому, применив лемму к векторам  $v$  и  $w$ , получим, что минимальный многочлен вектора  $v + w$  равен  $p_v f$ , причём степень этого многочлена строго больше степени многочлена  $p_v$ . Повторив эту операцию несколько раз, получим вектор, минимальный многочлен которого равен  $p$ .

Следует отметить, что теорема 3.4.1.1 о существовании канонической формы Фробениуса верна также и над конечными полями. Никаких изменений в доказательстве этой теоремы для конечных полей не требуется (нужно лишь заметить, что теорема 3.3.1.2, как мы только что доказали, верна и для конечных полей).

### 8.3.4. Квадратичные формы над конечными полями

Квадратичной формой над полем  $\mathbb{F}_q$  называют выражение вида  $f(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j$ , где  $a_{ij} \in \mathbb{F}_q$  (предполагается также, что пе-

ременные  $x_1, \dots, x_n$  принимают значения в  $\mathbb{F}_q$ ). По-другому можно сказать так: квадратичная форма  $f(x)$  равна  $B(x, x)$ , где  $B(x, y)$  — некоторая билинейная форма. Эта билинейная форма  $B(x, y) = \sum_{i,j=1}^n a_{ij}x_iy_j$  определена не однозначно, потому что квадратичная форма зависит только от коэффициентов  $a_{ii}$  и  $a_{ij} + a_{ji}$ .

Для поля нечётной характеристики билинейную форму  $B(x, y)$  можно выбрать симметрической, поскольку каждый член  $a_{ij}x_ix_j$  можно заменить на  $\frac{1}{2}(a_{ij} + a_{ji})x_ix_j$ ; такая симметрическая билинейная форма  $B(x, y)$  определена однозначно. Для поля чётной характеристики билинейную форму  $B(x, y)$  не всегда можно выбрать симметрической. Действительно, для поля чётной характеристики из равенства  $a_{ij} = a_{ji}$  следует, что  $a_{ij} + a_{ji} = 0$ . Поэтому для поля чётной характеристики симметрическую билинейную форму можно выбрать только для квадратичной формы вида  $\sum a_ix_i^2$ .

С билинейной формой  $B(x, y)$  связана матрица  $A$ , для которой  $B(x, y) = x^T Ay$ . Форма  $B$  симметрическая тогда и только тогда, когда матрица  $A$  симметрическая. В случае поля нечётной характеристики квадратичную форму называют *невырожденной*, если матрица  $A$ , соответствующая симметрической билинейной форме, невырожденная.

Если характеристика поля отлична от 2, то тождество  $f(x+y) - f(x) - f(y) = 2B(x, y)$  позволяет восстановить симметрическую билинейную форму  $B$  по квадратичной форме  $f$ . Таким образом, для полей, характеристика которых отлична от 2, есть взаимно однозначное соответствие между квадратичными формами и симметрическими билинейными функциями.

При замене координат матрица  $A$ , соответствующая квадратичной форме, заменяется на  $X^T AX$ , где  $X$  — матрица перехода. Две квадратичные формы называют *эквивалентными*, если они получаются друг из друга заменой координат.

Если уравнение  $f(x_1, \dots, x_n) = a$ , где  $f$  — квадратичная форма над  $\mathbb{F}_q$  и  $a \in \mathbb{F}_q$  — некоторый элемент, то говорят, что форма  $f$  *представляет* элемент  $a$ .

**Теорема 8.3.4.1.** Пусть квадратичная форма  $f$  от  $n \geq 2$  переменных над полем нечётной характеристики представляет элемент  $a \neq 0$ . Тогда она эквивалентна квадратичной форме вида  $ax_1^2 + g(x_2, \dots, x_n)$ .

**Доказательство.** Пусть  $f(c_1, \dots, c_n) = a \neq 0$ . Тогда набор  $(c_1, \dots, c_n)$  ненулевой, поэтому существует невырожденная матрица  $C$ , первый столбец которой имеет вид  $(c_1, \dots, c_n)^T$ . Пусть  $A$  — матрица, соответствующая данной квадратичной форме  $f$ . Тогда квадратичная форма,

соответствующая матрице  $C^T AC$ , имеет вид

$$\begin{aligned} ay_1^2 + 2b_2y_1y_2 + \dots + 2b_ny_1y_n + h(y_2, \dots, y_n) = \\ = a \left( y_1 + \frac{b_2}{a}y_2 + \dots + \frac{b_n}{a}y_n \right)^2 + g(x_2, \dots, x_n). \end{aligned}$$

□

**Теорема 8.3.4.2.** *Над полем нечётной характеристики любая квадратичная форма  $f$  эквивалентна квадратичной форме вида  $a_1x_1^2 + \dots + a_nx_n^2$ .*

**Доказательство.** Применим индукцию по  $n$ . При  $n = 1$  любая квадратичная форма имеет вид  $f(x_1) = a_{11}x_1^2$ . Предположим теперь, что  $n > 1$  и любая квадратичная форма от  $n - 1$  переменных эквивалентна диагональной квадратичной форме. Квадратичная форма с нулевыми коэффициентами является диагональной, поэтому будем предполагать, что хотя бы один коэффициент отличен от нуля. Возможны два случая:

- 1)  $a_{ii} \neq 0$  для некоторого  $i$ ;
- 2)  $a_{ii} = 0$  для всех  $i$ , но  $a_{ij} = a_{ji} \neq 0$  для некоторого  $i \neq j$ .

В первом случае квадратичная форма представляет  $a_{ii} \neq 0$  (нужно положить  $x_i = 1$  и  $x_j = 0$  при  $j \neq i$ ). Во втором случае квадратичная форма представляет  $2a_{ij} \neq 0$  (нужно положить  $x_i = x_j = 1$  и  $x_k = 0$  при  $k \neq i, k \neq j$ ). В обоих случаях квадратичная форма представляет ненулевой элемент  $a$ , поэтому согласно теореме 8.3.4.1 она эквивалентна квадратичной форме вида  $ax_1^2 + g(x_2, \dots, x_n)$ . Остаётся заметить, что согласно предположению индукции квадратичная форма  $g(x_2, \dots, x_n)$  эквивалентна диагональной квадратичной форме. □

### 8.3.5. Квадратичные формы над полями характеристики 2

Для поля характеристики 2 есть два разных (не эквивалентных) понятия невырожденной квадратичной формы. Во-первых, квадратичную форму  $f$  над полем чётной характеристики можно назвать *невырожденной*, если её нельзя представить в виде квадратичной формы от меньшего числа переменных. Ясно, что для любого натурального  $n$  существует квадратичная форма от  $n$  переменных, невырожденная в этом смысле. Во-вторых, невырожденность квадратичной формы можно определить на основе соответствующей ей билинейной формы. Напомним, что с каждой квадратичной формой  $q(x)$  над полем характеристики, отличной

от 2, связана симметрическая билинейная форма  $B(x, y)$ , для которой  $B(x, x) = q(x)$ . Тожество  $q(x + y) - q(x) - q(y) = 2B(x, y)$  позволяет восстановить  $B$  по  $q$ . Для поля характеристики 2 определение билинейной формы, связанной с квадратичной формой, изменяется следующим образом:  $q(x + y) + q(x) + q(y) = B(x, y)$ . Знаки здесь, конечно, несущественны; существенно лишь то, что  $2B(x, y)$  заменяется на  $B(x, y)$ , поскольку над полем характеристики 2 нельзя делить на 2. Квадратичную форму  $q(x)$  можно назвать *невырожденной*, если билинейная форма  $B$  невырождена.

Непосредственно из определения квадратичной формы видно, что  $B(x, y) = B(y, x)$  и  $B(x, x) = q(2x) + 2q(x) = 0$ . Поэтому точно такие же рассуждения, как для вещественных кососимметрических матриц, показывают, что если  $q$  — невырожденная (во втором смысле) квадратичная форма, то в пространстве  $V$ , на котором она определена, можно выбрать базис  $e_1, f_1, \dots, e_n, f_n$  так, что  $B(e_i, e_j) = B(f_i, f_j) = 0$  и  $B(e_i, f_j) = \delta_{ij}$ ; такой базис называют *симплектическим*.

Мы показали, что если квадратичная форма от  $n$  переменных невырождена во втором смысле, то число  $n$  чётно. Таким образом, эти два понятия невырожденности не эквивалентны.

Здесь мы обсудим невырожденность квадратичных форм над полем характеристики 2 в первом смысле и докажем теоремы 8.3.5.1 и 8.3.5.2. В параграфе 8.3.6 мы обсудим инвариант Арфа, который относится к квадратичным формам, невырожденным во втором смысле.

**Теорема 8.3.5.1.** *Над полем  $\mathbb{F}_q$ , где  $q$  чётно, невырожденная квадратичная форма  $f(x_1, \dots, x_n)$  от  $n \geq 3$  переменных эквивалентна квадратичной форме вида  $x_1x_2 + g(x_3, \dots, x_n)$ , где  $g$  — невырожденная квадратичная форма от  $n - 2$  переменных.*

**Доказательство.** Сначала докажем, что после замены  $f$  на эквивалентную ей квадратичную форму можно считать, что коэффициент при  $x_1^2$  равен 0. Пусть  $f(x_1, \dots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij}x_ix_j$ . Если  $a_{ii} = 0$ , то после изменения нумерации переменных получим  $a_{11} = 0$ . Поэтому можно считать, что все коэффициенты  $a_{ii}$  отличны от нуля. В таком случае все коэффициенты  $a_{ij}$ , где  $i < j$ , не могут быть равны нулю, поскольку тогда

$$f(x_1, \dots, x_n) = a_{11}x_1^2 + \dots + a_{nn}x_n^2 = (a_{11}^{q/2}x_1 + \dots + a_{nn}^{q/2}x_n)^2,$$

а это противоречит невырожденности квадратичной формы. Можно считать, что  $a_{23} \neq 0$ . Группируя члены, содержащие  $x_2$ , получаем

$$f(x_1, \dots, x_n) = a_{22}x_2^2 + x_2(a_{21}x_1 + a_{23}x_3 + \dots + a_{2n}x_n) + g_1(x_1, x_3, \dots, x_n).$$

Положим  $x_3 = a_{23}^{-1}(a_{12}y_1 + y_3 + a_{24}y_4 + \dots + a_{2n}y_n)$  и  $x_i = y_i$  при  $i \neq 3$ . В результате перейдём к эквивалентной квадратичной форме  $a_{22}y_2^2 + y_2y_3 + g_2(y_1, y_3, \dots, y_n)$ . Пусть в квадратичной форме  $g_2$  коэффициент при  $y_1^2$  равен  $b_{11}$ . Положим  $y_2 = (a_{22}^{-1}b_{11})^{q/2}z_1 + z_2$  и  $y_i = z_i$  при  $i \neq 2$ . В результате получим квадратичную форму, у которой коэффициент при  $z_1^2$  равен 0, что и требовалось.

Итак, мы можем считать, что  $f(x_1, \dots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij}x_i x_j$ , где  $a_{11} = 0$ . Квадратичная форма  $f$  невырожденная, поэтому все коэффициенты  $a_{1i}$  не могут быть равны 0. Для определённости будем считать, что  $a_{12} \neq 0$ . Положим  $x_2 = a_{12}^{-1}(y_2 + a_{13}y_3 + \dots + a_{1n}y_n)$  и  $x_i = y_i$  при  $i \neq 2$ . В результате перейдём к квадратичной форме вида  $y_1y_2 + \sum_{2 \leq i \leq j \leq n} a_{ij}y_i y_j$ . Наконец, положив  $y_1 = z_1 + c_{22}z_2 + \dots + c_{2n}z_n$  и  $y_i = z_i$  при  $i \neq 1$ , получим квадратичную форму  $z_1z_2 + g(z_2, \dots, z_n)$ . Из невырожденности формы  $f$  следует невырожденность  $g$ .  $\square$

**Теорема 8.3.5.2.** Пусть  $f(x_1, \dots, x_n)$  — невырожденная квадратичная форма от  $n$  переменных над полем  $\mathbb{F}_q$ , где  $q$  чётно. Если  $n$  нечётно, то  $f$  эквивалентна квадратичной форме

$$x_1x_2 + x_3x_4 + \dots + x_{n-2}x_{n-1} + x_n^2.$$

Если же  $n$  чётно, то  $f$  эквивалентна либо квадратичной форме

$$x_1x_2 + x_3x_4 + \dots + x_{n-1}x_n,$$

либо квадратичной форме

$$x_1x_2 + x_3x_4 + \dots + x_{n-1}x_n + x_{n-1}^2 + ax_n^2.$$

**Доказательство.** Пусть  $n$  нечётно. Применяя индукцию по  $n$  и теорему 8.3.5.1, получаем, что квадратичная форма  $f$  эквивалентна квадратичной форме  $x_1x_2 + x_3x_4 + \dots + x_{n-1}x_n + x_{n-1}^2 + ax_n^2$ , где  $a \neq 0$ . Заменяя  $x_n$  на  $a^{-q/2}x_n$ , получаем требуемую квадратичную форму.

Пусть теперь  $n$  чётно. Применяя индукцию по  $n$  и теорему 8.3.5.1, получаем, что квадратичная форма  $f$  эквивалентна квадратичной форме  $x_1x_2 + x_3x_4 + \dots + x_{n-3}x_{n-2} + bx_{n-1}^2 + cx_{n-1}x_n + dx_n^2$  для некоторых  $b, c, d \in \mathbb{F}_q$ . Прежде всего заметим, что  $c \neq 0$ , поскольку иначе равенство  $bx_{n-1}^2 + dx_n^2 = (b^{q/2}x_{n-1} + d^{q/2}x_n)^2$  позволило бы доказать вырожденность квадратичной формы  $f$ . Если  $b = 0$ , то квадратичная форма  $cx_{n-1}x_n + dx_n^2 = (cx_{n-1} + dx_n)x_n$  эквивалентна  $x_{n-1}x_n$ , и мы приходим к первой из указанных квадратичных форм. Если  $b \neq 0$ , то заменим



$x_{n-1}$  на  $b^{-q/2}x_{n-1}$ , а  $x_n$  на  $b^{q/2}c^{-1}x_n$ . В результате получим, что квадратичная форма  $bx_{n-1}^2 + cx_{n-1}x_n + dx_n^2$  эквивалентна квадратичной форме  $x_{n-1}x_n + x_{n-1}^2 + ax_n^2$ , и мы приходим ко второй из указанных форм. Отметим, что в этом случае элемент  $a$  должен быть таким, что многочлен  $x^2 + x + a$  неприводим над  $\mathbb{F}_q$ , поскольку иначе квадратичная форма  $x_{n-1}^2 + x_{n-1}x_n + ax_n^2 = (x_{n-1} + c_1x_n)(x_{n-1} + c_2x_n)$  вырожденная.  $\square$

### 8.3.6. Инвариант Арфа квадратичной формы

Для невырожденной (во втором смысле) квадратичной формы  $q$  над полем  $\mathbb{F}_2$  элемент  $c(q) = \sum_{i=1}^n q(e_i)q(f_i) \in \mathbb{F}_2$ , где  $e_1, \dots, f_n$  — симплектический базис, называют *инвариантом Арфа*.

**Теорема 8.3.6.1 (Арф [Ар]).** *Невырожденную квадратичную форму  $q$  можно привести к виду  $x_1y_1 + \dots + x_ny_n + c(q)(x_n^2 + y_n^2)$ , где  $c(q)$  — инвариант Арфа. При этом инвариант Арфа  $c(q)$  не зависит от выбора симплектического базиса.*

**Доказательство.** Рассмотрим сначала случай  $n = 1$ . Равенство  $q(e_1) + q(f_1) + q(e_1 + f_1) = 1$  показывает, что либо один из элементов  $q(e_1)$ ,  $q(f_1)$ ,  $q(e_1 + f_1)$  равен 1, а два других равны 0, либо все три элемента равны 1. Элементы  $e_1$ ,  $f_1$  и  $e_1 + f_1$  равноправны в том смысле, что любые два из них можно взять в качестве симплектического базиса, поскольку  $B(e_1 + f_1, e_1 + f_1) = 0$  и  $B(e_1 + f_1, e_1) = 1$ . Поэтому можно считать, что  $q(e_1 + f_1) = 1$ . Итак, в случае  $n = 1$  мы получаем две невырожденные квадратичные формы:  $q_0$  и  $q_1$ ; здесь  $q_0(e_1) = q_0(f_1) = 0$  и  $q_1(e_1) = q_1(f_1) = 1$ . В координатах эти формы можно записать так:  $q_0 = x_1y_1$  и  $q_1 = x_1y_1 + x_1^2 + y_1^2$ . Ясно, что  $c(q_0) = 0$  и  $c(q_1) = 1$ .

Чтобы доказать, что инвариант Арфа не зависит от выбора симплектического базиса, нужно доказать, что формы  $q_0$  и  $q_1$  не эквивалентны, т.е. не получаются друг из друга заменой базиса. Для этого достаточно заметить, что форма  $q_0$  принимает значение 1 ровно один раз, а форма  $q_1$  — три раза.

Рассмотрим теперь произвольную невырожденную квадратичную форму  $q$ . Выберем симплектический базис  $e_1, \dots, f_n$ . Пусть  $\varphi_i$  — ограничение формы  $q$  на подпространство, натянутое на векторы  $e_i$  и  $f_i$ . Тогда  $q = \varphi_1 \oplus \dots \oplus \varphi_n$ . Каждая форма  $\varphi_i$  эквивалентна  $q_0$  или  $q_1$ .

**Лемма 1.** *Квадратичные формы  $\psi_0 = q_0 \oplus q_0$  и  $\psi_1 = q_1 \oplus q_1$  эквивалентны.*

**Доказательство.** Пусть  $e_1, f_1, e_2, f_2$  — симплектический базис, для которого  $\psi_0(e_i) = \psi_0(f_i) = 0$  и  $\psi_1(e_i) = \psi_1(f_i) = 1$ . Рассмотрим базис

$e'_1 = e_1 + e_2$ ,  $f'_1 = e_1 + f_2$ ,  $e'_2 = e_1 + f_1 + e_2 + f_2$ ,  $f'_2 = e_1 + f_1 + f_2$ . Легко проверить, что этот базис симплектический. Проверим, что  $\psi_1(e'_i) = \psi_1(f'_i) = 0$ , т.е.  $\psi_1(e'_i) = \psi_0(e_i)$  и  $\psi_1(f'_i) = \psi_0(f_i)$ . Действительно,  $\psi_1(e_1 + e_2) = \psi_1(e_1) + \psi_1(e_2) + B(e_1, e_2) = 1 + 1 + 0 = 0$ ,  $\psi_1(e_1 + f_2) = \psi_1(e_1) + \psi_1(f_2) + B(e_1, f_2) = 1 + 1 + 0 = 0$ ,  $\psi_1(e_1 + f_1 + e_2 + f_2) = \psi_1(e_1 + e_2) + \psi_1(f_1 + f_2) + B(e_1 + e_2, f_1 + f_2) = 0 + 0 + 0 = 0$ ,  $\psi_1(e_1 + f_1 + f_2) = \psi_1(e_1 + f_2) + \psi_1(f_1) + B(e_1 + f_2, f_1) = 0 + 1 + 1 = 0$ .  $\square$

Введём следующее обозначение:  $\underbrace{q \oplus \cdots \oplus q}_{n \text{ раз}} = nq$ . Из леммы 1 следует,

что форма  $q$  эквивалентна либо форме  $nq_0$ , либо форме  $(n-1)q_0 \oplus q_1$ . В координатах эти формы записываются так:  $x_1y_1 + \cdots + x_ny_n$  и  $x_1y_1 + \cdots + x_ny_n + x_n^2 + y_n^2$ . Ясно также, что  $c(nq_0) = 0$  и  $c((n-1)q_0 \oplus q_1) = 1$ .

Чтобы доказать, что инвариант Арфа не зависит от выбора симплектического базиса, нужно доказать, что формы  $nq_0$  и  $(n-1)q_0 \oplus q_1$  не эквивалентны.

**Лемма 2.** Значение 1 форма  $nq_0$  принимает ровно  $2^{2n-1} - 2^{n-1}$  раз, а форма  $(n-1)q_0 \oplus q_1$  — ровно  $2^{2n-1} + 2^{n-1}$  раз.

**Доказательство.** Применим индукцию по  $n$ . При  $n = 1$  утверждение легко проверяется. Предположим, что мы уже доказали, что форма  $nq_0$  принимает значение 1 ровно  $2^{2n-1} - 2^{n-1}$  раз. Форма  $q_0$  принимает значение 1 ровно 1 раз, а значение 0 ровно 3 раза. Поэтому форма  $(n+1)q_0$  принимает значение 1 ровно  $3(2^{2n-1} - 2^{n-1}) + 2^{2n-1} + 2^{n-1} = 2^{2n+1} - 2^n$  раз (мы воспользовались тем, что форма  $nq_0$  принимает значение 0 ровно  $2^{2n-1} + 2^{n-1}$  раз). Форма  $q_1$  принимает значение 1 ровно 3 раза, а значение 0 ровно 1 раз. Поэтому форма  $nq_0 \oplus q_1$  принимает значение 1 ровно  $3(2^{2n-1} + 2^{n-1}) + 2^{2n-1} - 2^{n-1}$  раз.  $\square$

$\square$

## Задачи

**3.1.** Докажите, что определитель кососимметрической матрицы  $A$  нечётного порядка над полем нечётной характеристики равен нулю. Верно ли аналогичное утверждение для поля характеристики 2?

**3.2.** Приведите пример матриц  $X$  и  $Y$  над полем  $\mathbb{F}_p$ , удовлетворяющих соотношению  $XY - YX = I$ .

**3.3.** [Pe1] Пусть  $A$  — квадратная матрица с целочисленными элементами,  $p$  — простое число. Докажите, что  $\text{tr}(A^p) \equiv \text{tr } A \pmod{p}$ .

## 8.4. Результат

### 8.4.1. Матрица Сильвестра

Рассмотрим многочлены  $f(x) = \sum_{i=0}^n a_i x^{n-i}$  и  $g(x) = \sum_{i=0}^m b_i x^{m-i}$ , где  $a_0 \neq 0$  и  $b_0 \neq 0$ . Если  $f$  и  $g$  имеют общий корень, то  $f = \varphi p$  и  $g = \varphi q$  для некоторых многочленов  $\varphi$ ,  $p$  и  $q$ . Следовательно,  $fq = gp$ , причём  $\deg q \leq m-1$  и  $\deg p \leq n-1$ . Ясно также, что если  $f$  и  $g$  не имеют общих корней и  $fq = gp = h$ , то  $h$  делится на  $fg$ , а значит,  $\deg q \geq \deg g = m$ .

Пусть  $q = u_0 x^{m-1} + u_1 x^{m-2} + \dots + u_{m-1}$  и  $p = v_0 x^{n-1} + \dots + v_{n-1}$ . Равенство  $fq = gp$  можно записать в виде системы уравнений

$$\begin{aligned} a_0 u_0 &= b_0 v_0, \\ a_1 u_0 + a_0 u_1 &= b_1 v_0 + b_0 v_1, \\ a_2 u_0 + a_1 u_1 + a_0 u_2 &= b_2 v_0 + b_1 v_1 + b_0 v_2, \\ &\dots \dots \dots \end{aligned}$$

Многочлены  $f$  и  $g$  имеют общий корень тогда и только тогда, когда эта система уравнений имеет ненулевое решение  $(u_0, u_1, \dots, v_0, v_1, \dots)$ . Если, например,  $m = 3$  и  $n = 2$ , то определитель этой системы уравнений имеет вид

$$\begin{vmatrix} a_0 & 0 & 0 & -b_0 & 0 \\ a_1 & a_0 & 0 & -b_1 & -b_0 \\ a_2 & a_1 & a_0 & -b_2 & -b_1 \\ 0 & a_2 & a_1 & -b_3 & -b_2 \\ 0 & 0 & a_2 & 0 & -b_3 \end{vmatrix} = \pm \begin{vmatrix} a_0 & a_1 & a_2 & 0 & 0 \\ 0 & a_0 & a_1 & a_2 & 0 \\ 0 & 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & b_3 & 0 \\ 0 & b_0 & b_1 & b_2 & b_3 \end{vmatrix} = \pm |S(f, g)|.$$

Матрицу  $S(f, g)$  называют *матрицей Сильвестра* многочленов  $f$  и  $g$ . Определитель матрицы  $S(f, g)$  называют *результантом* многочленов  $f$  и  $g$ ; его обозначают  $R(f, g)$ . Ясно, что  $R(f, g)$  — однородный многочлен степени  $m$  по переменным  $a_i$  и степени  $n$  по переменным  $b_j$ . Многочлены  $f$  и  $g$  имеют общий корень тогда и только тогда, когда определитель рассматриваемой системы равен нулю, т.е.  $R(f, g) = 0$ .

### 8.4.2. Выражение результата через корни

**Теорема 8.4.2.1.** Пусть  $x_i$  — корни многочлена  $f$ ,  $y_j$  — корни многочлена  $g$ . Тогда  $R(f, g) = a_0^m b_0^n \prod_{i,j} (x_i - y_j) = a_0^m \prod_i g(x_i) = b_0^n \prod_j f(y_j)$ .

**Доказательство.** Так как  $f = a_0(x - x_1) \dots (x - x_n)$ , то  $a_k = \pm a_0 \sigma_k(x_1, \dots, x_n)$ , где  $\sigma_k$  — элементарная симметрическая функция. Аналогично  $b_k = \pm b_0 \sigma_k(y_1, \dots, y_m)$ . Результат является однородным многочленом степени  $m$  по переменным  $a_i$  и степени  $n$  по переменным  $b_j$ , поэтому

$$R(f, g) = a_0^m b_0^n P(x_1, \dots, x_n, y_1, \dots, y_m),$$

где  $P$  — симметрический многочлен от  $x_1, \dots, x_n$  и  $y_1, \dots, y_m$ , обращающийся в нуль при  $x_i = y_j$ . Формула  $x^k = (x_i - y_j)x_i^{k-1} + y_j x_i^{k-1}$  показывает, что  $P(x_1, \dots, y_m) = (x_i - y_j)Q(x_1, \dots, y_m) + T(x_1, \dots, \hat{x}_i, \dots, y_m)$ . Подставив в это равенство  $x_i = y_j$ , получим, что  $T(x_1, \dots, \hat{x}_i, \dots, y_m) \equiv 0$ , т.е.  $T$  — нулевой многочлен. Аналогичные рассуждения показывают, что многочлен  $P$  делится на  $S = a_0^m b_0^n \prod_{i,j} (x_i - y_j)$ .

Так как  $g(x) = b_0 \prod_j (x - y_j)$ , то  $\prod_i g(x_i) = b_0^n \prod_{i,j} (x_i - y_j)$ , а значит,  $S = a_0^m \prod_i g(x_i) = a_0^m \prod_i (b_0 x_i^m + b_1 x_i^{m-1} + \dots + b_m)$  — однородный многочлен степени  $n$  по переменным  $b_0, \dots, b_m$ . Для переменных  $a_0, \dots, a_n$  рассуждения аналогичны. Ясно также, что симметрический многочлен  $a_0^m \prod_i (b_0 x_i^m + b_1 x_i^{m-1} + \dots + b_m)$  является многочленом от  $a_0, \dots, a_n, b_0, \dots, b_m$ . Следовательно,  $R(a_0, \dots, b_m) = \lambda S(a_0, \dots, b_m)$ , где  $\lambda$  — некоторое число. С другой стороны, коэффициент при  $\prod_i x_i^m$  в многочленах  $a_0^m b_0^n R(x_1, \dots, y_m)$  и  $S(x_1, \dots, y_m)$  равен  $a_0^m b_0^n$ , поэтому  $\lambda = 1$   $\square$

### 8.4.3. Матрица Безу

Матрица Сильвестра имеет слишком большой порядок, поэтому вычислять результат с её помощью неудобно. Есть много разных способов понизить порядок матрицы для вычисления результата. Можно, например, заменить многочлен  $g$  на остаток от его деления на многочлен  $f$  (см. задачу 4.2). Другой способ понижения порядка матрицы состоит в переходе к матрице Безу. Мы ограничимся рассмотрением случая  $n = m$ , хотя для многочленов неравных степеней тоже существуют аналогичные, но менее симметричные формулы.

Запишем матрицу Сильвестра в блочном виде  $\begin{pmatrix} A_1 & A_2 \\ B_1 & B_2 \end{pmatrix}$ , где матри-

цы  $A_i, B_i$  квадратные. Легко проверить, что

$$A_1 B_1 = \begin{pmatrix} c_0 & c_1 & \dots & c_{n-1} & c_n \\ 0 & c_0 & \dots & c_{n-2} & c_{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_0 & c_1 \\ 0 & 0 & \dots & 0 & c_0 \end{pmatrix} = B_1 A_1,$$

где  $c_k = \sum_{i=0}^k a_i b_{k-i}$ . Поэтому

$$\begin{pmatrix} I & 0 \\ -B_1 & A_1 \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ B_1 & B_2 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_1 B_2 - B_1 A_2 \end{pmatrix},$$

а так как  $|A_1| = a_0^n \neq 0$ , то  $R(f, g) = |A_1 B_2 - B_1 A_2|$ .

Пусть  $c_{pq} = a_p b_q - a_q b_p$ . Легко проверить, что  $A_1 B_2 - B_1 A_2 = \|w_{ij}\|_1^n$ , где  $w_{ij} = \sum c_{pq}$  и суммирование ведётся по таким парам  $(p, q)$ , что  $p+q = n+j-i$ ,  $p \leq n-i$  и  $q \geq j$ . А так как  $c_{\alpha\beta} + c_{\alpha+1, \beta-1} + \dots + c_{\beta\alpha} = 0$  при  $\alpha \leq \beta$ , то можно ограничиться теми парами, для которых  $p \leq \min(n-i, j-1)$ . Например, при  $n=4$  получим матрицу

$$\begin{pmatrix} c_{04} & c_{14} & c_{24} & c_{34} \\ c_{03} & c_{04} + c_{13} & c_{14} + c_{23} & c_{24} \\ c_{02} & c_{03} + c_{12} & c_{04} + c_{13} & c_{14} \\ c_{01} & c_{02} & c_{03} & c_{04} \end{pmatrix}.$$

Если  $J = \begin{pmatrix} 0 & & & 1 \\ & \ddots & & \\ & & \ddots & \\ 1 & & & 0 \end{pmatrix}$ , то матрица  $Z = \|w_{ij}\|_1^n J$  симметрична; её

называют *безутианой* (или *матрицей Безу*) многочленов  $f$  и  $g$ . При этом  $z_{ij} = w_{i, n-j+1} = \sum c_{pq}$ , где  $p+q = 2n+1-i-j$  и  $p \leq \min(n-i, n-j)$ .

**Теорема 8.4.3.1.** Пусть  $f$  и  $g$  — многочлены степени  $n$ . Тогда

$$\frac{f(x)g(y) - f(y)g(x)}{x-y} = \sum_{p,q=0}^{n-1} z_{p+1, q+1} x^p y^q,$$

где  $Z = \|z_{ij}\|_1^n$  — матрица Безу многочленов  $f$  и  $g$ .

**Доказательство.** Если  $\alpha > \beta$ , то

$$\frac{x^\alpha y^\beta - x^\beta y^\alpha}{x-y} = x^\beta y^\beta \sum_{k+l=\beta-\alpha-1} x^k y^l = \sum_{k+l=\beta-\alpha-1} x^{k+\beta} y^{l+\beta}.$$

Ясно также, что коэффициент многочлена  $f(x)g(y) - f(y)g(x)$  при  $x^\alpha y^\beta - x^\beta y^\alpha$  равен  $a_{n-\alpha}b_{n-\beta} - a_{n-\beta}b_{n-\alpha}$ . Поэтому  $\frac{f(x)g(y) - f(y)g(x)}{x - y}$  — многочлен, причём если  $p \geq q$ , то его коэффициент при  $x^p y^q$  равен  $\sum (a_{n-\alpha}b_{n-\beta} - a_{n-\beta}b_{n-\alpha})$ , где суммирование ведётся по таким парам  $(\alpha, \beta)$ , что  $p + q = \beta + \alpha - 1$ , причём  $\alpha \leq n$  и  $\beta \leq q$ , т.е.  $p + 1 \leq \alpha \leq n$ . Результат суммирования не изменится, если нижнюю границу заменить на  $\max(p + 1, q + 1)$ . Остаётся заметить, что коэффициенты при  $x^p y^q$  и  $x^q y^p$  равны и  $z_{p+1, q+1} = \sum c_{n-\alpha, n-\beta}$ , где  $n - \alpha + n - \beta = 2n - p - q - 1$ , т.е.  $\alpha + \beta = p + q + 1$ , и  $0 \leq n - \alpha \leq \min(n - p - 1, n - q - 1)$ , т.е.  $\max(p + 1, q + 1) \leq \alpha \leq n$ .  $\square$

#### 8.4.4. Понижение порядка матрицы для вычисления результата

Опишем ещё один способ понижения порядка матрицы для вычисления результата. Для простоты будем считать, что  $a_0 = 1$ , т.е.  $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$  и  $g(x) = b_0 x^m + b_1 x^{m-1} + \dots + b_m$ . Рассмотрим матрицу

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-2} & -a_{n-3} & \dots & -a_1 \end{pmatrix}.$$

**Теорема 8.4.4.1.** Пусть  $R = g(A)$ . Тогда  $\det R = R(f, g)$ .

**Доказательство.** Пусть  $\alpha_i$  и  $\beta_i$  — корни многочленов  $f$  и  $g$ . Тогда  $g(x) = b_0(x - \beta_1) \dots (x - \beta_m)$ . Поэтому  $g(A) = b_0(A - \beta_1 I) \dots (A - \beta_m I)$ . А так как  $|A - \lambda I| = \prod_i (\alpha_i - \lambda)$  (см. §1.5), то  $\det g(A) = b_0^m \prod_{i,j} (\alpha_i - \beta_j) = R(f, g)$ .  $\square$

**Теорема 8.4.4.2.** При  $m \leq n$  матрицу  $R$  можно вычислять следующим рекуррентным способом. Пусть  $r_1, \dots, r_n$  — строки матрицы  $R$ . Тогда  $r_1 = (b_m, b_{m-1}, \dots, b_1, b_0, 0, \dots, 0)$  при  $m < n$ , а если  $m = n$ , то  $r_1 = (d_n, \dots, d_1)$ , где  $d_i = b_i - b_0 a_i$ . Кроме того,  $r_i = r_{i-1} A$  для  $i = 2, \dots, n$ .

**Доказательство.** Пусть  $e_i = (0, \dots, 1, \dots, 0)$ , где 1 стоит на  $i$ -м месте. При  $k < n$  первая строка матрицы  $A^k$  равна  $e_{k+1}$ . Поэтому строение  $r_1$  при  $m < n$  очевидно. При  $m = n$  нужно воспользоваться равенством  $A^n + \sum_{i=1}^n a_i A^{n-i} = 0$ .

Так как  $e_i = e_{i-1}A$  при  $i = 2, \dots, n$ , то  $r_i = e_i R = e_{i-1}AR = e_{i-1}RA = r_{i-1}A$ .  $\square$

**Теорема 8.4.4.3.** *Степень наибольшего общего делителя многочленов  $f$  и  $g$  равна  $n - \text{rk } R$ .*

**Доказательство.** Пусть  $\beta_1, \dots, \beta_s$  — корни многочлена  $g$  с кратностями  $k_1, \dots, k_s$  соответственно. Тогда  $g(x) = b_0 \prod_i (x - \beta_i)^{k_i}$  и  $R = g(A) = b_0 \prod_i (A - \beta_i I)^{k_i}$ .

При переходе к базису, в котором матрица  $A$  имеет жорданову нормальную форму  $J$ , матрица  $R$  заменится на матрицу  $b_0 \prod_i (J - \beta_i I)^{k_i}$ .

Характеристический многочлен матрицы  $A$  совпадает с минимальным многочленом, поэтому если  $\beta_i$  — корень кратности  $l_i$  многочлена  $f$ , то жорданова клетка  $J_i$  матрицы  $J$ , соответствующая собственному значению  $\beta_i$ , имеет порядок  $l_i$ . Ясно также, что  $\text{rk}(J_i - \beta_i I)^{k_i} = l_i - \min(k_i, l_i)$ . Теперь, рассматривая жордановы клетки матрицы  $J$  по отдельности, легко убедиться, что  $n - \text{rk } R = \sum_i \min(k_i, l_i)$ , а последняя сумма равна степени наибольшего общего делителя многочленов  $f$  и  $g$ .  $\square$

Доказательства теорем этого пункта следуют работе [Ba1].

#### 8.4.5. Дискриминант

Пусть  $x_1, \dots, x_n$  — корни многочлена  $f(x) = a_0 x^n + \dots + a_n$ , причём  $a_0 \neq 0$ . Величину  $D(f) = a_0^{2n-2} \prod_{i < j} (x_i - x_j)^2$  называют *дискриминантом* многочлена  $f$ . Ясно, что  $D(f) = 0$  тогда и только тогда, когда  $f$  имеет кратные корни, т.е.  $R(f, f') = 0$ .

**Теорема 8.4.5.1.**  $R(f, f') = \pm a_0 D(f)$ .

**Доказательство.** Согласно теореме 8.4.2.1  $R(f, f') = a_0^{n-1} \prod_i f'(x_i)$ . Легко проверить, что если  $x_i$  — корень многочлена  $f$ , то  $f'(x_i) = a_0 \prod_{j \neq i} (x_j - x_i)$ .

Поэтому  $R(f, f') = a_0^{2n-1} \prod_{j \neq i} (x_i - x_j) = \pm a_0^{2n-1} \prod_{i < j} (x_i - x_j)^2$ .  $\square$

**Следствие.** *Дискриминант является многочленом от коэффициентов многочлена  $f$ .*

**Теорема 8.4.5.2.** *Любая матрица является пределом матриц с некрратными собственными значениями.*

**Доказательство.** Пусть  $f(x)$  — характеристический многочлен матрицы  $A$ . Многочлен  $f$  имеет кратные корни тогда и только тогда, когда  $D(f) = 0$ . В результате получаем алгебраическое уравнение для элементов матрицы  $A$ . Ограничение уравнения  $D(f) = 0$  на прямую  $\{\lambda A + (1 - \lambda)B\}$ , где  $B$  — матрица с некрратными собственными значениями, имеет конечное число корней. Поэтому матрица  $A$  является пределом матриц с некрратными собственными значениями.  $\square$

### Задачи

**4.1.** Даны числа  $n$  и  $m$ . Докажите, что существуют такие многочлены  $\varphi$  и  $\psi$  с целыми коэффициентами от переменных  $x, a_0, \dots, a_n, b_0, \dots, b_m$ , что

$$\varphi(a, b, x)f(x) + \psi(a, b, x)g(x) = R(f, g)$$

для  $f(x) = \sum_{i=0}^n a_i x^{n-i}$  и  $g(x) = \sum_{i=0}^m b_i x^{m-i}$ .

**4.2.** Пусть  $r(x)$  — остаток от деления многочлена  $g(x)$  на многочлен  $f(x)$  и  $\deg r(x) = k$ . Докажите, что  $R(f, g) = a_0^{m-k} R(f, r)$ .

**4.3.** Пусть  $f(x) = a_0 x^n + \dots + a_n$ ,  $g(x) = b_0 x^m + \dots + b_m$  и  $r_k(x) = a_{k0} x^{n-1} + a_{k1} x^{n-2} + \dots + a_{k,n-1}$  — остаток от деления многочлена  $x^k g(x)$  на  $f(x)$ . Докажите, что

$$R(f, g) = a_0^m \begin{vmatrix} a_{n-1,0} & \dots & a_{n-1,n-1} \\ \dots & \dots & \dots \\ a_{00} & \dots & a_{0,n-1} \end{vmatrix}.$$

**4.4.** Характеристические многочлены матриц  $A$  и  $B$  порядков  $n$  и  $m$  соответственно равны  $f$  и  $g$ . Докажите, что результат многочленов  $f$  и  $g$  с точностью до знака равен определителю оператора  $X \mapsto AX - XB$  в пространстве матриц размера  $n \times m$ .

**4.5.** Пусть  $\alpha_1, \dots, \alpha_n$  — корни многочлена  $f(x) = \sum_{i=0}^n a_i x^{n-i}$  и  $s_k = \alpha_1^k + \dots + \alpha_n^k$ . Докажите, что  $D(f) = a_0^{2n-1} \det S$ , где

$$S = \begin{pmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \dots & \dots & \dots & \dots \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{pmatrix}.$$



**4.6.** Пусть  $f$  и  $g$  — многочлены степени  $n$ . Докажите, что их безугиана равна нулю тогда и только тогда, когда эти многочлены пропорциональны.

**4.7.** [Ba2] Пусть  $Z$  — матрица Безу многочленов  $f = x^n + a_1x^{n-1} + \dots + a_n$  и  $g(x) = b_0x^n + \dots + b_n$ ;  $R$  — матрица, определённая в §8.4.4. Докажите, что  $Z = TR$ , где

$$T = \begin{pmatrix} a_{n-1} & a_{n-2} & \dots & a_1 & 1 \\ a_{n-2} & a_{n-3} & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

## 8.5. Теорема Витта

### 8.5.1. Пространства Артина

Пусть  $g$  — симметрическая билинейная функция на пространстве  $V$ . Определим в пространстве  $V$  скалярное произведение формулой  $(x, y) = g(x, y)$ . Если для любого ненулевого вектора  $x \in V$  существует такой вектор  $y \in V$ , что  $(x, y) \neq 0$ , то скалярное произведение называют *невыврожденным*; пространство  $V$  с таким скалярным произведением называют *невыврожденным*.

*Изометрией* пространств  $(V_1, g_1)$  и  $(V_2, g_2)$  со скалярными произведениями называют такой изоморфизм  $F: V_1 \rightarrow V_2$ , что  $g_2(F(x), F(y)) = g_1(x, y)$  для любых векторов  $x$  и  $y$ . Пространства, между которыми существует изометрия, называют *изометричными*.

Пусть  $W \subset V$  — некоторое подпространство. *Ортогональным дополнением* пространства  $W$  относительно данного скалярного произведения называют подпространство

$$W^\perp = \{v \in V \mid (v, x) = 0 \quad \forall x \in W\}.$$

Ясно, что  $W \subset (W^\perp)^\perp$ . Равенство  $W = (W^\perp)^\perp$  выполняется не всегда; например, если  $(x, y) \equiv 0$ , то  $(W^\perp)^\perp = V$  для любого  $W \subset V$ . Но для невырожденного пространства  $V$  всегда выполняются равенства  $\dim W + \dim W^\perp = \dim V$  и  $(W^\perp)^\perp = W$  (см. задачу 5.1).

Векторы  $x$  и  $y$  называют *ортогональными*, если  $(x, y) = 0$ . Квадратичную форму  $q(x) = (x, x)$  можно привести к главным осям (см. §4.1.3),

поэтому в любом пространстве  $V$  над  $\mathbb{R}$  можно выбрать такой ортогональный базис  $x_1, \dots, x_s, u_1, \dots, u_{n-s}$ , что  $(x_i, x_i) = 0$  и  $(u_j, u_j) \neq 0$ . При этом пространство  $\langle x_1, \dots, x_s \rangle$  определено инвариантно: оно равно  $V^\perp$ .

Отметим, что подпространство в невырожденном пространстве может быть вырожденным. Рассмотрим, например, пространство с базисом  $x_1, y_1$  и матрицей скалярного произведения  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , т.е.  $(x_1, x_1) = 0$ ,  $(y_1, y_1) = 0$  и  $(x_1, y_1) = 1$ . Это пространство невырожденное, но подпространство, порождённое вектором  $x_1$ , вырожденное. В этом примере существенно, что скалярное произведение не является знакоопределённым.

**Теорема 8.5.1.1.** Пусть  $x_1, \dots, x_s, u_1, \dots, u_{k-s}$  — набор линейно независимых попарно ортогональных векторов в невырожденном пространстве  $V$ , причём  $(x_i, x_i) = 0$  и  $(u_i, u_i) \neq 0$ . Тогда в пространстве  $V$  существуют такие векторы  $y_1, \dots, y_s$ , что  $x_1, \dots, x_s, y_1, \dots, y_s, u_1, \dots, u_{k-s}$  — набор линейно независимых векторов и матрица ограничения скалярного произведения на каждое подпространство  $\langle x_i, y_i \rangle$  равна  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , причём подпространства  $\langle x_1, y_1 \rangle, \dots, \langle x_s, y_s \rangle, \langle u_1 \rangle, \dots, \langle u_{k-s} \rangle$  попарно ортогональны.

**Доказательство.** Пусть  $W_1 = \langle x_2, \dots, x_s, u_1, \dots, u_{k-s} \rangle$  и  $W_2 = \langle x_1, W_1 \rangle$ . Тогда  $\dim W_1^\perp = \dim W_2^\perp + 1$ , поэтому существует вектор  $z_1 \in W_1^\perp \setminus W_2^\perp$ . Можно считать, что  $(z_1, x_1) = 1$  и  $z_1 \perp W_1$ . Пусть  $y_1 = z_1 + \lambda x_1$ , где  $\lambda = -(z_1, z_1)/2$ . Тогда  $(y_1, x_1) = 1$ ,  $(y_1, x_1) = 0$  и  $y_1 \perp W_1$ . Легко проверить, что  $x_2, \dots, x_s, u_1, \dots, u_{k-s} \in \langle x_1, y_1 \rangle^\perp$  и  $\langle x_1, y_1 \rangle \cap \langle x_1, y_1 \rangle^\perp = 0$ . В самом деле, если  $v = \alpha x_1 + \beta y_1 \in \langle x_1, y_1 \rangle^\perp$ , то  $\alpha = (v, y_1) = 0$  и  $\beta = (v, x_1) = 0$ . Поэтому можно перейти к пространству  $\langle x_1, y_1 \rangle^\perp$  и аналогичным образом построить в нем вектор  $y_2$  и т.д.  $\square$

**Замечание.** Пространство  $\langle x_1, \dots, x_s, y_1, \dots, y_s, u_1, \dots, u_{k-s} \rangle$  невырождено.

Ненулевой вектор  $x$  называют *изотропным*, если  $(x, x) = 0$ . Пространство  $\langle x, y \rangle$  с матрицей скалярного произведения  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  называют *плоскостью Артина*. Прямую сумму попарно ортогональных плоскостей Артина называют *пространством Артина*.

### 8.5.2. Доказательство теоремы Витта

**Теорема 8.5.2.1 (Витт).** Пусть  $W$  и  $W'$  — подпространства невырожденного пространства  $V$  со скалярным произведением и  $f: W \rightarrow W'$  — неко-

торая изометрия. Тогда  $f$  можно продолжить до изометрического преобразования всего пространства  $V$ .

**Доказательство.** Подпространства  $W$  и  $W'$  можно считать невырожденными. В самом деле, выберем в пространстве  $W$  такой ортогональный базис  $x_1, \dots, x_s, u_1, \dots, u_{k-s}$ , что  $(x_i, x_i) = 0$  и  $(u_j, u_j) \neq 0$ , и дополним его векторами  $y_1, \dots, y_s$ , как указано в формулировке теоремы 8.5.1.1. Пусть  $x_i = f(x_i)$  и  $u'_i = f(u_i)$ . Базис  $x'_1, \dots, x'_s, u'_1, \dots, u'_{k-s}$  можно аналогичным образом дополнить векторами  $y'_1, \dots, y'_s$ . Положим  $\bar{f}(y_i) = y'_i$  и  $\bar{f}(w) = f(w)$  для всех  $w \in W$ . В результате получим изометрию невырожденных пространств  $\bar{W} = \langle W, y_1, \dots, y_s \rangle$  и  $\bar{W}' = \langle W', y'_1, \dots, y'_s \rangle$  и будем продолжать её.

Предположим сначала, что  $W$  и  $W'$  — одномерные невырожденные пространства. Пусть  $x$  — ненулевой элемент  $W$  и  $y = f(x) \in W'$ . По условию  $(x, x) = (y, y) \neq 0$ , поэтому  $(x + y, x + y) + (x - y, x - y) = 2(x, x) + 2(y, y) = 4(x, x) \neq 0$ . Следовательно, хотя бы один из векторов  $x + y$  и  $x - y$  неизотропен. Простые вычисления показывают, что для любого неизотропного вектора  $a$  преобразование

$$s_a(u) = u - \frac{(a, u)}{(a, a)}a$$

является изометрией. Если  $a = x \pm y$ , то  $(a, a) = 2(a, x)$ , поэтому  $s_{x \pm y}(x) = x - (x \pm y) = \mp y$ , а значит,  $-s_{x+y}$  или  $s_{x-y}$  — искомая изометрия.

Докажем теперь требуемое утверждение для невырожденных пространств  $W$  и  $W'$  произвольной размерности  $n$ . Доказательство проведём индукцией по  $n$ . Пусть  $x$  — неизотропный вектор пространства  $W$  и  $W_1 = \langle x \rangle$ . Тогда  $W = W_1 \oplus W_2$ , где  $W_2$  — ортогональное дополнение пространства  $W_1$  в  $W$ . Ограничение изометрии  $f$  на одномерное подпространство  $W_1$  можно продолжить до изометрии  $g$  всего пространства  $V$ . Остаётся построить такую изометрию  $h$  пространства  $W_1^\perp$ , что ограничения отображений  $h \circ g$  и  $f$  на  $W_2$  совпадают. Для этого достаточно рассмотреть изометрию  $f \circ g^{-1}$ , отображающую  $(n-1)$ -мерное пространство  $g(W_2)$  в  $f(W_2)$ , и продолжить её до изометрии пространства  $f(W_1)^\perp$ ; это можно сделать по предположению индукции.  $\square$

## Задачи

**5.1.** Докажите, что если скалярное произведение в пространстве  $V$  невырождено и  $W \subset V$ , то  $\dim V = \dim W + \dim W^\perp$  и  $(W^\perp)^\perp = W$ .

**5.2.** Пусть  $W_1$  и  $W_2$  — изометричные подпространства невырожденного пространства  $V$ . Докажите, что подпространства  $W_1^\perp$  и  $W_2^\perp$  тоже изометричны.

**5.3.** Пусть  $x$  и  $y$  — ненулевые изотропные векторы невырожденного пространства  $V$ . Докажите, что существует изометрия пространства  $V$ , переводящая вектор  $x$  в  $y$ .

**5.4.** Подпространство  $W \subset V$  называют изотропным, если ограничение скалярного произведения на него тождественно равно нулю. Изотропное подпространство, не содержащееся в другом изотропном подпространстве, называют максимальным изотропным подпространством. Докажите, что если  $W_1$  и  $W_2$  — максимальные изотропные подпространства невырожденного пространства  $V$ , то существует изометрия пространства  $V$ , переводящая  $W_1$  в  $W_2$ .

**5.5.** В пространстве  $V$  над  $\mathbb{R}$  задано скалярное произведение  $(x, y)$ , причём квадратичная форма  $q(x) = (x, x)$  невырождена. Докажите, что сигнатура  $q$  равна нулю тогда и только тогда, когда существует изотропное подпространство размерности  $(\dim V)/2$ .

## 8.6. Обобщённая обратная матрица. Матричные уравнения

### 8.6.1. Обобщённая обратная матрица

Матрицу  $X$  называют *обобщённой обратной* для матрицы  $A$  (не обязательно квадратной), если  $XAX = A$ ,  $AXA = A$  и матрицы  $AX$  и  $XA$  эрмитовы. Это понятие появилось независимо в работах [Mo5] и [Pe2]; эквивалентность определений Мура и Пенроуза показана в работе [Ra2]. Легко проверить, что для невырожденной матрицы  $A$  обобщённая обратная матрица совпадает с обратной матрицей.

**Теорема 8.6.1.1.** *Матрица  $X$  является обобщённой обратной для матрицы  $A$  тогда и только тогда, когда матрицы  $P = AX$  и  $Q = XA$  являются эрмитовыми проекторами на  $\operatorname{Im} A$  и  $\operatorname{Im} A^*$  соответственно.*

**Доказательство.** Предположим сначала, что матрицы  $P$  и  $Q$  являются эрмитовыми проекторами на  $\operatorname{Im} A$  и  $\operatorname{Im} A^*$  соответственно. Если  $v$  — произвольный вектор, то  $Av \in \operatorname{Im} A$ , поэтому  $PAv = Av$ , т.е.  $AXAv = Av$ . Кроме того,  $Xv \in \operatorname{Im} XA = \operatorname{Im} A^*$ , поэтому  $QXv = Xv$ , т.е.  $XXv = Xv$ .

Предположим теперь, что матрица  $X$  является обобщённой обратной для матрицы  $A$ . Тогда  $P^2 = (AXA)X = AX = P$  и  $Q^2 = (XAX)A = XA = Q$ ,

причём матрицы  $P$  и  $Q$  эрмитовы. Остаётся доказать, что  $\text{Im } P = \text{Im } A$  и  $\text{Im } Q = \text{Im } A^*$ . Так как  $P = AX$  и  $Q = Q^* = A^*X^*$ , то  $\text{Im } P \subset \text{Im } A$  и  $\text{Im } Q \subset \text{Im } A^*$ . С другой стороны,  $A = AXA = PA$  и  $A^* = A^*X^*A^* = Q^*A^* = QA^*$ , поэтому  $\text{Im } A \subset \text{Im } P$  и  $\text{Im } A^* \subset \text{Im } Q$ .  $\square$

**Теорема 8.6.1.2 (Мур–Пенроуз).** Для любой матрицы  $A$  существует единственная обобщённая обратная матрица  $X$ .

**Доказательство.** Если  $\text{rk } A = r$ , то матрицу  $A$  можно представить в виде произведения матриц  $C$  и  $D$  размера  $m \times r$  и  $r \times n$  соответственно (см. §2.4.2); при этом  $\text{Im } A = \text{Im } C$  и  $\text{Im } A^* = \text{Im } D^*$ ; ясно также, что матрицы  $C^*C$  и  $DD^*$  невырождены. Положим  $X = D^*(DD^*)^{-1}(C^*C)^{-1}C^*$ . Тогда  $AX = C(C^*C)^{-1}C^*$  и  $XA = D^*(DD^*)^{-1}D$ , т.е. матрицы  $AX$  и  $XA$  являются эрмитовыми проекторами на  $\text{Im } C = \text{Im } A$  и  $\text{Im } D^* = \text{Im } A^*$  соответственно (см. §4.7.3), а значит,  $X$  — обобщённая обратная матрица для матрицы  $A$ .

Предположим теперь, что  $X_1$  и  $X_2$  — обобщённые обратные матрицы для матрицы  $A$ . Тогда  $AX_1$  и  $AX_2$  — эрмитовы проекторы на  $\text{Im } A$ , поэтому  $AX_1 = AX_2$ . Аналогично  $X_1A = X_2A$ . Следовательно,  $X_1 = X_1(AX_1) = (X_1A)X_2 = X_2AX_2 = X_2$ .  $\square$

Обобщённую обратную матрицу для матрицы  $A$  будем обозначать  $A^+$ .

### 8.6.2. Решение несовместных систем уравнений

Обобщённая обратная матрица  $A^+$  применяется для решения систем линейных уравнений, как совместных, так и несовместных. Наиболее интересны её приложения к решению несовместных систем.

**Теорема 8.6.2.1.** Рассмотрим систему линейных уравнений  $Ax = b$ . Величина  $\|Ax - b\|$  минимальна для таких  $x$ , что  $Ax = AA^+b$ , а среди всех таких  $x$  наименьшую величину  $\|x\|$  имеет  $x_0 = A^+b$ .

**Доказательство.** Оператор  $P = AA^*$  — проектор, поэтому оператор  $I - P$  тоже проектор, причём  $\text{Im}(I - P) = \text{Ker } P$  (см. теорему 4.7.1.2). А так как оператор  $P$  эрмитов, то  $\text{Ker } P = (\text{Im } P)^\perp$ . Следовательно,  $\text{Im}(I - P) = \text{Ker } P = (\text{Im } P)^\perp = (\text{Im } A)^\perp$ , т.е. для любых векторов  $x$  и  $y$  векторы  $Ax$  и  $(I - AA^+)y$  перпендикулярны и

$$\|Ax + (I - AA^+)y\|^2 = \|Ax\|^2 + \|y - AA^+y\|^2.$$

Аналогично  $\|A^+x + (I - A^+A)y\|^2 = \|A^+x\|^2 + \|y - A^+Ay\|^2$ .

Так как  $Ax - b = A(x - A^+b) - (I - AA^+)b$ , то  $\|Ax - b\|^2 = \|Ax - AA^+b\|^2 + \|b - AA^+b\|^2 \geq \|b - AA^+b\|^2$ , причём равенство достигается тогда и только тогда, когда  $Ax = AA^+b$ . Если  $Ax = AA^+b$ , то  $\|x\|^2 = \|A^+b + (I - A^+A)x\|^2 = \|A^+b\|^2 + \|x - A^+Ax\|^2 \geq \|A^+b\|^2$ , причём равенство достигается тогда и только тогда, когда  $x = A^+Ax = A^+AA^+b = A^+b$ .  $\square$

**Замечание.** Равенство  $Ax = AA^+b$  эквивалентно равенству  $A^*Ax = A^*b$ . В самом деле, если  $Ax = AA^+b$ , то  $A^*b = A^*(A^+)^*A^*b = A^*AA^+b = A^*Ax$ , а если  $A^*Ax = A^*b$ , то  $Ax = AA^+Ax = (A^+)^*A^*Ax = (A^+)^*A^*b = AA^+b$ .

С помощью обобщённой обратной матрицы можно записать критерий совместности системы линейных уравнений и найти все её решения.

**Теорема 8.6.2.2.** Матричное уравнение  $AXB = C$  имеет решение тогда и только тогда, когда  $AA^+CB^+B = C$ . Решения этого уравнения имеют вид  $X = A^+CB^+ + Y - A^+AYBB^+$ , где  $Y$  — произвольная матрица.

**Доказательство.** Если  $AXB = C$ , то  $C = AXB = AA^+(AXB)B^+B = AA^+CB^+B$ . Наоборот, если  $C = AA^+CB^+B$ , то  $X_0 = A^+CB^+$  — частное решение уравнения  $AXB = C$ . Остаётся доказать, что общее решение уравнения  $AXB = 0$  имеет вид  $X = Y - A^+AYBB^+$ . Ясно, что  $A(Y - A^+AYBB^+)B = 0$ . С другой стороны, если  $AXB = 0$ , то  $X = Y - A^+AYBB^+$ , где  $Y = X$ .  $\square$

**Следствие.** а) Матричное уравнение  $AX = C$  разрешимо тогда и только тогда, когда  $AA^+C = C$ ; решения этого уравнения имеют вид  $X = A^+C + Y - A^+AY$ , где  $Y$  — произвольная матрица.

б) Матричное уравнение  $XA = C$  разрешимо тогда и только тогда, когда  $CA^+A = C$ ; решения этого уравнения имеют вид  $X = CA^+ + Y - YA^+A$ , где  $Y$  — произвольная матрица.

### 8.6.3. Матричные уравнения

**Теорема 8.6.3.1.** [Ro2] Пусть  $A \in M_{m,m}$ ,  $B \in M_{n,n}$  и  $C \in M_{m,n}$ .

а) Уравнение  $AX - XB = C$  имеет решение  $X \in M_{m,n}$  тогда и только тогда, когда матрицы  $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$  и  $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$  подобны.

б) Уравнение  $AX - YB = C$  имеет решение  $X, Y \in M_{m,n}$  тогда и только тогда, когда матрицы  $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$  и  $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$  одного ранга.

**Доказательство.** [F12] а) Пусть  $U = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$ . Предположим сначала, что указанные матрицы подобны. Рассмотрим для  $i = 0, 1$  отображения  $\varphi_i: M_{m,n} \rightarrow M_{m,n}$ , заданные формулами

$$\begin{aligned}\varphi_0(U) &= \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} U - U \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} AP - PA & AQ - QB \\ BR - RA & BS - SB \end{pmatrix}, \\ \varphi_1(U) &= \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} U - U \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} AP + CR - PA & AQ + CS - QB \\ BR - RA & BS - SB \end{pmatrix}.\end{aligned}$$

Уравнения  $FU = UF$  и  $GFG^{-1}U' = U'F$  имеют изоморфные пространства решений; этот изоморфизм задаётся формулой  $U = G^{-1}U'$ . Поэтому  $\dim \text{Ker } \varphi_0 = \dim \text{Ker } \varphi_1$ . Если  $U \in \text{Ker } \varphi_i$ , то  $BR = RA$  и  $BS = SB$ . Поэтому можно рассмотреть пространство

$$V = \{(R, S) \in M_{n,m+n} \mid BR = RA, BS = SB\}$$

и определить проекции  $\mu_i: \text{Ker } \varphi_i \rightarrow V$ , где  $\mu_i(U) = (R, S)$ . Легко проверить, что

$$\text{Ker } \mu_i = \left\{ \begin{pmatrix} P & Q \\ 0 & 0 \end{pmatrix} \mid AP = PA, AQ = QB \right\}.$$

Для  $\mu_0$  это очевидно, а для  $\mu_1$  следует из того, что  $CR = 0$  и  $CS = 0$ , так как  $R = 0$  и  $S = 0$ .

Докажем, что  $\text{Im } \mu_0 = \text{Im } \mu_1$ . Если  $(R, S) \in V$ , то  $\begin{pmatrix} 0 & 0 \\ R & S \end{pmatrix} \in \text{Ker } \varphi_0$ .

Поэтому  $\text{Im } \mu_0 = V$ , а значит,  $\text{Im } \mu_1 \subset \text{Im } \mu_0$ . С другой стороны,  $\dim \text{Im } \mu_0 + \dim \text{Ker } \mu_0 = \dim \text{Ker } \varphi_0 = \dim \text{Ker } \varphi_1 = \dim \text{Im } \mu_1 + \dim \text{Ker } \mu_1$ .

Матрица  $\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$  лежит в  $\text{Ker } \varphi_0$ , поэтому  $(0, -I) \in \text{Im } \mu_0 = \text{Im } \mu_1$ .

Следовательно, в  $\text{Ker } \varphi_1$  есть матрица вида  $\begin{pmatrix} P & Q \\ 0 & -I \end{pmatrix}$ . Значит,  $AQ + CS - QB = 0$ , где  $S = -I$ . Поэтому  $X = Q$  — решение уравнения  $AX - XB = C$ .

Обратно, если  $X$  — решение данного уравнения, то

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & AX \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & C + XB \\ 0 & B \end{pmatrix} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} A & C \\ 0 & B \end{pmatrix},$$

а значит,  $\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}^{-1} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ .

б) Предположим сначала, что указанные матрицы одного ранга.

Пусть  $U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}$  и  $W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$ . Рассмотрим для  $i = 0, 1$

отображения  $\psi_i: M_{m+n, 2(m+n)} \rightarrow M_{m+n, m+n}$ , заданные формулами

$$\begin{aligned}\psi_0(U, W) &= \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} U - W \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \\ &= \begin{pmatrix} AU_{11} - W_{11}A & AU_{12} - W_{12}B \\ BU_{21} - W_{21}A & BU_{22} - W_{22}B \end{pmatrix}, \\ \psi_1(U, W) &= \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} U - W \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \\ &= \begin{pmatrix} AU_{11} + CU_{21} - W_{11}A & AU_{12} + CU_{22} - W_{12}B \\ BU_{21} - W_{21}A & BU_{22} - W_{22}B \end{pmatrix}.\end{aligned}$$

Пространства решений уравнений  $FU = WF$  и  $GFG^{-1}U' = W'F$  изоморфны; этот изоморфизм задаётся формулами  $U = G^{-1}U'$  и  $W = G^{-1}W'$ . Поэтому  $\dim \text{Ker } \varphi_0 = \dim \text{Ker } \varphi_1$ .

Рассмотрим пространство

$$Z = \{(U_{21}, U_{22}, W_{21}, W_{22}) \mid BU_{21} = W_{21}A, BU_{22} = W_{22}B\}$$

и определим отображения  $v_i: \text{Ker } \varphi_i \rightarrow Z$ , где  $v_i(U, W) = (U_{21}, U_{22}, W_{21}, W_{22})$ . Тогда  $\text{Im } v_1 \subset \text{Im } v_0 = Z$  и  $\text{Ker } v_1 = \text{Ker } v_0$ . Поэтому  $\text{Im } v_1 = \text{Im } v_0$ . Матрица  $(U, W)$ , где  $U = W = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ , лежит в  $\text{Ker } \psi_0$ . Поэтому в  $\text{Ker } \psi_1$  тоже существует элемент, для которого  $U_{22} = -I$ . Для этого элемента равенство  $AU_{12} + CU_{22} = W_{12}B$  эквивалентно равенству  $AU_{21} - W_{12}B = C$ .

Обратно, если решение  $X, Y$  данного уравнения существует, то

$$\begin{pmatrix} I & -Y \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & AX - YB \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}.$$

□

### Задачи

**6.1.** Докажите, что если  $C = AX = YB$ , то существует такая матрица  $Z$ , что  $C = AZB$ .

**6.2.** Докажите, что любое решение системы матричных уравнений  $AX = 0, XB = 0$  имеет вид  $X = (I - A^+A)Y(I - BB^+)$ , где  $Y$  — произвольная матрица.



**6.3.** Докажите, что система уравнений  $AX = C$ ,  $XB = D$  имеет решение тогда и только тогда, когда каждое из уравнений  $AX = C$  и  $XB = D$  имеет решение и  $AD = CB$ .

**6.4.** Операторы  $A$  и  $B$  аннулируют пространство  $V^\perp$ , а их ограничения на пространство  $V$  являются взаимно обратными операторами. Докажите, что  $A^+ = B$  и  $B^+ = A$ .

**6.5.** Докажите, что если  $A$  — ненулевая нильпотентная матрица, то матрицы  $A$  и  $A^+$  не коммутируют.

**6.6.** [Wi2] Пусть  $A \in M_{m,m}$ ,  $B \in M_{n,n}$  и  $C \in M_{m,n}$ . Докажите, что уравнение  $X - AXB = C$  имеет решение  $X \in M_{m,n}$  тогда и только тогда, когда матрицы  $A + xI_m$  и  $I_n + xB$  одного ранга.

## 8.7. Ганкелевы и тёплицевы матрицы

Матрицу  $\|a_{ij}\|_0^{n-1}$ , где  $a_{ij} = s_{i+j}$  для некоторых комплексных чисел  $s_0, s_1, \dots, s_{2n-2}$ , называют *ганкелевой*. Ганкелева матрица всегда симметрическая. Ганкелева матрица является эрмитовой тогда и только тогда, когда все числа  $s_0, \dots, s_{2n-2}$  вещественные. *Ганкелевой формой* называют квадратичную форму  $\sum_{i,j=0}^{n-1} s_{i+j} x_i x_j$ , где  $s_0, \dots, s_{2n-2}$  — вещественные числа.

Матрицу  $\|a_{ij}\|_0^{n-1}$ , где  $a_{ij} = c_{i-j}$  для некоторых комплексных чисел  $c_0, c_{\pm 1}, \dots, c_{\pm(n-1)}$ , называют *тёплицевой*. Тёплицева матрица является эрмитовой тогда и только тогда, когда  $c_{-k} = \bar{c}_k$  для всех  $k = 0, 1, \dots, n-1$ . Эрмитова тёплицева матрица задаёт эрмитову *тёплицеву форму*  $\sum_{i,j=0}^{n-1} c_{i-j} z_i \bar{z}_j$ . Примером эрмитовой тёплицевой формы служит форма

$$\|z_0 e_0 + \dots + z_{n-1} e_{n-1}\|^2 = \sum_{i,j=0}^{n-1} (e_i, e_j) z_i \bar{z}_j.$$

Если векторы  $e_0, \dots, e_{n-1}$  линейно независимы, то эта эрмитова форма положительно определённая.

### 8.7.1. Корни многочленов

Пусть  $P_n(t) = a_0 t^n + a_1 t^{n-1} + \dots + a_n$  — многочлен с вещественными коэффициентами,  $\alpha_1, \dots, \alpha_n$  — его корни (с учётом кратностей). Тогда числа  $s_0 = n$ ,  $s_k = \alpha_1^k + \dots + \alpha_n^k$  вещественные, поскольку суммы степеней выражаются через элементарные симметрические функции (см. §1.4.1).

**Теорема 8.7.1.1.** *Количество различных вещественных корней многочлена  $P_n(t)$  равно сигнатуре ганкелевой квадратичной формы  $\sum_{i,j=0}^{n-1} s_{i+j} x_i x_j$ .*

**Доказательство.** Рассматриваемая квадратичная форма представляется в виде суммы квадратов следующим образом:

$$\begin{aligned} \sum_{i,j=0}^{n-1} s_{i+j} x_i x_j &= \sum_{i,j=0}^{n-1} \sum_{k=1}^n \alpha_k^{p+q} x_p x_q = \sum_{i,j=0}^{n-1} \sum_{k=1}^n \alpha_k^p x_p \alpha_k^q x_q = \\ &= \sum_{k=1}^n \sum_{i,j=0}^{n-1} \alpha_k^p x_p \alpha_k^q x_q = \\ &= \sum_{k=1}^n (x_0 + \alpha_k x_1 + \alpha_k^2 x_2 + \cdots + \alpha_k^{n-1} x_{n-1})^2. \end{aligned}$$

Слагаемые, относящиеся к одному и тому же кратному корню, сгруппируем вместе. В результате получим  $r$  слагаемых, относящихся к  $r$  вещественным корням, и  $2s$  слагаемых, относящихся к  $s$  парам комплексно сопряжённых корней. Для вещественных корней мы уже получаем то, что нужно. Для пары комплексно сопряжённых корней применим следующее преобразование:

$$(f(x) + ig(x))^2 + (f(x) - ig(x))^2 = 2(f(x))^2 - 2(g(x))^2.$$

Исходные линейные формы  $x_0 + \alpha_k x_1 + \alpha_k^2 x_2 + \cdots + \alpha_k^{n-1} x_{n-1}$  линейно независимы, поэтому после этих преобразований мы тоже получим линейно независимые формы. Таким образом, рассматриваемая квадратичная форма приводится к виду

$$y_1^2 + \cdots + y_r^2 + (y_{r+1}^2 - y_{r+2}^2) + \cdots + (y_{r+2s-1}^2 - y_{r+2s}^2).$$

Сигнатура этой квадратичной формы равна  $r$ .  $\square$

Аналогичное утверждение имеет место и для корней многочлена  $P_n(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n$  с комплексными коэффициентами, удовлетворяющими соотношению  $a_k = \bar{a}_{n-k}$ . Для такого многочлена получаем  $\overline{P_n(z)} = \bar{z}^n P_n(1/\bar{z})$ , поэтому его корни, не лежащие на окружности  $|z| = 1$ , разбиваются на пары вида  $\{\alpha, \bar{\alpha}^{-1}\}$ .

**Теорема 8.7.1.2 (Герглouc–Крейн).** *Количество различных корней многочлена  $P_n(z)$  равно сигнатуре тѐплицевой эрмитовой формы  $\sum_{i,j=0}^{n-1} c_{i-j} z_i \bar{z}_j$ .*

**Доказательство.** Прежде всего нужно убедиться, что указанная форма эрмитова, т.е.  $s_{-k} = \bar{s}_k$ . Введём обозначение  $\alpha^* = 1/\bar{\alpha}$ , где  $\alpha$  — комплексное число. Пусть  $\alpha_1, \dots, \alpha_r$  — корни многочлена  $P_n(z)$ , лежащие на единичной окружности,  $n_1, \dots, n_r$  — их кратности;  $\alpha_{r+1}, \dots, \alpha_{r+s}, \alpha_{r+1}^*, \dots, \alpha_{r+s}^*$  — остальные корни многочлена  $P_n(z)$ ,  $n_{r+1}, \dots, n_{r+s}$  — их кратности. Тогда

$$\begin{aligned} s_{-k} &= \sum_{j=1}^r n_j \alpha_j^{-k} + \sum_{j=r+1}^s n_j (\alpha_j^{-k} + (\alpha_j^*)^{-k}) = \\ &= \sum_{j=1}^r n_j \bar{\alpha}_j^k + \sum_{j=r+1}^s n_j ((\bar{\alpha}_j^*)^k + \alpha_j^k) = \bar{s}_k. \end{aligned}$$

Записав  $s_{p-q}$  в виде

$$s_{p-q} = \sum_{j=1}^r n_j \alpha_j^p \bar{\alpha}_j^q + \sum_{j=r+1}^s n_j (\alpha_j^p (\bar{\alpha}_j^*)^q + \bar{\alpha}_j^p \alpha_j^{*q}),$$

получим

$$\sum_{p,q=0}^{n-1} s_{p-q} z_p \bar{z}_q = \sum_{j=1}^r n_j |X_j(z)|^2 + \sum_{j=r+1}^s n_j (Y_j(z) \bar{Z}_j(z) + \bar{Y}_j(z) Z_j(z)),$$

где

$$\begin{aligned} X_j(z) &= z_0 + z_1 \alpha_j + \dots + z_{n-1} \alpha_j^{n-1}, \\ Y_j(z) &= z_0 + z_1 \alpha_j + \dots + z_{n-1} \alpha_j^{n-1}, \\ Z_j(z) &= z_0 + z_1 \alpha_j^* + \dots + z_{n-1} \alpha_j^{*n-1}. \end{aligned}$$

Заменяя каждое выражение  $Y_j \bar{Z}_j + \bar{Y}_j Z_j$  на  $\frac{1}{2}(|Y_j + Z_j|^2 - |Y_j - Z_j|^2)$ , получим требуемое представление эрмитовой формы.  $\square$

### 8.7.2. Ганкелевы матрицы и рациональные функции

Рассмотрим правильную рациональную функцию

$$R(z) = \frac{a_1 z^{m-1} + \dots + a_m}{b_0 z^m + b_1 z^{m-1} + \dots + b_m},$$

где  $b_0 \neq 0$ . Эту функцию можно разложить в ряд

$$R(z) = s_0 z^{-1} + s_1 z^{-2} + s_2 z^{-3} + \dots$$

При этом

$$\begin{aligned} b_0 s_0 &= a_1, \\ b_0 s_1 + b_1 s_0 &= a_2, \\ b_0 s_2 + b_1 s_1 + b_2 s_0 &= a_3, \\ &\dots\dots\dots \\ b_0 s_{m-1} + \dots + b_{m-1} s_0 &= a_m. \end{aligned} \quad (1)$$

Кроме того,  $b_0 s_q + \dots + b_m s_{q-m} = 0$  при  $q \geq m$ . Таким образом, при всех  $q \geq m$  имеет место равенство

$$s_q = \alpha_1 s_{q-1} + \dots + \alpha_m s_{q-m}, \quad (2)$$

где  $\alpha_i = -b_i/b_0$ .

Рассмотрим бесконечную ганкелеву матрицу

$$S = \begin{pmatrix} s_0 & s_1 & s_2 & \dots \\ s_1 & s_2 & s_3 & \dots \\ s_2 & s_3 & s_4 & \dots \\ \dots\dots\dots \end{pmatrix}.$$

Соотношение (2) означает, что  $(m+1)$ -я строка матрицы  $S$  является линейной комбинацией первых  $m$  строк (с коэффициентами  $\alpha_1, \dots, \alpha_m$ ). Отбросив в каждой из рассматриваемых строк первый элемент, получим, что  $(m+2)$ -я строка матрицы  $S$  является линейной комбинацией предшествующих ей  $m$  строк, а значит, и линейной комбинацией первых  $m$  строк. Продолжив эти рассуждения, получим, что любая строка матрицы  $S$  выражается через  $m$  первых строк, т.е.  $\text{rk } S \leq m$ .

Таким образом, если ряд

$$s_0 z^{-1} + s_1 z^{-2} + s_2 z^{-3} + \dots \quad (3)$$

соответствует рациональной функции  $R(z)$ , то построенная по числам  $s_0, s_1, \dots$  ганкелева матрица  $S$  имеет конечный ранг. Предположим теперь, что ганкелева матрица  $S$  имеет конечный ранг  $m$ ; построим по матрице  $S$  ряд (3). Докажем, что этот ряд соответствует рациональной функции. Строки  $1, 2, \dots, m+1$  матрицы  $S$  линейно зависимы, поэтому строка с номером  $h+1$ , где  $h \leq m$ , линейно выражается через строки  $1, 2, \dots, h$ . Как было показано, в этом случае все строки матрицы  $S$  выражаются через первые  $h$  строк. Следовательно,  $h = m$ . Таким образом, числа  $s_i$  связаны соотношением (2) при всех  $q \geq m$ . Коэффициенты  $\alpha_i$  в этом соотношении позволяют определить числа  $b_0 = 1, b_1 = \alpha_1, \dots$ ,

$b_m = \alpha_m$ . Затем с помощью соотношения (1) можно определить числа  $a_1, \dots, a_m$ . Для определённых таким образом чисел  $a_i, b_j$  выполняется равенство

$$\frac{s_0}{z} + \frac{s_1}{z^2} + \dots = \frac{a_1 z^{m-1} + \dots + a_m}{b_0 z^m + b_1 z^{m-1} + \dots + b_m},$$

т.е.  $R(z)$  — рациональная функция.

## 8.8. Функции от матриц. Дифференцирование матриц

### 8.8.1. Экспонента

По аналогии с экспонентой числа можно определить экспоненту матрицы  $A$  как сумму ряда  $\sum_{k=0}^{\infty} \frac{A^k}{k!}$ . Докажем, что этот ряд сходится. Если  $A$  и  $B$  — квадратные матрицы порядка  $n$  и  $|a_{ij}| \leq a, |b_{ij}| \leq b$ , то элементы матрицы  $AB$  по модулю не превосходят  $nab$ . Поэтому элементы матрицы  $A^k$  по модулю не превосходят  $n^{k-1}a^k = (na)^k/n$ . А так как  $\frac{1}{n} \sum_{k=0}^{\infty} \frac{(na)^k}{k!} = \frac{1}{n} e^{na}$ , то ряд  $\sum_{k=0}^{\infty} \frac{A^k}{k!}$  сходится к некоторой матрице  $e^A = \exp A$ ; эту матрицу называют *экспонентой* матрицы  $A$ .

Если  $A_1 = P^{-1}AP$ , то  $A_1^k = P^{-1}A^kP$ . Следовательно,  $\exp(P^{-1}AP) = P^{-1}(\exp A)P$ . Поэтому вычисление экспоненты произвольной матрицы сводится к вычислению экспоненты жордановой клетки.

Пусть  $J = \lambda I + N$  — жорданова клетка порядка  $n$ . Тогда  $(\lambda I + N)^k = \sum_{m=0}^k \binom{k}{m} \lambda^{k-m} N^m$ . Поэтому  $\exp(tJ) = \sum_{k=0}^{\infty} \frac{t^k J^k}{k!} = \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{t^k \binom{k}{m} \lambda^{k-m} N^m}{k!} = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} \frac{(\lambda t)^{k-m}}{(k-m)!} \frac{t^m N^m}{m!} = \sum_{m=0}^{\infty} \frac{t^m}{m!} e^{\lambda t} N^m = \sum_{m=0}^{n-1} \frac{t^m}{m!} e^{\lambda t} N^m$ , так как  $N^m = 0$  при  $m \geq n$ .

Приведя матрицу  $A$  к жордановой нормальной форме, получим следующее утверждение.

**Теорема 8.8.1.1.** Если минимальный многочлен матрицы  $A$  равен  $(x - \lambda_1)^{n_1} \dots (x - \lambda_k)^{n_k}$ , то элементы матрицы  $e^{At}$  имеют вид  $p_1(t)e^{\lambda_1 t} + \dots + p_k(t)e^{\lambda_k t}$ , где  $p_i(t)$  — многочлен степени не более  $n_i - 1$ .

**Теорема 8.8.1.2.**  $\det(e^A) = e^{\text{tr } A}$ .

**Доказательство.** Можно считать, что  $A$  — верхняя треугольная матрица с элементами  $\lambda_1, \dots, \lambda_n$  на диагонали. Тогда  $A^k$  — верхняя треугольная матрица с элементами  $\lambda_1^k, \dots, \lambda_n^k$  на диагонали. Поэтому  $e^A$  — верхняя треугольная матрица с элементами  $\exp \lambda_1, \dots, \exp \lambda_n$  на диагонали.  $\square$

### 8.8.2. Дифференцирование матриц

Рассмотрим семейство матриц  $X(t) = \|x_{ij}(t)\|_1^n$ , элементы которых являются дифференцируемыми функциями от  $t$ . Пусть  $\dot{X}(t) = \frac{dX(t)}{dt}$  — поэлементная производная матричной функции  $X(t)$ .

**Теорема 8.8.2.1.**  $(XY)^\cdot = \dot{X}Y + X\dot{Y}$ .

**Доказательство.** Если  $Z = XY$ , то  $z_{ij} = \sum_k x_{ik}y_{kj}$ , а значит,  $\dot{z}_{ij} = \sum_k \dot{x}_{ik}y_{kj} + \sum_k x_{ik}\dot{y}_{kj}$ . Поэтому  $\dot{Z} = \dot{X}Y + X\dot{Y}$ .  $\square$

**Теорема 8.8.2.2.** а)  $(X^{-1})^\cdot = -X^{-1}\dot{X}X^{-1}$ .

б)  $\text{tr}(X^{-1}\dot{X}) = -\text{tr}((X^{-1})^\cdot X)$ .

**Доказательство.** а) С одной стороны,  $(X^{-1}X)^\cdot = \dot{I} = 0$ . С другой стороны,  $(X^{-1}X)^\cdot = (X^{-1})^\cdot X + X^{-1}\dot{X}$ . Поэтому  $(X^{-1})^\cdot X = -X^{-1}\dot{X}$ , а значит,  $(X^{-1})^\cdot = -X^{-1}\dot{X}X^{-1}$ .

б) Так как  $\text{tr}(X^{-1}X) = n$ , то  $0 = [\text{tr}(X^{-1}X)]^\cdot = \text{tr}((X^{-1})^\cdot X) + \text{tr}(X^{-1}\dot{X})$ .

$\square$

**Теорема 8.8.2.3.**  $(e^{At})^\cdot = Ae^{At}$

**Доказательство.** Так как ряд  $\sum_{k=0}^{\infty} \frac{(tA)^k}{k!}$  сходится абсолютно, то  $\frac{d}{dt}(e^{At}) = \sum_{k=0}^{\infty} \frac{d}{dt} \left( \frac{(tA)^k}{k!} \right) = \sum_{k=0}^{\infty} \frac{kt^{k-1}A^k}{k!} = A \sum_{k=1}^{\infty} \frac{(tA)^{k-1}}{(k-1)!} = Ae^{At}$ .  $\square$

Однородный многочлен  $p(x, y) = \sum a_k x^{n-k} y^k$  можно отождествить с вектором  $(a_0, \dots, a_n)^T$ . Легко видеть, что многочлен

$$p_\varphi(x, y) = p(x \cos \varphi + y \sin \varphi, -x \sin \varphi + y \cos \varphi)$$

получается из многочлена  $p(x, y)$  некоторым линейным преобразованием:  $p_\varphi(x, y) = L_\varphi p(x, y)$ .

**Теорема 8.8.2.4.**  $L_\varphi = e^{\varphi A_n}$ , где

$$A_n = \begin{pmatrix} 0 & n & & & \\ -1 & 0 & n-1 & & \\ & -2 & n-2 & & \\ \dots & \dots & \dots & \dots & \dots \\ & & & -(n-1) & 0 & 1 \\ & & & & -n & 0 \end{pmatrix}.$$

**Доказательство.** Продифференцировав по  $\varphi$  выражение  $a_k(x \cos \varphi + y \sin \varphi)^{n-k}(-x \sin \varphi + y \cos \varphi)^k$ , получим  $\frac{d}{d\varphi} L_\varphi = A_n L_\varphi$ . Остаётся воспользоваться теоремой 8.8.2.3 и заметить, что  $L_0 = I$ .  $\square$

### 8.8.3. Системы дифференциальных уравнений

Систему  $n$  линейных дифференциальных уравнений первого порядка для  $n$  неизвестных можно записать в виде  $\dot{X} = AX$ , где  $X$  — столбец длиной  $n$ ,  $A$  — матрица порядка  $n$ . Если матрица  $A$  постоянна, то  $X(t) = e^{At}C$  — решение этого уравнения с начальным условием  $X(0) = C$  (см. теорему 8.8.2.3); решение данного уравнения с данным начальным условием единственно. Общий вид элементов матрицы  $e^{At}$  даёт теорема 8.8.1.1; используя эту теорему, приходим к следующему утверждению.

**Теорема 8.8.3.1.** Если минимальный многочлен матрицы  $A$  равен  $(\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_k)^{n_k}$ , то решения  $x_1(t), \dots, x_n(t)$  уравнения  $\dot{X} = AX$  (т.е. координаты вектора  $X$ ) имеют вид  $x_i(t) = p_{i1}(t)e^{\lambda_1 t} + \dots + p_{ik}(t)e^{\lambda_k t}$ , где  $p_{ij}(t)$  — многочлен, степень которого не превосходит  $n_j - 1$ .

Непосредственной подстановкой легко проверить, что  $X(t) = e^{At}Ce^{Bt}$  — решение уравнения  $\dot{X} = AX + XB$  с начальным условием  $X(0) = C$ .

**Теорема 8.8.3.2.** Пусть  $A(t)$  и  $X(t)$  — матрицы порядка  $n$ , причём  $X(t)$  — решение уравнения  $\dot{X} = A(t)X$ . Тогда

$$\det X = \exp \left( \int_0^t \operatorname{tr} A(s) ds \right) \det X(0).$$

**Доказательство.** Согласно задаче 8.6, а)  $(\det X)' = (\det X)(\operatorname{tr} \dot{X} X^{-1})$ . В нашем случае  $\dot{X} X^{-1} = A(t)$ . Следовательно, функция  $y(t) = \det X(t)$  удовлетворяет условию  $(\ln y)' = \dot{y}/y = \operatorname{tr} A(t)$ . Поэтому  $y(t) = c \exp \left( \int_0^t \operatorname{tr} A(s) ds \right)$ , где  $c = y(0) = \det X(0)$ .  $\square$

### 8.8.4. Вронскиан

Пусть функции  $f_1(t), \dots, f_n(t)$  обладают на интервале  $(a, b)$  непрерывными производными порядка  $n - 1$ . Их *вронскианом*, или *определителем Вронского*, называют определитель

$$W(f_1, \dots, f_n)(t) = \begin{vmatrix} f_1 & f_2 & \dots & f_n \\ f_1' & f_2' & \dots & f_n' \\ \dots & \dots & \dots & \dots \\ f_1^{(n-1)} & f_2^{(n-1)} & \dots & f_n^{(n-1)} \end{vmatrix}.$$

Функции  $f_1, \dots, f_n$  называют *линейно зависимыми* на  $(a, b)$ , если существуют константы  $c_1, \dots, c_n$ , не все равные нулю, для которых  $c_1 f_1(t) + \dots + c_n f_n(t) = 0$  при всех  $t \in (a, b)$ . Ясно, что если функции  $f_1, \dots, f_n$  линейно зависимы, то  $W(f_1, \dots, f_n) = 0$ . Если на функции  $f_i$  не наложено никаких дополнительных ограничений, то обратное утверждение неверно. Например, пусть  $f_1(t) = t^2$  и  $f_2(t) = t|t|$ . Тогда на  $\mathbb{R}$  функции  $f_1$  и  $f_2$  линейно независимы, но  $W(f_1, f_2) = 0$ . Тем не менее, для многих классов функций из тождественного обращения в нуль вронскиана вытекает их линейная зависимость.

**Пример 8.8.4.1.** Пусть  $p_0, p_1, \dots, p_{n-1}$  — непрерывные функции,  $f_1, \dots, f_n$  — решения однородного дифференциального уравнения

$$y^{(n)} + p_{n-1}(t)y^{(n-1)} + \dots + p_1(t)y' + p_0(t)y = 0; \quad (1)$$

$W(t) = W(f_1, \dots, f_n)(t)$  — их вронскиан. Тогда

$$W(t_1) = W(t_0) \exp \int_{t_1}^{t_0} p_{n-1}(t) dt \quad (2)$$

и функции  $f_1, \dots, f_n$  линейно зависимы тогда и только тогда, когда их вронскиан обращается в нуль в некоторой точке.

**Доказательство.** Дифференциальное уравнение (1) сводится к системе линейных однородных дифференциальных уравнений  $\dot{X} = AX$ , где  $X$  — столбец  $(y, y', \dots, y^{(n-1)})$  и

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ -p_0 & -p_1 & -p_2 & -p_3 & \dots & -p_{n-1} \end{pmatrix}.$$



Для определителя матрицы, составленной из столбцов—решений, выполняется равенство

$$W(t) = W(0) \exp \int_0^t (-p_{n-1}(\tau)) d\tau$$

(теорема 8.8.3.2). Из этого следует формула (2).

Если вронскиан функций  $f_1, \dots, f_n$  обращается в нуль в некоторой точке  $t_0$ , то векторы  $(f_i(t_0), f'_i(t_0), \dots, f_i^{(n-1)}(t_0))$ ,  $i = 1, \dots, n$ , линейно зависимы. Поэтому можно выбрать константы  $c_1, \dots, c_n$  (не все равные нулю) так, что функция  $\varphi(t) = c_1 f_1(t) + \dots + c_n f_n(t)$  и её производные  $\varphi', \dots, \varphi^{(n-1)}$  обращаются в нуль в точке  $t_0$ . При этом  $\varphi(t)$  — решение уравнения (1), поскольку линейная комбинация решений снова будет решением. Согласно теореме единственности решение  $f(t)$  уравнения (1) однозначно задаётся числами  $f(t_0), f'(t_0), \dots, f^{(n-1)}(t_0)$ . Поэтому  $\varphi(t) \equiv 0$ .  $\square$

Из тождественного обращения в нуль на интервале  $(a, b)$  вронскиана функций можно сделать вывод о линейной зависимости этих функций на некотором меньшем интервале  $(c, d) \subset (a, b)$ . Для этого нам понадобится следующее вспомогательное утверждение.

**Лемма.** Пусть функция  $g$  дифференцируема  $n - 1$  раз. Тогда

$$W(gf_1, \dots, gf_n) = g^n W(f_1, \dots, f_n).$$

**Доказательство.** Если  $g(t) = 0$ , то требуемое равенство очевидно. Будем считать, что  $g \neq 0$ . Элементы второй строки определителя  $W(gf_1, \dots, gf_n)$  имеют вид  $gf'_i + g'f_i$ . Поэтому после вычитания из второй строки первой строки, умноженной на  $g'/g$ , получится вторая строка определителя  $W(f_1, \dots, f_n)$ , умноженная на  $g$ . Элементы третьей строки имеют вид  $gf''_i + 2g'f'_i + g''f_i$ , поэтому из неё нужно вычесть первую строку, умноженную на  $g''/g$ , и новую вторую строку, умноженную на  $g'/g$ , и т.д.  $\square$

**Следствие.** Если  $f_1 \neq 0$ , то

$$W(f_1, \dots, f_n) = f_1^n W\left(\left(\frac{f_2}{f_1}\right)', \dots, \left(\frac{f_n}{f_1}\right)'\right).$$

**Доказательство.** Согласно лемме

$$\begin{aligned} W(f_1, \dots, f_n) &= f_1^n W\left(1, \frac{f_2}{f_1}, \dots, \frac{f_n}{f_1}\right) = \\ &= f_1^n \begin{vmatrix} 1 & f_2/f_1 & \dots & f_n/f_1 \\ 0 & (f_2/f_1)' & \dots & (f_n/f_1)' \\ \dots & \dots & \dots & \dots \\ 0 & (f_2/f_1)^{(n-1)} & \dots & (f_n/f_1)^{(n-1)} \end{vmatrix} = \\ &= f_1^n W\left(\left(\frac{f_2}{f_1}\right)', \dots, \left(\frac{f_n}{f_1}\right)'\right). \end{aligned}$$

□

**Теорема 8.8.4.1.** Если  $W(f_1, \dots, f_n) = 0$  на интервале  $(a, b)$ , то функции  $f_1, \dots, f_n$  линейно зависимы на некотором интервале  $(c, d) \subset (a, b)$ .

**Доказательство.** [Kr2] Применим индукцию по  $n$ . При  $n = 1$  из равенства  $W(f_1) = 0$  следует, что  $f_1 = 0$ . Пусть  $n > 1$ . Если  $f_1 = 0$  на  $(a, b)$ , то линейная зависимость очевидна. Поэтому будем считать, что можно выбрать интервал, на котором  $f_1$  не обращается в нуль. На этом интервале

$$W\left(\left(\frac{f_2}{f_1}\right)', \dots, \left(\frac{f_n}{f_1}\right)'\right) = \frac{1}{f_1^n} W(f_1, \dots, f_n) = 0.$$

По предположению индукции можно выбрать интервал, на котором функции  $(f_2/f_1)', \dots, (f_n/f_1)'$  линейно зависимы, т.е.

$$\left(\frac{c_2 f_2 + \dots + c_n f_n}{f_1}\right)' \equiv 0.$$

Следовательно,  $c_2 f_2 + \dots + c_n f_n \equiv c f_1$ . □

## Задачи

**8.1.** Пусть  $A = \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix}$ . Вычислите  $e^A$ .

**8.2.** а) Докажите, что если  $[A, B] = 0$ , то  $e^{A+B} = e^A e^B$ .

б) Докажите, что если  $e^{(A+B)t} = e^{At} e^{Bt}$  при всех  $t$ , то  $[A, B] = 0$ .

**8.3.** Докажите, что для любой унитарной матрицы  $U$  существует такая эрмитова матрица  $H$ , что  $U = e^{iH}$ .

**8.4.** а) Докажите, что если вещественная матрица  $X$  кососимметрична, то матрица  $e^X$  ортогональна.

б) Докажите, что любую ортогональную матрицу  $U$  с определителем 1 можно представить в виде  $e^X$ , где  $X$  — вещественная кососимметрическая матрица.

**8.5.** а) Пусть  $A$  — вещественная матрица. Докажите, что  $\det e^A = 1$  тогда и только тогда, когда  $\operatorname{tr} A = 0$ .

б) Пусть  $B$  — вещественная матрица и  $\det B = 1$ . Всегда ли найдётся такая вещественная матрица  $A$ , что  $B = e^A$ ?

**8.6.** а) Докажите, что  $(\det A)' = \operatorname{tr}(\dot{A} \operatorname{adj} A^T) = (\det A) \operatorname{tr}(\dot{A} A^{-1})$ .

б) Пусть  $A$  — матрица порядка  $n$ . Докажите, что  $\operatorname{tr}(A(\operatorname{adj} A^T)') = (n-1) \operatorname{tr}(\dot{A} \operatorname{adj} A^T)$ .

**8.7.** Докажите, что если матрица  $A$  нильпотентна, то матрица  $e^A$  унитарна. Верно ли обратное?

**8.8.** Докажите, что

$$e^{tX_1} e^{tX_2} \dots e^{tX_n} = e^{t \sum X_i + \frac{t^2}{2} \sum_{i < j} [X_i, X_j]} + O(t^3).$$

**8.9.** Пусть  $\partial_i = \partial/\partial x_i$ . Докажите, что

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ \partial_1 & \partial_2 & \dots & \partial_n \\ \dots & \dots & \dots & \dots \\ \partial_1^{n-1} & \partial_2^{n-1} & \dots & \partial_n^{n-1} \end{vmatrix} \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = 1!2! \dots n!.$$

**8.10.** [Ai] Рассмотрим отображение  $F: M_{n,n} \rightarrow M_{n,n}$ . Пусть  $\Omega F(X) = \|\omega_{ij}(X)\|_1^n$ , где  $\omega_{ij}(X) = \frac{\partial}{\partial x_{ji}} \operatorname{tr} F(X)$ . Докажите, что если  $F(X) = X^m$ , где  $m$  — целое число, то  $\Omega F(X) = mX^{m-1}$ .

## 8.9. Пары Лакса и интегрируемые системы

### 8.9.1. Уравнение Лакса

Рассмотрим систему дифференциальных уравнений  $\dot{x}(t) = f(x, t)$ , где  $x = (x_1, \dots, x_n)$ . Функцию  $F(x_1, \dots, x_n)$ , отличную от константы, называют *первым интегралом* этой системы, если  $\frac{d}{dt} F(x_1(t), \dots, x_n(t)) = 0$

для любого решения  $(x_1(t), \dots, x_n(t))$  данной системы. Первый интеграл позволяет понизить порядок системы на 1.

Пусть  $A$  и  $L$  — квадратные матрицы, элементы которых зависят от  $x_1, \dots, x_n$ . Дифференциальное уравнение  $\dot{L} = AL - LA$  называют *уравнением Лакса*.

**Теорема 8.9.1.1.** *Функции  $f_k(x_1, \dots, x_n) = \text{tr}(L^k)$  являются первыми интегралами уравнения Лакса (если они не константы).*

**Доказательство.** Пусть  $B(t)$  — решение уравнения  $\dot{B} = -BA$  с начальным условием  $B(0) = I$ . Тогда  $\det B(t) = \exp\left(\int_0^t A(s) ds\right) \neq 0$  (см. теорему 8.8.3.2) и

$$\begin{aligned} (BLB^{-1})' &= \dot{B}LB^{-1} + B\dot{L}B^{-1} + BL(B^{-1})' = \\ &= -BALB^{-1} + B(AL - LA)B^{-1} + BLB^{-1}(BA)B^{-1} = 0. \end{aligned}$$

Поэтому жорданова нормальная форма матрицы  $L$  не зависит от  $t$ , а значит, её собственные значения тоже не зависят от  $t$ .  $\square$

Представление систем дифференциальных уравнений в виде уравнения Лакса является важным способом нахождения первых интегралов гамильтоновых систем дифференциальных уравнений. Например, в виде уравнения Лакса легко записать *уравнение Эйлера*  $\dot{M} = M \times \omega$ , описывающее движение твёрдого тела с неподвижной точкой. Для этого нужно взять

$$L = \begin{pmatrix} 0 & -M_3 & M_2 \\ M_3 & 0 & -M_1 \\ -M_2 & M_1 & 0 \end{pmatrix} \quad \text{и} \quad A = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}.$$

Первым интегралом этого уравнения является  $\text{tr } L^2 = -2(M_1^2 + M_2^2 + M_3^2)$ .

### 8.9.2. Цепочка Тоды

Более нетривиальным примером является *цепочка Тоды*

$$\ddot{x}_i = -\frac{\partial}{\partial x_i} U, \quad \text{где} \quad U = \exp(x_1 - x_2) + \dots + \exp(x_{n-1} - x_n).$$

Эту систему уравнений можно переписать в виде уравнения Лакса с

$$L = \begin{pmatrix} b_1 & a_1 & 0 & \dots & \dots & 0 \\ a_1 & b_2 & a_2 & \dots & \dots & 0 \\ 0 & a_2 & b_3 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{n-2} & a_{n-1} \\ 0 & 0 & 0 & \dots & a_{n-1} & b_n \end{pmatrix} \quad \text{и} \quad A = \begin{pmatrix} 0 & a_1 & 0 & \dots & \dots & 0 \\ -a_1 & 0 & a_2 & \dots & \dots & 0 \\ 0 & -a_2 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & a_{n-1} \\ 0 & 0 & 0 & \dots & -a_{n-1} & 0 \end{pmatrix},$$

где  $2a_k = \exp \frac{1}{2}(x_k - x_{k+1})$  и  $2b_k = -\dot{x}_k$ . В самом деле, уравнение  $\dot{L} = [A, L]$  эквивалентно системе уравнений

$$\begin{aligned} \dot{b}_1 &= 2a_1^2, \quad \dot{b}_2 = 2(a_2^2 - a_1^2), \dots, \quad \dot{b}_n = -2a_{n-1}^2, \\ \dot{a}_1 &= a_1(b_2 - b_1), \dots, \quad \dot{a}_{n-1} = a_{n-1}(b_n - b_{n-1}). \end{aligned}$$

Из уравнения

$$\dot{a}_k = a_k(b_{k+1} - b_k) = a_k \frac{\dot{x}_k - \dot{x}_{k+1}}{2}$$

следует, что

$$\ln a_k = \frac{1}{2}(x_k - x_{k+1}) + c_k, \quad \text{т.е.} \quad a_k = d_k \exp \frac{1}{2}(x_k - x_{k+1}).$$

Поэтому уравнение  $\dot{b}_k = 2(a_k^2 - a_{k-1}^2)$  эквивалентно уравнению

$$-\frac{\ddot{x}_k}{2} = 2(d_k^2 \exp(x_k - x_{k+1}) - d_{k-1}^2 \exp(x_{k-1} - x_k)).$$

Если  $d_1 = \dots = d_{h-1} = \frac{1}{2}$ , то получаем требуемые уравнения.

### 8.9.3. Уравнения движения твёрдого тела

Движение многомерного твёрдого тела с матрицей инерции  $J$  описывается уравнением

$$\dot{M} = [M, \omega], \quad (8.1)$$

где  $\omega$  — кососимметрическая матрица,  $M = J\omega + \omega J$ ; матрицу  $J$  можно при этом считать диагональной. Уравнение (8.1) уже записано в форме Лакса, поэтому  $I_k = \text{tr } M^{2k}$  при  $k = 1, \dots, [n/2]$  — первые интегралы этого уравнения (остальные функции  $\text{tr } M^p$  либо нулевые, либо выражаются через указанные функции  $I_k$ ). Но можно получить гораздо больше первых интегралов, записав уравнение (8.1) в виде

$$(M + \lambda J^2)' = [M + \lambda J^2, \omega + \lambda J], \quad (8.2)$$

как это было сделано в статье [Ma]; здесь  $\lambda$  — произвольная константа. Для доказательства эквивалентности уравнений (8.1) и (8.2) достаточно заметить, что

$$[M, J] = -J^2\omega + \omega J^2 = -[J^2, \omega].$$

Первыми интегралами уравнения (8.2) будут все ненулевые коэффициенты многочленов

$$P_k(\lambda) = \text{tr}(M + \lambda J^2)^k = \sum b_s \lambda^s.$$

Так как  $M^T = -M$  и  $J^T = J$ , то

$$P_k(\lambda) = \text{tr}(-M + \lambda J^2)^k = \sum (-1)^{k-s} b_s \lambda^s.$$

Поэтому если  $k - s$  нечётно, то  $b_s = 0$ .

#### 8.9.4. Уравнения Вольтерра

В виде семейства уравнений Лакса, зависящего от параметра  $\lambda$ , можно записать также систему уравнений *Вольтерра*

$$\dot{a}_i = a_i \left( \sum_{k=1}^{p-1} a_{i+k} - \sum_{k=1}^{p-1} a_{i-k} \right), \quad (8.1)$$

где  $p \geq 2$  и  $a_{i+n} = a_i$ . Такое представление приведено в книге [БЗ]. Пусть  $M = \|m_{ij}\|_1^n$  и  $A = \|a_{ij}\|_1^n$ , где в каждой матрице отличны от нуля лишь  $n$  элементов  $m_{i,i+1} = 1$  и  $a_{i,i+1-p} = a_i$ . Рассмотрим уравнение

$$(A + \lambda M)' = [A + \lambda M, -B - \lambda M^p]. \quad (8.2)$$

Если  $B = \sum_{j=0}^{p-1} M^{p-1-j} A M^j$ , то  $[M, B] + [A, M^p] = 0$ , поэтому уравнение (8.2) эквивалентно уравнению  $\dot{A} = -[A, B]$ . Легко проверить, что  $b_{ij} = a_{i+p-1,j} + \dots + a_{i,j+p-1}$ . Поэтому  $b_{ij} = 0$  при  $i \neq j$  и  $b_i = b_{ii} = \sum_{k=0}^{p-1} a_{i+k}$ . Уравнение  $\dot{A} = -[A, B]$  эквивалентно системе уравнений  $\dot{a}_{ij} = a_{ij}(b_i - b_j)$ , причём  $a_{ij} \neq 0$  лишь при  $j = i + 1 - p$ . В итоге получаем систему уравнений

$$\dot{a}_i = a_i \left( \sum_{k=0}^{p-1} a_{i+k} - \sum_{k=1}^{p-1} a_{i+1-p+k} \right) = a_i \left( \sum_{k=0}^{p-1} a_{i+k} - \sum_{k=1}^{p-1} a_{i-k} \right).$$

Таким образом, уравнение (8.1) имеет первые интегралы  $I_k = \text{tr}(A + \lambda M)^{kp}$ .

## 8.10. Матрицы с предписанными собственными значениями

### 8.10.1. Фиксированная диагональ

**Теорема 8.10.1.1.** Для любого многочлена  $f(x) = x^n + c_1x^{n-1} + \dots + c_n$  и любых чисел  $a_1, \dots, a_{n-1}$  существует матрица порядка  $n$  с характеристическим многочленом  $f$  и элементами  $a_1, \dots, a_n$  на диагонали (последний диагональный элемент  $a_n$  задаётся соотношением  $a_1 + \dots + a_n = -c_1$ ).

**Доказательство.** [Fa2] Многочлены  $u_0 = 1, u_1 = x - a_1, \dots, u_n = (x - a_1) \dots (x - a_n)$  образуют базис в пространстве многочленов степени не более  $n$ , поэтому  $f = u_n + \lambda_1 u_{n-1} + \dots + \lambda_n u_0$ . Приравнявая коэффициенты при  $x^{n-1}$  в левой и правой части, получаем  $c_1 = -(a_1 + \dots + a_n) + \lambda_1$ , т.е.  $\lambda_1 = c_1 + (a_1 + \dots + a_n) = 0$ . Пусть

$$A = \begin{pmatrix} a_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & a_2 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & 1 \\ -\lambda_n & -\lambda_{n-1} & -\lambda_{n-2} & \dots & -\lambda_2 & a_n \end{pmatrix}.$$

Разложив определитель матрицы  $xI - A$  по последней строке, получим  $|xI - A| = \lambda_n + \lambda_{n-1}u_1 + \dots + \lambda_2u_{n-2} + u_n = f$ , т.е.  $A$  — искомая матрица.  $\square$

**Теорема 8.10.1.2.** Для любого многочлена  $f(x) = x^n + c_1x^{n-1} + \dots + c_n$  и любой матрицы  $B$  порядка  $n - 1$ , характеристический многочлен которой совпадает с минимальным многочленом, существует такая матрица  $A$ , что  $B$  является подматрицей  $A$  и характеристический многочлен  $A$  равен  $f$ .

**Доказательство.** [Fa2] Пусть  $A = \begin{pmatrix} B & P \\ Q^T & b \end{pmatrix}$ , где  $P$  и  $Q$  — произвольные столбцы длиной  $n - 1$ ,  $b$  — произвольное число. Ясно, что  $\det(xI_n - A) = (x - b) \det(xI_{n-1} - B) - Q^T \operatorname{adj}(xI_{n-1} - B)P$  (см. теорему 1.3.1.3). Докажем, что  $\operatorname{adj}(xI_{n-1} - B) = \sum_{r=0}^{n-2} u_r(x)B^r$ , где многочлены  $u_0, \dots, u_{n-2}$  образуют базис пространства многочленов степени не более  $n - 2$ . Пусть  $g(x) = \det(xI_{n-1} - B) = x^{n-1} + t_1x^{n-2} + \dots$  и  $\varphi(x, \lambda) = (g(x) - g(\lambda))/(x - \lambda)$ . Тогда  $(xI - B)\varphi(x, B) = g(x)I - g(B) = g(x)I$ , так как по теореме Гамильтона–Кэли  $g(B) = 0$ . Поэтому  $\varphi(x, B) = g(x)(xI - B)^{-1} = \operatorname{adj}(xI -$

*B*). Кроме того, так как  $(x^k - \lambda^k)/(x - \lambda) = \sum_{s=0}^{k-1} x^{k-1-s}\lambda^s$ , то  $\varphi(x, \lambda) = \sum_{r=0}^{n-2} t_{n-r-2} \sum_{s=0}^r x^{r-s}\lambda^s = \sum_{s=0}^{n-2} \lambda^s \sum_{r=s}^{n-2} t_{n-r-2} x^{r-s}$ , а значит,  $\varphi(x, \lambda) = \sum_{s=0}^{n-2} \lambda^s u_s(x)$ , где  $u_s = x^{n-s-2} + t_1 x^{n-s-3} + \dots + t_{n-s-2}$ . Итак,  $\det(xI_n - A) = (x-b)(x^{n-1} + t_1 x^{n-2} + \dots) - \sum_{s=0}^{n-2} u_s Q^T B^s P = x^n + (t_1 - b)x^{n-1} + h(x) - \sum_{s=0}^{n-2} u_s Q^T B^s P$ , где  $h$  — многочлен степени не более  $n-2$  и многочлены  $u_0, \dots, u_{n-2}$  образуют базис пространства многочленов степени не более  $n-2$ . А так как характеристический многочлен матрицы  $B$  совпадает с минимальным многочленом, то столбцы  $Q$  и  $P$  можно выбрать так, что  $(Q^T P, \dots, Q^T B^{n-2} P)$  — произвольный заданный набор чисел (см. §3.4).  $\square$

### 8.10.2. Фиксированные внедиагональные элементы

Для доказательства теоремы 8.10.2.2 нам понадобится следующее утверждение о решениях систем уравнений.

#### Теорема 8.10.2.1. Многочлены

$$f_i(x_1, \dots, x_n) = x_i^{m_i} + P_i(x_1, \dots, x_n),$$

где  $i = 1, \dots, n$ , таковы, что  $\deg P_i < m_i$ ;  $I(f_1, \dots, f_n)$  — идеал, порождённый многочленами  $f_1, \dots, f_n$ .

а) Пусть  $P(x_1, \dots, x_n)$  — ненулевой многочлен вида  $\sum a_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n}$ , где  $i_k < m_k$  при всех  $k = 1, \dots, n$ . Тогда  $P \notin I(f_1, \dots, f_n)$ .

б) Система уравнений  $x_i^{m_i} + P_i(x_1, \dots, x_n) = 0$  ( $i = 1, \dots, n$ ) всегда разрешима над  $\mathbb{C}$ , причём число решений конечно.

**Доказательство.** а) Подставляя вместо  $x_i^{m_i t_i + q_i}$ , где  $0 < t_i$  и  $0 \leq q_i < m_i$ , многочлен  $(f_i - P_i)^{t_i} x^{q_i}$ , имеющий по переменным  $x_1, \dots, x_n$  строго меньшую степень, получим, что любой многочлен  $Q(x_1, \dots, x_n)$  можно представить в виде  $Q(x_1, \dots, x_n) = Q^*(x_1, \dots, x_n, f_1, \dots, f_n) = \sum a_{js} x_1^{j_1} \dots x_n^{j_n} f_1^{s_1} \dots f_n^{s_n}$ , где  $j_1 < m_1, \dots, j_n < m_n$ . Докажем, что такое представление  $Q^*$  определено однозначно. Достаточно проверить, что подстановкой  $f_i = x_i^{m_i} + P_i(x_1, \dots, x_n)$  любой ненулевой многочлен  $Q^*(x_1, \dots, x_n, f_1, \dots, f_n)$  приводится к ненулевому многочлену  $\tilde{Q}(x_1, \dots, x_n)$ . Среди всех членов многочлена  $Q^*$  выберем тот, для которого сумма  $(s_1 m_1 + j_1) + \dots + (s_n m_n + j_n) = m$  максимальна. Ясно, что  $\deg \tilde{Q} \leq m$ . Вычислим коэффициент при мономе  $x_1^{s_1 m_1 + j_1} \dots x_n^{s_n m_n + j_n}$  в  $\tilde{Q}$ . Так как сумма  $(s_1 m_1 + j_1) + \dots + (s_n m_n + j_n)$  максимальна, этот моном



может возникнуть лишь из монома  $x_1^{j_1} \dots x_n^{j_n} f_1^{s_1} \dots f_n^{s_n}$ . Следовательно, коэффициенты при этих двух мономах равны и  $\deg \tilde{Q} = m$ .

Ясно, что  $Q(x_1, \dots, x_n) \in I(f_1, \dots, f_n)$  тогда и только тогда, когда  $Q^*(x_1, \dots, x_n, f_1, \dots, f_n)$  — сумма мономов, для которых  $s_1 + \dots + s_n \geq 1$ . Кроме того, если  $P(x_1, \dots, x_n) = \sum a_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n}$ , где  $i_k < m_k$ , то  $P^*(x_1, \dots, x_n, f_1, \dots, f_n) = P(x_1, \dots, x_n)$ . Поэтому  $P \notin I(f_1, \dots, f_n)$ .

б) Если многочлены  $f_1, \dots, f_n$  не имеют общего нуля, то по теореме Гильберта о нулях (см., например, [П2]) идеал  $I(f_1, \dots, f_n)$  совпадает с кольцом всех многочленов, а значит,  $P \in I(f_1, \dots, f_n)$ , что противоречит а). Следовательно, данная система уравнений разрешима. Пусть  $\xi = (\xi_1, \dots, \xi_n)$  — решение этой системы. Тогда  $\xi_i^{m_i} = -P_i(\xi_1, \dots, \xi_n)$ , где  $\deg P_i < m_i$ , поэтому любой многочлен  $Q(\xi_1, \dots, \xi_n)$  представим в виде  $Q(\xi_1, \dots, \xi_n) = \sum a_{i_1 \dots i_n} \xi_1^{i_1} \dots \xi_n^{i_n}$ , где  $i_k < m_k$ . Следовательно,  $\dim K[\xi_1, \dots, \xi_n] \leq m_1 \dots m_n = m$ , а значит,  $m + 1$  элементов  $1, \xi_i, \dots, \xi_i^m$  линейно зависимы, т.е. любой элемент  $\xi_i$  удовлетворяет полиномиальному соотношению  $\sum b_k \xi_i^k = 0$ . Ясно, что число решений уравнения  $\sum b_k x^k = 0$  конечно, а это уравнение едино для всех решений системы и зависит лишь от номера  $i$ .  $\square$

**Теорема 8.10.2.2.** Пусть в комплексной матрице  $A$  заданы все внедиагональные элементы. Тогда диагональные элементы  $x_1, \dots, x_n$  можно подобрать так, что собственные значения матрицы  $A$  совпадут с данными комплексными числами; при этом количество таких наборов  $x_1, \dots, x_n$  конечно.

**Доказательство.** [Fr3] Ясно, что  $\det(A + \lambda I) = (x_1 + \lambda) \dots (x_n + \lambda) + \sum_{k \leq n-2} a_{i_1 \dots i_k} (x_{i_1} + \lambda) \dots (x_{i_k} + \lambda) = \sum \lambda^{n-k} \sigma_k(x_1, \dots, x_n) + \sum_{k \leq n-2} \lambda^{n-k} p_k(x_1, \dots, x_k)$ , где  $p_k$  — многочлен,  $\deg p_k \leq k - 2$ . Уравнение  $\det(A + \lambda I) = 0$  имеет корни  $\lambda_1, \dots, \lambda_n$  тогда и только тогда, когда  $\sigma_k(\lambda_1, \dots, \lambda_n) = \sigma_k(x_1, \dots, x_n) + p_k(x_1, \dots, x_n)$ . Таким образом, наша задача сводится к решению системы уравнений  $\sigma_k(\lambda_1, \dots, \lambda_n) = q_k(x_1, \dots, x_n)$ , где  $k = 1, \dots, n$  и степень многочлена  $q_k$  не превосходит  $k - 1$  (и даже  $k - 2$  при  $k \geq 2$ ).

Пусть  $x = x_i$ ,  $\sigma_k = \sigma_k(x_1, \dots, x_n)$  и  $\sigma'_k = \sigma_k(x_1, \dots, \hat{x}_i, \dots, x_n)$ . Тогда  $\sigma_k = x \sigma'_{k-1} + \sigma'_k$ . Поэтому  $x_i^n - \sigma_1 x_i^{n-1} + \sigma_2 x_i^{n-2} - \dots + (-1)^n \sigma_n = x_i^n - (x + \sigma'_1) x_i^{n-1} + (x \sigma'_1 + \sigma_2) x_i^{n-2} - \dots = 0$ . Рассмотрим многочлены  $f_i(x_1, \dots, x_n) = x_i^n + q_1 x_i^{n-1} - q_2 x_i^{n-2} + \dots + (-1)^n q_n = x_i^n + r_i(x_1, \dots, x_n)$ , где степень  $r_i$  меньше  $n$ . Тогда  $f_i = f_i - (x_i^n - \sigma_1 x_i^{n-1} + \sigma_2 x_i^{n-2} - \dots + (-1)^n \sigma_n) = x_i^{n-1} g_1 + \dots + x_i^{n-2} g_2 + \dots + g_n$ , где  $g_i = (-1)^{i-1} (\sigma_i + q_i)$ , т.е.  $F = VG$ , где  $F$  и  $G$  — столбцы  $(f_1, \dots, f_n)$  и  $(g_n, \dots, g_1)$ ,  $V = \|x_{i+1}^j\|_0^{n-1}$ . Поэтому

$G = V^{-1}F$ , а так как  $V^{-1} = W^{-1}V_1$  где  $W = \det V = \prod_{i>j} (x_i - x_j)$  и  $V_1$  — матрица, элементы которой — многочлены от  $x_1, \dots, x_n$ , то  $Wg_1, \dots, Wg_n \in I[f_1, \dots, f_n]$ , где  $I[f_1, \dots, f_n]$  — идеал кольца многочленов над  $\mathbb{C}$ , порождённый  $f_1, \dots, f_n$ .

Предположим, что многочлены  $g_1, \dots, g_n$  не имеют общих корней. Тогда по теореме Гильберта о нулях  $1 = \sum v_i g_i$ , поэтому  $W = \sum v_i (Wg_i) \in I[f_1, \dots, f_n]$ . С другой стороны,  $W = \sum a_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n}$ , где  $i_k < n$ . Поэтому  $W \notin I[f_1, \dots, f_n]$  (теорема 8.10.2.1 а). Следовательно, многочлены  $g_1, \dots, g_n$  имеют общий корень.

Докажем теперь, что количество общих корней многочленов  $g_1, \dots, g_n$  конечно. Пусть  $\xi = (x_1, \dots, x_n)$  — корень многочленов  $g_1, \dots, g_n$ . Тогда  $\xi$  — корень многочленов  $f_1, \dots, f_n$ , так как  $f_i = x_i^{n-1} g_1 + \dots + g_n$ . Поэтому  $x_i^n + r_i(x_1, \dots, x_n) = f_i = 0$ , причём степень многочлена  $r_i$  меньше  $n$ . Но такая система уравнений имеет конечное число решений (теорема 8.10.2.1 б). Поэтому число различных наборов  $x_1, \dots, x_n$  конечно.  $\square$

### 8.10.3. Предписанная диагональ

**Теорема 8.10.3.1.** Пусть  $\lambda_1 \leq \dots \leq \lambda_n$ ,  $d_1 \leq \dots \leq d_n$ ,  $d_1 + \dots + d_k \geq \lambda_1 + \dots + \lambda_k$  при  $k = 1, \dots, n-1$  и  $d_1 + \dots + d_n = \lambda_1 + \dots + \lambda_n$ . Тогда существует такая ортогональная матрица  $P$ , что на диагонали матрицы  $P\Lambda P$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ , стоят числа  $d_1, \dots, d_n$ .

**Доказательство.** [Ch2] Пусть сначала  $n = 2$ . Тогда  $\lambda_1 \leq d_1 \leq d_2 \leq \lambda_2$  и  $d_2 = \lambda_1 + \lambda_2 - d_1$ . Если  $\lambda_1 = \lambda_2$ , то можно положить  $P = I$ . Если же  $\lambda_1 < \lambda_2$ , то матрица

$$P = \frac{1}{\sqrt{\lambda_2 - \lambda_1}} \begin{pmatrix} \sqrt{\lambda_2 - d_1} & -\sqrt{d_1 - \lambda_1} \\ \sqrt{d_1 - \lambda_1} & \sqrt{\lambda_2 - d_1} \end{pmatrix}$$

искомая.

Предположим теперь, что утверждение верно для некоторого  $n \geq 2$  и рассмотрим наборы из  $n+1$  чисел. Так как  $\lambda_1 \leq d_1 \leq d_{n+1} \leq \lambda_{n+1}$ , то существует такой номер  $j > 1$ , что  $\lambda_{j-1} \leq d_1 \leq \lambda_j$ . Пусть  $P_1$  — такая матрица перестановки, что  $P_1^T \Lambda P_1 = \text{diag}(\lambda_1, \lambda_j, \lambda_2, \dots, \hat{\lambda}_j, \dots, \lambda_{n+1})$ . Легко проверить, что  $\lambda_1 \leq \min(d_1, \lambda_1 + \lambda_j - d_1) \leq \max(d_1, \lambda_1 + \lambda_j - d_1) \leq \lambda_j$ . Поэтому существует такая ортогональная матрица  $Q$  порядка 2, что на диагонали матрицы  $Q^T \text{diag}(\lambda_1, \lambda_j) Q$  стоят числа  $d_1$  и  $\lambda_1 + \lambda_j - d_1$ . Рассмотрим матрицу  $P_2 = \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix}$  порядка  $n+1$ . Ясно, что  $P_2^T (P_1^T \Lambda P_1) P_2 = \begin{pmatrix} d_1 & b^T \\ b & \Lambda_1 \end{pmatrix}$ , где  $\Lambda_1 = \text{diag}(\lambda_1 + \lambda_j - d_1, \lambda_2, \dots, \hat{\lambda}_j, \dots, \lambda_{n+1})$ . Упорядоченные

по возрастанию диагональные элементы матрицы  $\Lambda_1$  и числа  $d_2, \dots, d_{n+1}$  удовлетворяют условию теоремы. В самом деле,  $d_2 + \dots + d_k \geq (k-1)d_1 \geq \lambda_2 + \dots + \lambda_k$  при  $k = 2, \dots, j-1$  и  $d_2 + \dots + d_k = d_1 + \dots + d_k - d_1 \geq \lambda_1 + \dots + \lambda_k - d_1 = (\lambda_1 + \lambda_j - d_1) + \lambda_2 + \dots + \lambda_{j-1} + \lambda_{j+1} + \dots + \lambda_k$  при  $k = j, \dots, n+1$ ; в обоих случаях правые части неравенств не меньше суммы  $k-1$  минимальных диагональных элементов матрицы  $\Lambda_1$ . Поэтому существует такая ортогональная матрица  $Q_1$ , что на диагонали матрицы  $Q_1^T \Lambda_1 Q_1$  стоят числа  $d_2, \dots, d_{n+1}$ . Пусть  $P_3 = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$ . Тогда матрица  $P = P_1 P_2 P_3$  искомая.  $\square$

**Замечание.** Если  $\lambda_1 \leq \dots \leq \lambda_n$ ,  $d_1 \leq \dots \leq d_n$  и на диагонали матрицы  $P \Lambda P$ , где  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  и  $P$  — ортогональная матрица, стоят числа  $d_1, \dots, d_n$ , то  $d_1 + \dots + d_k \geq \lambda_1 + \dots + \lambda_k$  при  $k = 1, \dots, n-1$  и  $d_1 + \dots + d_n = \lambda_1 + \dots + \lambda_n$ .

## 8.11. Числовой образ оператора

### 8.11.1. Теорема Тёплица–Хаусдорфа

Пусть  $A$  — оператор в эрмитовом пространстве  $V$ . Числовым образом оператора  $A$  называют множество  $W(A)$  всех комплексных чисел вида  $(Av, v)$ , где  $v$  пробегает все векторы в  $V$  единичной длины. Это понятие ввёл Тёплиц [То]. Он доказал, что внешняя граница множества  $W(A)$  представляет собой выпуклую кривую, и высказал предположение, что  $W(A)$  — выпуклое множество. Выпуклость множества  $W(A)$  вскоре доказал Хаусдорф [Ha2]. А именно, он доказал, что точки множества  $W(A)$  с фиксированной вещественной частью образуют отрезок, и воспользовался тем, что ранее доказал Тёплиц.

**Теорема 8.11.1.1 (Тёплиц–Хаусдорф).** Числовой образ оператора является выпуклым множеством.

**Доказательство.** Пусть  $v$  и  $w$  — единичные векторы в  $V$ ,  $a = (Av, v)$  и  $b = (Aw, w)$ . Требуется доказать, что отрезок, соединяющий точки  $a$  и  $b$ , целиком лежит в  $W(A)$ . Если  $a = b$ , то это очевидно. Если же  $a \neq b$ , то можно выбрать комплексные числа  $\lambda$  и  $\mu$  так, что  $\lambda a + \mu = 1$  и  $\lambda b + \mu = 0$ . Рассмотрим вместо оператора  $A$  оператор  $A' = \lambda A + \mu I$ . Легко проверить, что  $W(A') = \lambda W(A) + \mu$ , поэтому  $W(A')$  содержит точки 0 и 1. Достаточно доказать, что  $W(A')$  содержит отрезок  $[0, 1]$ . Действительно, если  $u \in V$

— единичный вектор, для которого  $(A'u, u) = t$ , т.е.  $\lambda A(u, u) + \mu = t$ , то

$$\begin{aligned}\lambda A(u, u) + \mu &= t(\lambda a + \mu) + (1 - t)(\lambda b + \mu) = \\ &= \lambda(ta + (1 - t)b) + \mu.\end{aligned}$$

Учитывая, что  $\lambda \neq 0$ , получаем  $(Au, u) = ta + (1 - t)b$ . Таким образом, можно считать, что  $(Av, v) = 1$  и  $(Aw, w) = 0$ .

Запишем оператор  $A$  в виде  $A = B + iC$ , где  $B$  и  $C$  — эрмитовы операторы. Числа  $(Av, v)$  и  $(Aw, w)$  вещественные, поэтому  $(Cv, v) = 0$  и  $(Cw, w) = 0$ . Если заменить  $v$  на  $\lambda v$ , где  $|\lambda| = 1$ , то  $(Av, v)$  не изменится, а  $(Cv, w)$  заменится на  $\lambda(Cv, w)$ . Поэтому можно считать, что число  $(Cv, w)$  чисто мнимое.

Для каждого  $t \in [0, 1]$  рассмотрим вектор  $u(t) = tv + (1 - t)w$ . Векторы  $v$  и  $w$  линейно независимы, поскольку  $(Av, v) = 1$  и  $(A\lambda w, \lambda w) = |\lambda|^2(Aw, w) = 0$ . Поэтому  $u(t) \neq 0$  для всех  $t$ . При этом

$$(Cu(t), u(t)) = t^2(Cv, v) + 2t(1 - t)\operatorname{Re}(Cv, w) + (1 - t)^2(Cw, w) = 0,$$

поскольку  $(Cv, v) = (Cw, w) = 0$  и число  $(Cv, w)$  чисто мнимое. Следовательно, число  $(Au(t), u(t))$  вещественное для всех  $t \in [0, 1]$ . Вещественная функция  $f(t) = \frac{(Au(t), u(t))}{\|u(t)\|^2}$  непрерывна на отрезке  $[0, 1]$ , причём  $f(0) = 0$  и  $f(1) = 1$ . Поэтому  $f(t)$  принимает все значения от 0 до 1.  $\square$

По поводу других доказательств выпуклости числового образа оператора см. теорему 8.11.6.3 и задачу 11.4.

### 8.11.2. Числовой образ нормального оператора

Если  $\lambda$  — собственное значение оператора  $A$ , а  $u$  — собственный вектор единичной длины, соответствующий этому собственному значению, то  $(Au, u) = \lambda(u, u) = \lambda$ . Поэтому все собственные значения оператора принадлежат его числовому образу. Значит, согласно теореме 8.11.1.1 числовой образ оператора всегда содержит выпуклую оболочку всех его собственных значений.

**Теорема 8.11.2.1.** *Если оператор  $A$  нормален, то его числовой образ совпадает с выпуклой оболочкой его собственных значений.*

**Доказательство.** Согласно теореме Шура (теорема 3.7.1.1) матрицу нормального оператора  $A$  можно представить в виде  $A = U^* \Lambda U$ , где  $U$  — унитарная матрица, а  $\Lambda$  — диагональная матрица. При этом  $(Ax, x) =$

$x^*Ax = x^*U^*\Lambda Ux = y^*\Lambda y$ , где  $y = Ux$ . Оператор  $U$  унитарный, поэтому  $y$  тоже пробегает все векторы единичной длины. Таким образом,  $(Ax, x) = \lambda_1|y_1|^2 + \dots + \lambda_n|y_n|^2$ , где  $\lambda_1, \dots, \lambda_n$  — диагональные элементы матрицы  $\Lambda$ , т.е. собственные значения матрицы  $A$  (см. замечание 2 к теореме 3.1.3.2). Числа  $|y_1|^2, \dots, |y_n|^2$  неотрицательны и их сумма равна 1, причём, когда  $x$  пробегает все векторы единичной длины, мы получаем все возможные такие наборы чисел, т.е. получаем выпуклую оболочку точек  $\lambda_1, \dots, \lambda_n$ .  $\square$

### 8.11.3. Числовой образ и нормальные расширения

*Расширением* матрицы  $A$  называют матрицу вида  $\begin{pmatrix} A & X \\ Y & X \end{pmatrix}$ . По поводу некоторых свойств расширений матриц см задачи 4.9, 5.6 и 7.5.

Ясно, что если квадратная матрица  $B$  является расширением квадратной матрицы  $A$ , то  $W(A) \subset W(B)$ .

**Теорема 8.11.3.1.** [На1] Числовой образ оператора совпадает с пересечением числовых образов его нормальных расширений.

**Доказательство.** Числовой образ оператора  $A$  содержится в числовом образе любого его расширения, поэтому он содержится в пересечении числовых образов его нормальных расширений. Нужно лишь доказать обратное включение. Множество  $W(A)$  замкнуто (оно является образом компактного множества при непрерывном отображении) и выпукло (теорема 8.11.1.1). Поэтому достаточно доказать, что для любой замкнутой полуплоскости, содержащей  $W(A)$ , найдётся такое нормальное нормальное расширение  $N$  оператора  $A$ , что  $W(N)$  содержится в той же самой полуплоскости. Используя перенос и поворот, т.е. заменяя оператор  $A$  на  $\alpha A + \beta I$ , где  $|\alpha| = 1$ , рассматриваемую полуплоскость можно заменить на правую полуплоскость  $x \geq 0$ . Если  $(Av, v) = x + iy$ , то  $(A^*v, v) = x - iy$ . Поэтому  $W(A)$  содержится в замкнутой правой полуплоскости тогда и только тогда, когда оператор  $A + A^*$  неотрицательно определён. Таким образом, нужно доказать, что если  $A + A^* \geq 0$ , то у оператора  $A$  есть нормальное расширение  $N$ , для которого  $N + N^* \geq 0$ . Требуемый оператор строится явно, а именно,  $N = \begin{pmatrix} A & A^* \\ A^* & A \end{pmatrix}$ . Равенство  $N^*N = NN^*$  проверяется непосредственно. Оператор  $N + N^* = \begin{pmatrix} A+A^* & A+A^* \\ A+A^* & A+A^* \end{pmatrix}$  неотрицательно определён согласно задаче 1.8.  $\square$

### 8.11.4. Приложение к коммутаторам

**Теорема 8.11.4.1.** [Sh2] Пусть матрица  $A$  нормальна и  $0 \notin W(B)$ . Тогда если матрица  $C = ABA^{-1}B^{-1}$  коммутирует с  $A$ , то и матрица  $B$  коммутирует с  $A$ .

**Доказательство.** Для нормальной матрицы  $A$  можно выбрать унитарную матрицу  $U$  так, что матрица  $U^*AU$  диагональная. Поэтому, заменив матрицы  $B$  и  $C$  на  $U^*BU$  и  $U^*CU$ , можно считать, что матрица  $A$  диагональная; более того,  $A$  имеет блочно диагональный вид с диагональными блоками  $\alpha_1 I_{n_1}, \dots, \alpha_k I_{n_k}$ , где числа  $\alpha_1, \dots, \alpha_k$  попарно различны. Тогда из условия  $[A, C] = 0$  следует, что матрица  $C$  имеет блочно диагональный вид с некоторыми блоками  $C_1, \dots, C_k$  на диагонали (это доказывается аналогично решению задачи 1.1 а).

Запишем матрицу  $B$  в блочном виде с блоками  $B_{ij}$  в соответствии с блочной записью матрицы  $A$ . Равенство  $C = ABA^{-1}B^{-1}$  можно переписать в виде  $CBA = AB$ . Диагональные блоки с номером  $i$  матриц  $CBA$  и  $AB$  равны  $\alpha_i C_i B_{ii}$  и  $\alpha_i B_{ii}$ . По условию матрица  $A$  обратима, поэтому  $\alpha_i \neq 0$ , а значит,  $C_i B_{ii} = B_{ii}$ . Если матрица  $B_{ii}$  вырожденная, то  $0 \in W(B_{ii}) \subset W(B)$ , что противоречит условию. Следовательно, матрица  $B_{ii}$  невырожденная и  $C_i = I_{n_i}$ . Поэтому  $C = I$ , т.е.  $AB = BA$ .  $\square$

**Замечание.** Условие  $0 \notin W(B)$  выполняется, например, когда  $B$  — унитарная матрица, все собственные значения которой лежат на дуге единичной окружности, длина которой меньше  $\pi$ .

### 8.11.5. Числовой образ вещественных квадратичных форм

Пусть  $q_1, \dots, q_k$  — вещественные квадратичные формы от  $n$  переменных. Рассмотрим в  $\mathbb{R}^k$  множество

$$R(q_1, \dots, q_k) = \{(q_1(v), \dots, q_k(v)) \mid v \in \mathbb{R}^n, \|v\| = 1\}.$$

Если  $k = 2$ , то можно отождествить  $\mathbb{R}^2$  с  $\mathbb{C}$  и рассматривать  $R(q_1, q_2)$  как множество в  $\mathbb{C}$ . При таком отождествлении для квадратичных форм, соответствующих вещественным симметрическим матрицам  $A$  и  $B$ , множество  $R(A, B)$  содержится в  $W(A + iB)$ . Действительно,  $W(A + iB)$  строится аналогично, только  $v$  берётся в  $\mathbb{C}^n$ , а не в  $\mathbb{R}^n$ . Оказывается, что если  $n \geq 3$ , то множество  $R(A, B)$  выпуклое и, более того, оно совпадает с  $W(A + iB)$ .

Если  $n = 2$ , то множество  $R(q_1, q_2)$  не обязательно выпуклое. Например, если  $q_1(v) = v_1^2 - v_2^2$  и  $q_2 = 2v_1v_2$ , то  $R(q_1, q_2)$  — это единичная окружность.

Если  $k \geq 3$ , то множество  $R(q_1, \dots, q_k)$  тоже не обязательно выпуклое. Например, если  $n = 3$  и  $q_1(v) = v_1^2 - v_2^2$ ,  $q_2 = 2v_1v_2$  и  $q_3(v) = v_3$ , то  $R(q_1, q_2, q_3)$  — это единичная сфера.

Для доказательства теоремы 8.11.5.2 нам потребуется следующее свойство квадрик (гиперповерхностей, заданных уравнением второй степени) в проективном пространстве  $\mathbb{R}P^n$ .

**Теорема 8.11.5.1.** *При  $n \geq 2$  любая квадрика в  $\mathbb{R}P^n$  линейно связна.*

**Доказательство.** Применим индукцию по  $n$ . При  $n = 2$  невырожденная квадрика с топологической точки зрения представляет собой окружность; вырожденная квадрика — пустое множество, точка, прямая, пара прямых или вся проективная плоскость. Каждое из этих множеств линейно связно. Предположим, что требуемое утверждение доказано для квадрики в  $\mathbb{R}P^{n-1}$ . Рассмотрим квадрику  $Q_n$  в  $\mathbb{R}P^n$ . Достаточно доказать, что любые две точки  $a$  и  $b$  в  $Q_n$  можно соединить путём. Проведём через точки  $a$  и  $b$  в  $\mathbb{R}P^n$  гиперплоскость  $\mathbb{R}P^{n-1}$ . Ясно, что  $Q_{n-1} = Q_n \cap \mathbb{R}P^{n-1}$  — квадрика в  $\mathbb{R}P^{n-1}$ . По предположению индукции квадрика  $Q_{n-1}$  линейно связна, поэтому точки  $a$  и  $b$  можно соединить путём, лежащим в  $Q_{n-1}$ .  $\square$

**Теорема 8.11.5.2.** [Br2] Пусть  $A$  и  $B$  — вещественные симметрические матрицы порядка  $n \geq 3$ . Тогда  $R(A, B)$  — выпуклое множество.

**Доказательство.** Множество  $R(A, B)$  является образом вещественного проективного пространства  $\mathbb{R}P^{n-1}$  при отображении

$$v \mapsto \left( \frac{v^T A v}{\|v\|}, \frac{v^T B v}{\|v\|} \right) \in \mathbb{R}^2.$$

Достаточно доказать, что прообраз любой прямой  $ax + by + c = 0$  на плоскости линейно связан. Действительно, мы можем взять прямую, проходящую через две точки  $P$  и  $Q$  множества  $R(A, B)$ , рассмотреть её прообраз и соединить в нём некоторым путём точку из прообраза  $P$  с точкой из прообраза  $Q$ . Образ этого пути содержит отрезок  $PQ$ .

Прообраз прямой  $ax + by + c = 0$  задаётся в  $\mathbb{R}P^{n-1}$  уравнением

$$av^T A v + bv^T B v + cv^T v = 0.$$

Это уравнение задаёт квадрику, а при  $n \geq 3$  квадрика в  $\mathbb{R}P^{n-1}$  линейно связна (теорема 8.11.5.1).  $\square$

**Теорема 8.11.5.3.** [Br2] Пусть  $A$  и  $B$  — вещественные симметрические матрицы порядка  $n \geq 3$ . Тогда  $R(A, B) = W(A + iB)$ .

**Доказательство.** Нужно лишь проверить, что  $W(A+iB) \subset R(A, B)$ . Пусть  $x = v^T A \bar{v}$  и  $y = v^T B \bar{v}$  для некоторого комплексного вектора  $v$  единичной длины. Запишем  $v$  в виде  $v = v_1 + iv_2$ , где  $v_1$  и  $v_2$  — вещественные векторы. Тогда

$$\begin{aligned} x &= (v_1^T + v_2^T)A(v_1 - iv_2) = v_1^T A v_1 + v_2^T A v_2, \\ y &= v_1^T B v_1 + v_2^T B v_2. \end{aligned}$$

Ясно, что  $v_1^T v_1 + v_2^T v_2 = 1$ , поэтому если  $v_1 = 0$  или  $v_2 = 0$ , то  $(x, y) \in R(A, B)$ . Если же  $v_1 \neq 0$  и  $v_2 \neq 0$ , то мы можем рассмотреть единичные векторы  $u_1 = v_1/\|v_1\|$  и  $u_2 = v_2/\|v_2\|$  и записать

$$v_1^T A v_1 = v_1^T v_1 \frac{v_1^T}{\|v_1\|} A \frac{v_1}{\|v_1\|} = \|v_1\|^2 u_1^T A u_1$$

и т.д. Пусть  $t = \|v_1\|^2$ . Тогда  $\|v_2\|^2 = 1 - t$  и

$$\begin{aligned} x &= t u_1^T A u_1 + (1 - t) u_2^T A u_2, \\ y &= t u_1^T B u_1 + (1 - t) u_2^T B u_2. \end{aligned}$$

Согласно теореме 8.11.5.2 множество  $R(A, B)$  выпуклое, поэтому  $(x, y) \in R(A, B)$ .  $\square$

### 8.11.6. Отображение, заданное двумя эрмитовыми матрицами

**Теорема 8.11.6.1.** *[Au1] Пусть  $A$  и  $B$  — вещественные симметрические (соответственно комплексные эрмитовы) матрицы порядка  $n$ , причём  $(u^* A u)^2 + (u^* B u)^2 > 0$  для любого ненулевого вектора  $u$  с вещественными (комплексными) координатами. Тогда если  $n \geq 3$  (соответственно  $n \geq 2$ ), то образ единичной сферы при отображении  $f$ , заданном формулой*

$$f(u) = \left( \frac{u^* A u}{\sqrt{(u^* A u)^2 + (u^* B u)^2}}, \frac{u^* B u}{\sqrt{(u^* A u)^2 + (u^* B u)^2}} \right),$$

*представляет собой дугу единичной окружности длины меньше  $\pi$ .*

**Доказательство.** Рассматриваемое отображение непрерывно, поэтому его образ — компактное связное подмножество окружности, т.е. некоторая дуга. Таким образом, достаточно доказать, что не существует ненулевых векторов  $u_1$  и  $u_2$ , для которых  $u_1^* A u_1 = -u_2^* A u_2$  и  $u_1^* B u_1 = -u_2^* B u_2$ .



Предположим, что такие векторы  $u_1$  и  $u_2$  существуют. Ясно, что они должны быть линейно независимыми. По условию  $n \geq 3$  в вещественном случае и  $n \geq 2$  в комплексном случае. Следовательно, можно выбрать вектор  $u_3$  так, что векторы  $u_1$ ,  $u_2$  и  $u_3$  линейно независимы над  $\mathbb{R}$ . Для любых трёх вещественных чисел  $(x, y, z)$  получаем

$$(xu_1 + yu_2 + zu_3)^* A(xu_1 + yu_2 + zu_3) = a_0(x^2 - y^2) + a_1xy + a_2xz + a_3yz + a_4z^2,$$

$$(xu_1 + yu_2 + zu_3)^* B(xu_1 + yu_2 + zu_3) = b_0(x^2 - y^2) + b_1xy + b_2xz + b_3yz + b_4z^2,$$

где  $a_0 = u_1^* A u_1$ ,  $b_0 = u_1^* B u_1$ ,  $a_1 = u_1^* A u_2 + u_2^* A u_1$  и т.д. (все числа вещественные).

**Лемма.** Система уравнений

$$a_0(x^2 - y^2) + a_1xy + a_2xz + a_3yz + a_4z^2 = 0,$$

$$b_0(x^2 - y^2) + b_1xy + b_2xz + b_3yz + b_4z^2 = 0,$$

с произвольными вещественными коэффициентами имеет ненулевое вещественное решение.

**Доказательство.** Рассмотрим сначала случай, когда  $a_0b_1 = a_1b_0$ . Положив  $z = 0$ , получим систему уравнений

$$a_0(x^2 - y^2) + a_1xy = 0, \quad b_0(x^2 - y^2) + b_1xy = 0.$$

Поскольку  $a_0b_1 = a_1b_0$ , эта система сводится к одному уравнению, причём это уравнение имеет ненулевое вещественное решение.

Рассмотрим теперь случай, когда  $a_0b_1 \neq a_1b_0$ . Положив  $z = 1$ , получим систему уравнений

$$a_0(x^2 - y^2) + a_1xy + a_2x + a_3y + a_4 = 0,$$

$$b_0(x^2 - y^2) + b_1xy + b_2x + b_3y + b_4 = 0.$$

Поскольку  $a_0b_1 \neq a_1b_0$ , эта система уравнений эквивалентна следующей:

$$x^2 - y^2 + a'_2x + a'_3y + a'_4 = 0, \quad xy + b'_2x + b'_3y + b'_4 = 0,$$

где  $a'_2, a'_3, a'_4, b'_2, b'_3, b'_4$  — вещественные числа. Эта система уравнений, очевидно, имеет вещественное решение.  $\square$

Воспользовавшись этой леммой, найдём тройку вещественных чисел  $(x_0, y_0, z_0) \neq (0, 0, 0)$ , для которой  $u^* A u = 0$  и  $u^* B u = 0$ , где  $u = x_0 u_1 + y_0 u_2 + z_0 u_3$  — ненулевой вектор. Но это противоречит тому, что  $(u^* A u)^2 + (u^* B u)^2 > 0$  для любого ненулевого вектора  $u$ .  $\square$

**Теорема 8.11.6.2.** Пусть  $A$  и  $B$  — вещественные симметрические (соответственно комплексные эрмитовы) матрицы порядка  $n$ , причём  $(u^*Au)^2 + (u^*Bu)^2 > 0$  для любого ненулевого вектора  $u$  с вещественными (комплексными) координатами. Тогда если  $n \geq 3$  (соответственно  $n \geq 2$ ), то можно выбрать вещественные числа  $\lambda$  и  $\mu$  так, что  $\lambda A + \mu B > 0$ .

**Доказательство.** [Au1] Условия теоремы 8.11.6.1 выполнены, поэтому образ отображения  $f$  из формулировки этой теоремы, представляет собой дугу длины меньше  $\pi$ . Пусть  $(\lambda, \mu)$  — середина этой дуги. Для любого ненулевого  $u$  угол  $\varphi$  между единичными векторами  $(\lambda, \mu)$  и  $f(u)$  меньше  $\pi/2$ , поэтому

$$\lambda \frac{u^*Au}{\sqrt{(u^*Au)^2 + (u^*Bu)^2}} + \mu \frac{u^*Bu}{\sqrt{(u^*Au)^2 + (u^*Bu)^2}} = \cos \varphi > 0.$$

Это означает, что матрица  $\lambda A + \mu B$  положительно определённая.  $\square$

**Замечание.** Напомним, что если для эрмитовых матриц  $A$  и  $B$  можно выбрать вещественные числа  $\lambda$  и  $\mu$  так, что  $\lambda A + \mu B > 0$ , то матрицы  $A$  и  $B$  одновременно диагонализуются (теорема 4.2.2.1).

Доказательство теоремы 8.11.6.2 позволяет получить ещё одно доказательство выпуклости числового образа оператора. А именно, можно доказать следующее утверждение.

**Теорема 8.11.6.3.** Пусть  $A$  и  $B$  — вещественные симметрические (соответственно комплексные эрмитовы) матрицы порядка  $n$ . Тогда если  $n \geq 3$  (соответственно  $n \geq 2$ ), то множество

$$R(A, B) = \{(u^*Au, u^*Bu) \mid u^*u = 1\}$$

выпукло.

**Доказательство.** Пусть  $u_1$  и  $u_2$  — единичные векторы,  $a_1 = u_1^*Au_1$ ,  $b_1 = u_1^*Bu_1$ ,  $a_2 = u_2^*Au_2$  и  $b_2 = u_2^*Bu_2$ . Предположим, что для некоторого  $t \in (0, 1)$  точка  $(ta_1 + (1-t)a_2, tb_1 + (1-t)b_2)$  не принадлежит  $R(A, B)$ . Это означает, что для любого единичного вектора  $u$

$$(u^*Au, u^*Bu) \neq (ta_1 + (1-t)a_2, tb_1 + (1-t)b_2).$$

Поэтому если мы рассмотрим эрмитовы матрицы  $\tilde{A} = A - (ta_1 + (1-t)a_2)I$  и  $\tilde{B} = B - (tb_1 + (1-t)b_2)I$ , то  $(u^*\tilde{A}u, u^*\tilde{B}u) \neq (0, 0)$  для любого ненулевого вектора  $u$ . При доказательстве теоремы 8.11.6.2 фактически

было получено, что при таком условии ни для каких ненулевых векторов  $u_1$  и  $u_2$  векторы  $(u_1^* \tilde{A}u_1, u_1^* \tilde{B}u_1)$  и  $(u_2^* \tilde{A}u_2, u_2^* \tilde{B}u_2)$  не могут быть противоположно направлены. Но, как показывают простые вычисления,

$$\begin{aligned}(u_1^* \tilde{A}u_1, u_1^* \tilde{B}u_1) &= (1-t)(a_1 - a_2, b_1 - b_2), \\ (u_2^* \tilde{A}u_2, u_2^* \tilde{B}u_2) &= -t(a_1 - a_2, b_1 - b_2).\end{aligned}$$

Получено противоречие.  $\square$

Следующее доказательство вещественной части теоремы 8.11.6.2 будет интересно топологам. (Необходимые для него сведения из топологии можно найти в [ПЗ]).

**Теорема 8.11.6.4.** Пусть  $p(x)$  и  $q(x)$  — квадратичные формы от  $n$  вещественных переменных, причём  $n \geq 3$ . Предположим, что  $p$  и  $q$  не обращаются в нуль одновременно ни для какого ненулевого вектора  $x \in \mathbb{R}^n$ . Тогда можно выбрать вещественные числа  $\lambda$  и  $\mu$  так, что форма  $\lambda p + \mu q$  положительно определена.

**Доказательство.** [Ca1] Поскольку  $(p(\lambda x), q(\lambda x)) = \lambda^2(p(x), q(x))$ , отображение  $x \mapsto (p(x), q(x))$  после соответствующей факторизации можно рассматривать как отображение  $f : \mathbb{R}P^{n-1} \rightarrow \mathbb{R}P^1$ , где  $\mathbb{R}P^k$  — вещественное проективное пространство размерности  $k$ . Поскольку  $\pi_1(\mathbb{R}P^{n-1}) = \mathbb{Z}_2$  при  $n \geq 3$ , а  $\pi_1(\mathbb{R}P^1) = \mathbb{Z}$ , то индуцированный гомоморфизм  $f_* : \pi_1(\mathbb{R}P^{n-1}) \rightarrow \pi_1(\mathbb{R}P^1)$  тривиален. Следовательно, для отображения  $f$  существует поднятие  $\tilde{f} : \mathbb{R}P^{n-1} \rightarrow \mathbb{R}$  в универсальное накрытие:

$$\begin{array}{ccc} & & \mathbb{R} \\ & \nearrow \tilde{f} & \downarrow p \\ \mathbb{R}P^{n-1} & \xrightarrow{f} & \mathbb{R}P^1. \end{array}$$

Пространство  $\mathbb{R}P^{n-1}$  компактно и связно, поэтому его образ  $\tilde{f}(\mathbb{R}P^{n-1})$  является замкнутым отрезком  $I \subset \mathbb{R}$ . Возможны два случая.

**Случай 1.**  $p(I) = \mathbb{R}P^1$ .

В этом случае существует точка  $(a : b) \in \mathbb{R}P^1$ , прообраз которой  $p^{-1}(a : b)$  содержит по крайней мере две различные точки  $t_1, t_2 \in I$ . Прообраз  $f^{-1}(a : b)$  — это квадрика  $Q$ , заданная уравнением  $bp(x) - aq(x) = 0$  в  $\mathbb{R}P^{n-1}$ . Прообразы  $\tilde{f}^{-1}(t_1)$  и  $\tilde{f}^{-1}(t_2)$  — это непустые подмножества

в квадрике  $Q$ , которые одновременно открыты и замкнуты. Действительно, при ограничении на  $Q$  мы получаем непрерывное отображение  $\tilde{f}|_Q : Q \rightarrow p^{-1}(a : b)$ . Точки  $t_1$  и  $t_2$  одновременно открыты и замкнуты в дискретном пространстве  $p^{-1}(a : b)$ , поэтому их прообразы тоже одновременно открыты и замкнуты. В итоге мы получаем противоречие с тем, что квадратика в  $\mathbb{R}P^{n-1}$  при  $n \geq 3$  связна (теорема 8.11.5.1).

**Случай 1.**  $p(I) \neq \mathbb{R}P^1$ .

В этом случае дополнение к  $p(I)$  в окружности  $\mathbb{R}P^1$  — непустое множество  $J$ . Если  $(a : b)$  — произвольная точка множества  $J$ , то  $bp(x) - aq(x) \neq 0$  для любого ненулевого вектора  $x \in \mathbb{R}^n$ . Следовательно, квадратичная форма  $bp - aq$  либо положительно определена, либо отрицательно определена (но тогда квадратичная форма  $-bp + aq$  положительно определена).  $\square$

## Задачи

**11.1.** Докажите, что диагональные элементы матрицы  $A$  принадлежат её числовому образу  $W(A)$ .

**11.2.** Докажите, что числовой образ оператора  $A$  в 2-мерном пространстве представляет собой эллипс (с фокусами в собственных значениях  $\lambda_1$  и  $\lambda_2$  и малой полуосью  $\sqrt{\operatorname{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$ ), который вырождается в отрезок или точку тогда и только тогда, когда оператор нормальный.

**11.3.** Пусть  $U \subset V$  — комплексное подпространство,  $P : V \rightarrow V$  — эрмитов проектор на  $U$ . Сопоставим оператору  $A$  в  $V$  оператор  $\tilde{A} = (PA)|_U$  в  $U$ . Докажите, что  $W(\tilde{A}) \subset W(A)$ .

**11.4.** Получите ещё одно доказательство выпуклости числового образа оператора, воспользовавшись задачами 11.2 и 11.3.

**11.5.** [Fi2] Докажите, что диагональ матрицы комплексного оператора  $A$  в некотором ортонормированном базисе имеет вид  $(\operatorname{tr} A, 0, \dots, 0)$  тогда и только тогда, когда  $\operatorname{tr} A \in W(A)$ .

**11.6.** Докажите теорему 3.5.2.1, воспользовавшись задачей 11.5.

## Решения

**1.1.** Легко проверить, что произведение указанных кватернионов равно  $-(x_1x_2 + y_1y_2 + z_1z_2) + (y_1z_2 - z_1y_2)i + (z_1x_2 - z_2x_1)j + (x_1y_2 - x_2y_1)k$ .

**1.2.** Пусть  $q = a + v$ , где  $a$  — вещественная часть кватерниона,  $v$  — его мнимая часть. Тогда  $(a + v)^2 = a^2 + 2av + v^2$ . Согласно теореме 8.1.2.1  $v^2 = -\bar{v}v = -|v|^2 \leq 0$ . Поэтому кватернион  $a^2 + 2av + v^2$  вещественный тогда и только тогда, когда  $av$  — вещественный кватернион, т.е.  $a = 0$  или  $u = 0$ .

**1.3.** Из решения задачи 1.2 видно, что  $q^2 = -1$  тогда и только тогда, когда  $q = xi + yj + zk$ , где  $x^2 + y^2 + z^2 = 1$ .

**1.4.** Пусть кватернион  $q = a + v$ , где  $a$  — вещественная часть кватерниона  $q$ , коммутирует с любым чисто мнимым кватернионом  $w$ . Тогда  $(a + v)w = w(a + v)$  и  $aw = wa$ , поэтому  $vw = wv$ . А так как  $\overline{vw} = \bar{w}\bar{v} = wv$ , то  $vw$  — вещественный кватернион. Остаётся заметить, что если  $v \neq 0$  и кватернион  $w$  не пропорционален  $\bar{v}$ , то  $vw \notin \mathbb{R}$ .

**1.5.** Пусть  $B = W_1 + W_2j$ , где  $W_1$  и  $W_2$  — комплексные матрицы. Тогда  $AB = Z_1W_1 + Z_2jW_1 + Z_1W_2j + Z_2jW_2j$  и

$$A_c B_c = \begin{pmatrix} Z_1W_1 - Z_2\overline{W_2} & Z_1W_2 + Z_2\overline{W_1} \\ -\overline{Z_2}W_1 - \overline{Z_1}W_2 & -\overline{Z_2}W_2 + \overline{Z_1}W_1 \end{pmatrix}.$$

Поэтому достаточно доказать, что  $Z_2jW_1 = Z_2\overline{W_1}j$  и  $Z_2jW_2j = -Z_2\overline{W_2}$ . Так как  $ji = -ij$ , то  $jW_1 = \overline{W_1}j$ , а так как  $jj = -1$  и  $jij = i$ , то  $jW_2j = -\overline{W_2}$ .

**1.6.** Ясно, что  $(qx, qy) = (\bar{x}\bar{q}qy + \bar{y}\bar{q}qx)/2 = |q|^2(x, y)$ . Отображение  $x \mapsto qx/|q|$  ортогонально и сохраняет ориентацию. Поэтому определитель отображения  $x \mapsto qx$  равен  $|q|^4$ .

**1.7.** Тетраэдр можно расположить в пространстве кватернионов. Пусть  $a, b, c$  и  $d$  — кватернионы, соответствующие его вершинам. Можно считать, что  $c$  и  $d$  — вещественные кватернионы. Тогда

$$(a - b)(c - d) + (b - c)(a - d) = (b - d)(a - c).$$

В самом деле, все сомножители, содержащие  $b$ , сокращаются, а кватернионы  $a, c$  и  $d$  попарно коммутируют. Следовательно,  $|a - b||c - d| + |b - c||a - d| \geq |b - d||a - c|$ .

**1.8.** а) Пусть  $x = a + be$  и  $y = u + ve$ . Тогда

$$\begin{aligned} (xy)y &= [(au - \bar{v}b)u - \bar{v}(b\bar{u} + va)] + [(b\bar{u} + va)\bar{u} + v(au - \bar{v}b)]e, \\ x(yu) &= [a(u^2 - \bar{v}v) - (\bar{v}\bar{u} + v\bar{u})b] + [b(\bar{u}^2 - \bar{v}v) + (v\bar{u} + v\bar{u})a]e. \end{aligned}$$

А так как  $\bar{v}v = v\bar{v}$  и  $u + \bar{u}$  — действительные числа, то

$$\begin{aligned} a(u^2 - \bar{v}v) - (\bar{v}\bar{u} + v\bar{u})b &= au^2 - a\bar{v}v - (u + \bar{u})\bar{v}b = \\ &= au^2 - \bar{v}va - \bar{v}b(u + \bar{u}) = (au - \bar{v}b)u - \bar{v}(b\bar{u} + va), \\ \overline{b(u^2 - \bar{v}v)} + (v\bar{u} + v\bar{u})a &= b\bar{u}^2 - b\bar{v}v + v(\bar{u} + u)a = \\ &= b\bar{u}^2 - v\bar{v}b + va(\bar{u} + u) = (b\bar{u} + va)\bar{u} + v(au - \bar{v}b). \end{aligned}$$

Равенство  $x(xy) = (xx)y$  доказывается аналогично.

б) Рассмотрим трилинейное отображение  $f(a, x, y) = (ax)y - a(xy)$ . Подставив  $b = x + y$  в равенство  $(ab)b = a(bb)$  и учтя, что  $(ax)x = a(xx)$  и  $(ay)y = a(yy)$ , получим  $(ax)y - a(yx) = a(xy) - (ay)x$ , т.е.  $f(a, x, y) = -f(a, y, x)$ . Аналогично, подставив  $b = x + y$  в равенство  $b(ba) = (bb)a$ , получим  $f(x, y, a) = -f(y, x, a)$ . Значит,  $f(a, x, y) = -f(a, y, x) = f(y, a, x) = -f(y, x, a)$ , т.е.  $(ax)y + (yx)a = a(xy) + y(xa)$ . При  $a = y$  получаем  $(yx)y = y(xy)$ .

2.1. Так как  $A_i S = \sum_j A_i A_j = \sum_k A_k = S$ , то  $S^2 = \sum_i A_i S = nS$ . Поэтому все собственные значения матрицы  $S$  равны 0 или  $n$ , а так как  $\text{tr } S = 0$ , то все собственные значения равны 0. Следовательно, матрица  $S - nI$  невырождена и из равенства  $S(S - nI) = 0$  получаем  $S = 0$ .

3.1. Легко проверить, что  $2|A| = 0$  (см. задачу 1.1). Для поля нечётной характеристики из этого равенства следует, что  $|A| = 0$ .

Для поля характеристики 2 аналогичное утверждение неверно; в этом случае, например, единичная матрица является кососимметрической.

### 3.2. Матрицы

$$X = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \quad \text{и} \quad Y = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p-1 & 0 \end{pmatrix}$$

удовлетворяют требуемому соотношению.

3.3. Первое решение. Пусть  $f(x) = \det(xI - A)$ , где  $I$  — единичная матрица. Любому многочлену  $f$  с целыми коэффициентами соответствует многочлен  $\tilde{f}$  над полем  $\mathbb{F}_p$  (конечное поле из  $p$  элементов). Присоединим к  $\mathbb{F}_p$  корни  $x_1, \dots, x_n$  многочлена  $\tilde{f}$ . В результате получим поле  $\mathbb{F}_{p^k}$ .

Для любых элементов  $x, y \in \mathbb{F}_{p^k}$  выполняется равенство  $(x + y)^p = x^p + y^p$ , поскольку  $\binom{p}{m} \equiv 0 \pmod{p}$  при  $m = 1, 2, \dots, p-1$ . Поэтому

$$x_1^p + \dots + x_n^p = (x_1 + \dots + x_n)^p = x_1 + \dots + x_n$$

(второе равенство следует из того, что  $x_1 + \dots + x_n \in \mathbb{F}_p$ ). Но  $x_1 + \dots + x_n = \text{tr } A \pmod{p}$  и  $(x_1 + \dots + x_n)^p = \text{tr}(A^p) \pmod{p}$ . Первое равенство очевидно, а второе следует из того, что если  $f(x) = (x - \lambda_1) \dots (x - \lambda_n)$ , то  $\lambda_1^p + \dots + \lambda_n^p$  выражается как многочлен с целыми коэффициентами через коэффициенты многочлена  $f$ , причём  $x_1^p + \dots + x_n^p$  выражается через коэффициенты многочлена  $\tilde{f}$  точно так же.

Второе решение. Докажем сначала, что если  $A$  и  $B$  — матрицы с целочисленными элементами, то  $\text{tr}(A + B)^p \equiv (\text{tr } A^p + \text{tr } B^p) \pmod{p}$ . Если из выражения для  $(A + B)^p$  исключить матрицы  $A^p$  и  $B^p$ , то все остальные мономы разбиваются на наборы по  $p$  мономов в каждом, причём все мономы в одном наборе имеют один и тот же след. А именно, каждый набор состоит из матриц  $A_1 A_2 \dots A_p, A_2 A_3 \dots A_p A_1, \dots, A_p A_1 A_2 \dots A_{p-1}$ , где  $A_i = A$  или  $B$ . Ясно,

что любые два таких набора либо совпадают, либо не пересекаются. Кроме того,  $\text{tr } A_1(A_2 \dots A_p) = \text{tr}(A_2 \dots A_p)A_1 = \dots$

Индукцией по  $n$  теперь легко получить сравнение  $\text{tr}(X_1 + \dots + X_n)^p \equiv (\text{tr } X_1^p + \dots + \text{tr } X_n^p) \pmod{p}$ .

Представим матрицу  $A$  в виде  $A = X_1 + \dots + X_n$ , где каждая матрица  $X_i$  имеет ровно один ненулевой элемент. Для такой матрицы имеет место сравнение  $\text{tr}(X_i^p) \equiv \text{tr } X_i \pmod{p}$ . Действительно, если ненулевой элемент расположен вне диагонали, то  $\text{tr}(X_i^p) = 0$  и  $\text{tr } X_i = 0$ . А если ненулевой элемент расположен на диагонали и равен  $a$ , то  $\text{tr}(X_i^p) = a^p$  и  $\text{tr } X_i = a$ . Но  $a^p \equiv a \pmod{p}$  согласно малой теореме Ферма.

**4.1.** Пусть  $c_0, \dots, c_{n+m-1}$  — столбцы матрицы Сильвестра  $S(f, g)$  и  $y_k = x^{m+n-k-1}$ . Тогда  $y_0 c_0 + \dots + y_{n+m-1} c_{n+m-1} = c$ , где  $c$  — столбец  $(x^{m-1}f(x), \dots, f(x), x^{n-1}g(x), \dots, g(x))$ . Рассмотрим это равенство как систему линейных уравнений относительно  $y_0, \dots, y_{n+m-1}$  и применим для нахождения  $y_{n+m-1}$  правило Крамера. В результате получим

$$y_{n+m-1} \det(c_0, \dots, c_{n+m-1}) = \det(c_0, \dots, c_{n+m-2}, c). \quad (8.1)$$

Остаётся заметить, что  $y_{n+m-1} = 1$ ,  $\det(c_0, \dots, c_{n+m-1}) = R(f, g)$ , а определитель, стоящий в правой части равенства (8.1), можно представить в требуемом виде.

**4.2.** Согласно теореме 8.4.2.1  $R(f, g) = a_0^m \prod_i g(x_i)$  и  $R(f, r) = a_0^k \prod_i r(x_i)$ . Кроме того,  $f(x_i) = 0$ , поэтому  $g(x_i) = f(x_i)q(x_i) + r(x_i) = r(x_i)$ .

**4.3.** Пусть  $c_0, \dots, c_{n+m-1}$  — столбцы матрицы Сильвестра  $S(f, g)$  и  $y_k = x^{n+m-k-1}$ . Тогда  $y_0 c_0 + \dots + y_{n+m-1} c_{n+m-1} = c$ , где  $c$  — столбец  $(x^{m-1}f(x), \dots, f(x), x^{n-1}g(x), \dots, g(x))$ . Ясно, что если  $k \leq n-1$ , то  $x^k g(x) = \sum \lambda_i x^i f(x) + r_k(x)$ , где  $\lambda_i$  — некоторые числа и  $i \leq m-1$ . Поэтому, прибавляя к последним  $n$  элементам столбца  $c$  линейные комбинации первых  $m$  элементов, этот столбец можно привести к виду  $(x^{m-1}f(x), \dots, f(x), r_{n-1}(x), \dots, r_0(x))$ . Аналогичные преобразования строк матрицы  $S(f, g)$  приводят эту матрицу к виду

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}, \quad \text{где } A = \begin{pmatrix} a_0 & & * \\ & \ddots & \\ 0 & & a_0 \end{pmatrix}, \quad C = \begin{pmatrix} a_{n-1,0} & \dots & a_{n-1,n-1} \\ \dots & \dots & \dots \\ a_{0,0} & \dots & a_{0,n-1} \end{pmatrix}.$$

**4.4.** Рассматриваемому оператору соответствует оператор  $I_m \otimes A - B^T \otimes I_m$  (см. §5.1.6). Собственные значения этого оператора равны  $\alpha_i - \beta_i$ , где  $\alpha_i$  — корни многочлена  $f$ ,  $\beta_i$  — корни многочлена  $g$  (см. §5.1.5). Поэтому его определитель равен  $\prod_{i,j} (\alpha_i - \beta_j) = \pm R(f, g)$ .

**4.5.** Легко проверить, что  $S = V^T V$ , где

$$V = \begin{pmatrix} 1 & \alpha_1 & \dots & \alpha_1^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & \alpha_n & \dots & \alpha_n^{n-1} \end{pmatrix}.$$

Поэтому  $\det S = (\det V)^2 = \prod_{i < j} (\alpha_i - \alpha_j)^2$ .

**4.6.** Воспользуемся теоремой 8.4.3.1. Ясно, что

$$\lim_{x \rightarrow y} \frac{f(x)g(y) - f(y)g(x)}{x - y} = \lim_{\varepsilon \rightarrow 0} \frac{f(y + \varepsilon)g(y) - f(y)g(y + \varepsilon)}{\varepsilon} =$$

$$= f'(y)g(y) - f(y)g'(y).$$

Значит, если безугиана многочленов  $f$  и  $g$  равна нулю, то  $f'g = g'f$  и  $(f/g)' = (f'g - g'f)/g^2 \equiv 0$ . Обратное утверждение очевидно.

**4.7.** Пусть  $z_1, \dots, z_n$  — строки матрицы  $Z$ ,  $r_1, \dots, r_n$  — строки матрицы  $R$ ,  $s_1, \dots, s_n$  — строки матрицы  $TR$ . Тогда

$$s_n = r_1 = (b_n - b_0 a_n, \dots, b_1 - b_0 a_1) = z_n$$

и  $s_i = a_{n-i}r_1 + a_{n-i-1}r_2 + \dots + r_{n-i+1} = a_{n-i}r_1 + (a_{n-i-1}r_1 A + \dots + r_{n-i} A) = a_{n-i}s_n + s_{i+1}A$  при  $i = n-1, \dots, 2, 1$ . Поэтому остаётся проверить, что  $z_i = a_{n-i}z_n + z_{i+1}A$  при  $i = 1, 2, \dots, n-1$ . Ясно, что  $z_{ij} - a_{n-i}z_{nj} - (z_{i+1}A)_j = z_{ij} - a_{n-i}c_{0,n+1-j} - z_{i+1,j-1} + a_{n+1-j}c_{0,n-i} + a_{n+1-j}c_{0,n-i} - a_{n-i}c_{0,n+1-j} = c_{n+1-j,n-i}$ . При  $i = j-1$ ,  $i > j-1$  и  $i < j-1$  величина  $z_{i+1,j-1} - z_{ij}$  равна соответственно 0,  $c_{n-j+1,n-i}$  и  $-c_{n-i,n-j+1}$ . Поэтому  $z_{i+1,j-1} - z_{ij} = c_{n-j+1,n-i}$ .

**5.1.** Рассмотрим отображение  $f: V \rightarrow V^*$ , сопоставляющее элементу  $x \in V$  линейную функцию  $h(y) = (x, y)$ . Так как скалярное произведение невырождено, то  $\text{Ker } f = 0$ , а значит,  $f$  — изоморфизм. Поэтому  $\dim f(W) = \dim W$ . Ясно также, что пространство  $W^\perp$  совпадает с аннулятором пространства  $f(W)$  (см. §2.1.8). Следовательно,  $\dim W^\perp = \dim V - \dim W$  и  $\dim (W^\perp)^\perp = \dim V - \dim W^\perp = \dim W$ . А так как  $W \subset (W^\perp)^\perp$ , то  $W = (W^\perp)^\perp$ .

**5.2.** Согласно теореме Витта существует изометрия пространства  $V$ , переводящая  $W_1$  в  $W_2$ . Ясно, что эта изометрия переводит  $W_1^\perp$  в  $W_2^\perp$ .

**5.3.** Линейное отображение пространства  $\langle x \rangle$  в пространство  $\langle y \rangle$ , переводящее  $x$  в  $y$ , является изометрией. По теореме Витта эту изометрию можно продолжить до изометрии всего пространства  $V$ .

**5.4.** Докажем сначала, что  $\dim W_1 = \dim W_2$ . Предположим, что  $\dim W_1 < \dim W_2$ . Тогда  $W_2 = W_2' \oplus W_2''$ , где  $\dim W_2' = \dim W_1$ . Любой изоморфизм  $f: W_2' \rightarrow W_1$  является изометрией. По теореме Витта эту изометрию можно продолжить до изометрии  $\bar{f}$  всего пространства  $V$ . Тогда  $W_1 \subset \bar{f}(W_2)$ , причём  $\dim \bar{f}(W_2) = \dim W_2 > \dim W_1$ , т.е. изотропное подпространство не максимально. Получено противоречие, поэтому  $\dim W_1 = \dim W_2$ . Следовательно, существует изоморфизм  $g: W_1 \rightarrow W_2$ . Этот изоморфизм является изометрией, которую по теореме Витта можно продолжить до изометрии всего пространства  $V$ .

**5.5.** Предположим сначала, что сигнатура формы  $q$  равна нулю. Тогда в пространстве  $V$  существует такой базис  $e_1, \dots, e_n, \varepsilon_1, \dots, \varepsilon_n$ , что  $(e_i, e_j) = \delta_{ij}$ ,  $(\varepsilon_i, \varepsilon_j) = -\delta_{ij}$  и  $(e_i, \varepsilon_j) = 0$ . Легко проверить, что подпространство  $U = \langle e_1 + \varepsilon_1, \dots, e_n + \varepsilon_n \rangle$  изотропно и имеет размерность  $(\dim V)/2$ .



Предположим теперь, что  $U$  — изотропное подпространство и  $\dim U = (\dim V)/2$ . Пусть  $V'_+ = \{x \in V \mid q(x) > 0\}$ ,  $V'_- = \{x \in V \mid q(x) < 0\}$  и  $V_\pm = V'_\pm \cup \{0\}$ . Тогда  $U \cap V_\pm = 0$ , поэтому  $\dim V_\pm \leq \dim V - \dim U = (\dim V)/2$ . А так как  $\dim V_+ + \dim V_- = \dim V$ , то  $\dim V_\pm = (\dim V)/2$ .

**6.1.** Уравнения  $AX = C$  и  $YB = C$  разрешимы, поэтому  $AA^+C =$  и  $CB^+B = C$  (см. §8.6.2). Следовательно,  $AA^+C = AA^+CB^+B = AZB$ , где  $Z = A^+CB^+$ .

**6.2.** Если  $X$  — матрица размера  $m \times n$  и  $\text{rk } X = r$ , то  $X = PQ$ , где  $P$  и  $Q$  — матрицы размеров  $m \times r$  и  $r \times n$  (см. §2.4.2). Пространства, порождённые столбцами матриц  $X$  и  $P$ , совпадают, поэтому из равенства  $AX = 0$  следует равенство  $AP = 0$ , т.е.  $P = (I - A^+A)Y_1$  (см. §8.6.2). Аналогично из равенства  $XB = 0$  следует, что  $Q = Y_2(I - BB^+)$ . Поэтому  $X = PQ = (I - A^+A)Y(I - BB^+)$ , где  $Y = Y_1Y_2$ . Ясно также, что если  $X = (I - A^+A)Y(I - BB^+)$ , то  $AX = 0$  и  $XB = 0$ .

**6.3.** Если  $AX = C$  и  $XB = D$ , то  $AD = AXB = CB$ . Если же  $AX_1 = C$ ,  $X_2B = D$  и  $AD = CB$ , то  $AA^+C = C$  и  $DB^+B = D$ . Поэтому  $A(A^+C + DB^+ - A^+ADB^+) = C$  и  $(A^+C + DB^+ - A^+ADB^+)B = D$ , т.е.  $X_0 = A^+C + DB^+ - A^+ADB^+$  — частное решение рассматриваемой системы уравнений.

**6.4.** Выберем в пространстве  $V$  ортонормированный базис и дополним его до ортонормированного базиса всего пространства. В этом базисе матрицы операторов  $A$  и  $B$  имеют блочный вид  $\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}$  и  $\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ . Поэтому  $AB = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix} = BA$ , где  $m = \dim V$ . Теперь ясно, что  $ABA = A$  и  $BAB = B$ , а  $AB$  и  $BA$  — эрмитовы операторы.

**6.5.** Предположим, что  $AA^+ = A^+A$ . Для обобщённой обратной матрицы выполняется равенство  $AA^+A = A$ , поэтому  $A^2A^+ = A$ . Затем получаем  $A^k(A^+)^{k-1} = A$  для всех натуральных  $k$ . По условию матрица  $A$  нильпотентная, поэтому  $A^k = 0$  для достаточно большого  $k$ . Значит,  $A = 0$ , что противоречит условию.

**6.6.** Уравнение  $X - AXB = C$  эквивалентно системе двух уравнений  $X - AY = C$ ,  $Y = XB$ . В свою очередь, эта система уравнений эквивалентна одному уравнению над кольцом многочленов от переменной  $x$ :

$$(A + xI_m)Y - X(I_n + xB) = -C.$$

Согласно теореме Рота это уравнение имеет решение тогда и только тогда, когда матрицы  $A + xI_m$  и  $I_n + xB$  одного ранга.

**8.1.** Пусть  $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . Тогда  $A^2 = -t^2I$ ,  $A^3 = -t^3J$ ,  $A^4 = t^4I$ ,  $A^5 = t^5J$  и т.д. Поэтому

$$\begin{aligned} e^A &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots\right)I + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots\right)J = \\ &= (\cos t)I + (\sin t)J = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

**8.2. а)** Для коммутирующих матриц справедлива формула бинома Ньютона, поэтому

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{\infty} \frac{(A+B)^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{\binom{n}{k} A^k B^{n-k}}{k!} = \\ &= \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{A^k}{k!} \cdot \frac{B^{n-k}}{(n-k)!} = e^A e^B. \end{aligned}$$

б) Так как  $e^{(A+B)t} = I + (A+B)t + (A^2 + AB + BA + B^2)t^2/2 + \dots$  и  $e^{At}e^{Bt} = I + (A+B)t + (A^2 + 2AB + B^2)t^2/2 + \dots$ , то  $A^2 + AB + BA + B^2 = A^2 + 2AB + B^2$ , а значит,  $AB = BA$ .

**8.3.** Существует такая унитарная матрица  $V$ , что  $U = VDV^{-1}$ , где  $D = \text{diag}(\exp(i\alpha_1), \dots, \exp(i\alpha_n))$ . Пусть  $\Lambda = \text{diag}(\alpha_1, \dots, \alpha_n)$ . Тогда  $U = e^{iH}$ , где  $H = V\Lambda V^{-1} = V\Lambda V^* -$  эрмитова матрица.

**8.4. а)** Пусть  $U = e^X$  и  $X^T = -X$ . Тогда  $UU^T = e^X e^{X^T} = e^X e^{-X} = I$ , так как матрицы  $X$  и  $-X$  коммутируют.

б) Для матрицы  $U$  существует такая ортогональная матрица  $V$ , что  $U = V \text{diag}(A_1, \dots, A_k, I) V^{-1}$ , где  $A_i = \begin{pmatrix} \cos \varphi_i & -\sin \varphi_i \\ \sin \varphi_i & \cos \varphi_i \end{pmatrix}$  (см. теорему 3.1.3.3). Ясно также, что матрицу  $A_i$  можно представить в виде  $e^X$ , где  $X = \begin{pmatrix} 0 & -\varphi_i \\ \varphi_i & 0 \end{pmatrix}$  (см. задачу 8.1).

**8.5. а)** Достаточно заметить, что  $\det(e^A) = e^{\text{tr } A}$  (см. теорему 8.8.1.2) и  $\text{tr } A -$  вещественное число.

б) Пусть  $\lambda_1$  и  $\lambda_2$  — собственные значения вещественной матрицы  $A$  и  $\lambda_1 + \lambda_2 = \text{tr } A = 0$ . Числа  $\lambda_1$  и  $\lambda_2$  либо оба вещественные, либо  $\lambda_1 = \overline{\lambda_2}$ , т.е.  $\lambda_1 = -\overline{\lambda_1}$ . Поэтому собственные значения матрицы  $e^A$  равны либо  $e^\alpha$  и  $e^{-\alpha}$ , либо  $e^{i\alpha}$  и  $e^{-i\alpha}$ , где в обоих случаях  $\alpha -$  вещественное число. Следовательно, матрица  $B = \begin{pmatrix} -2 & 0 \\ 0 & -1/2 \end{pmatrix}$  не является экспонентой вещественной матрицы.

**8.6. а)** Пусть  $A_{ij} -$  алгебраическое дополнение элемента  $a_{ij}$ . Тогда  $\text{tr}(A \text{adj } A^T) = \sum_{i,j} \dot{a}_{ij} A_{ij}$ .

Так как  $\det A = a_{ij} A_{ij} + \dots$ , где многоточием обозначены члены, не содержащие  $a_{ij}$ , то  $(\det A)' = \dot{a}_{ij} A_{ij} + a_{ij} \dot{A}_{ij} + \dots = \dot{a}_{ij} A_{ij} + \dots$ , где многоточием обозначены члены, не содержащие  $\dot{a}_{ij}$ . Поэтому  $(\det A)' = \sum_{i,j} \dot{a}_{ij} A_{ij}$ . Кроме того,  $\text{adj } A^T = (\det A) A^{-1}$ .

б) Так как  $A \text{adj } A^T = (\det A) I$ , то  $\text{tr}(A \text{adj } A^T) = n \det A$ , а значит,  $n(\det A)' = \text{tr}(A \text{adj } A^T) + \text{tr}(A(\text{adj } A^T)')$ . Остаётся воспользоваться результатом задачи а).

**8.7.** Требуется доказать, что если матрица  $A$  нильпотентна, то матрица  $e^A - I$  тоже нильпотентна. Пусть  $\lambda_1, \dots, \lambda_n -$  собственные значения матрицы  $A$ . Тогда

$e^{\lambda_1} - 1, \dots, e^{\lambda_n} - 1$  — собственные значения матрицы  $e^A - I$ . Поэтому если все собственные значения матрицы  $A$  нулевые, то все собственные значения матрицы  $e^A - I$  тоже нулевые.

Попутно мы доказали, что матрица  $e^A$  унитарна тогда и только тогда, когда  $e^{\lambda_1} = 1, \dots, e^{\lambda_n} = 1$ , т.е.  $\lambda_1 = 2k_1\pi i, \dots, \lambda_n = 2k_n\pi i$ , где  $k_1, \dots, k_n$  — целые числа. Если хотя бы одно из чисел  $k_1, \dots, k_n$  отлично от нуля, то матрица  $A$  не нильпотентна.

**8.8.** Легко проверить тождество

$$(\sum X_i)^2 + \sum_{i < j} [X_i, X_j] = \sum X_i^2 + 2 \sum_{i < j} X_i X_j.$$

Таким образом,

$$e^{tX_1} e^{tX_2} \dots e^{tX_n} = I + t \sum X_i + \frac{t^2}{2} (\sum X_i)^2 + \frac{t^2}{2} \sum_{i < j} [X_i, X_j] + O(t^3).$$

С другой стороны,

$$e^{t \sum X_i + \frac{t^2}{2} \sum_{i < j} [X_i, X_j]} = I + t \sum X_i + \frac{t^2}{2} \sum_{i < j} [X_i, X_j] + \frac{t^2}{2} (\sum X_i)^2 + O(t^3).$$

**8.9.** Рассматриваемые определители имеют вид  $\sum \pm \partial_1^{r_1} \partial_2^{r_2} \dots \partial_n^{r_n}$  и  $\sum \pm x_1^{r_1} x_2^{r_2} \dots x_n^{r_n}$ , где  $(r_1, r_2, \dots, r_n)$  — перестановка чисел  $(0, 1, \dots, n-1)$ . Произведение этих сумм содержит ровно  $n!$  ненулевых членов, а именно, оно равно

$$\sum \partial_1^{r_1} \partial_2^{r_2} \dots \partial_n^{r_n} x_1^{r_1} x_2^{r_2} \dots x_n^{r_n}.$$

Каждый из этих членов равен  $0!1!2! \dots (n-1)!$ .

**8.10.** Предположим сначала, что  $m > 0$ . Тогда

$$(X^m)_{ij} = \sum_{a,b,\dots,p,q} x_{ia} x_{ab} \dots x_{pq} x_{qj},$$

$$\text{tr } X^m = \sum_{a,b,\dots,p,q,r} x_{ra} x_{ab} \dots x_{pq} x_{qr}.$$

Поэтому

$$\begin{aligned} \frac{\partial x_{ra}}{\partial x_{ji}} (\text{tr } X^m) &= \sum_{a,b,\dots,p,q,r} \frac{\partial}{\partial x_{ji}} x_{ab} \dots x_{pq} x_{qr} + \dots \\ &\dots + \sum_{a,b,\dots,p,q,r} x_{ra} x_{ab} \dots x_{pq} \frac{\partial x_{qr}}{\partial x_{ji}} = \\ &= \sum_{b,\dots,p,q} x_{ib} \dots x_{pq} x_{qj} + \dots + \sum_{a,b,\dots,p} x_{ia} x_{ab} \dots x_{pj} = m(X^{m-1})_{ij}. \end{aligned}$$

Предположим теперь, что  $m < 0$ . Пусть  $X^{-1} = \|y_{ij}\|_1^n$ . Тогда  $y_{ij} = X_{ji}/\Delta$ , где  $X_{ij}$  — алгебраическое дополнение элемента  $x_{ij}$  в матрице  $X$  и  $\Delta = \det X$ . Согласно теореме Якоби (теорема 1.2.6.3)

$$\begin{vmatrix} X_{i_1 j_1} & X_{i_1 j_2} \\ X_{i_2 j_1} & X_{i_2 j_2} \end{vmatrix} = (-1)^\sigma \begin{vmatrix} x_{i_3 j_3} & \dots & x_{i_3 j_n} \\ \dots & \dots & \dots \\ x_{i_n j_3} & \dots & x_{i_n j_n} \end{vmatrix} \Delta$$

и

$$X_{i_1 j_1} = (-1)^\sigma \begin{vmatrix} x_{i_2 j_2} & \dots & x_{i_2 j_n} \\ \dots & \dots & \dots \\ x_{i_n j_2} & \dots & x_{i_n j_n} \end{vmatrix}, \quad \text{где } \sigma = \begin{pmatrix} i_1 & \dots & i_n \\ j_1 & \dots & j_n \end{pmatrix}.$$

Поэтому

$$\begin{vmatrix} X_{i_1 j_1} & X_{i_1 j_2} \\ X_{i_2 j_1} & X_{i_2 j_2} \end{vmatrix} = \Delta \frac{\partial}{\partial x_{i_2 j_2}} (X_{i_1 j_1}).$$

Следовательно,  $-X_{j\alpha}X_{\beta i} = \Delta \frac{\partial}{\partial x_{ji}}(X_{\beta\alpha}) - X_{\beta\alpha}X_{ji} = \Delta \frac{\partial}{\partial x_{ji}}(X_{\beta\alpha}) - X_{\beta\alpha} \frac{\partial}{\partial x_{ji}}(\Delta) = \Delta^2 \frac{\partial}{\partial x_{ji}} \left( \frac{X_{\beta\alpha}}{\Delta} \right)$ , т.е.  $\frac{\partial}{\partial x_{ij}} y_{\alpha\beta} = -y_{\alpha j} y_{i\beta}$ . А так как  $(X^m)_{ij} = \sum_{a,b,\dots,q} y_{ia} y_{ab} \dots y_{qj}$  и  $\text{tr } X^m = \sum_{a,b,\dots,q,r} y_{ra} y_{ab} \dots y_{qr}$ , то

$$\begin{aligned} \frac{\partial}{\partial x_{ji}} (\text{tr } X^m) &= - \sum_{a,b,\dots,q,r} y_{rj} y_{ia} y_{ab} \dots y_{qr} - \dots \\ &\dots - \sum_{a,b,\dots,q,r} y_{ra} y_{ab} \dots y_{qj} y_{ir} = m(X^{m-1})_{ij}. \end{aligned}$$

**11.1.** Пусть  $e_i = (0, \dots, 1, \dots, 0)^T$ , где 1 стоит на  $i$ -м месте. Тогда  $e_i^* A e_i = a_{ii}$ .

**11.2.** Мы приведём два решения этой задачи; оба решения используют следующие факты. Согласно теореме Шура (теорема 3.7.1.1) квадратную матрицу  $A$  можно представить в виде  $A = UTU^*$ , где  $U$  — унитарная,  $T$  — треугольная матрица. Унитарное преобразование переводит единичную сферу в единичную сферу, поэтому  $W(A) = W(T)$ . Кроме того, если оператор  $A$  заменить на  $A' = \lambda A + \mu I$ , то  $W(A') = \lambda W(A) + \mu$ . Легко проверить, что при замене оператора  $A$  на  $UTU^*$  величина  $\sqrt{\text{tr}(A^* A) - |\lambda_1|^2 - |\lambda_2|^2}$  не изменяется, а при замене на  $\lambda A + \mu I$  умножается на  $|\lambda|$ . Поэтому достаточно доказать требуемое утверждение для преобразованных матриц.

**Первое решение.** [Li] Если матрица  $A$  нормальна, то существует унитарная матрица  $U$ , для которой  $U^* A U = D = \text{diag } \lambda_1, \lambda_2$ . В этом случае

$$W(A) = W(D) = \{\lambda_1 |z_1|^2 + \lambda_2 |z_2|^2 \mid z_1, z_2 \in \mathbb{C}, \quad |z_1|^2 + |z_2|^2 = 1\}.$$

Это множество представляет собой отрезок с концами в точках  $\lambda_1$  и  $\lambda_2$ .

Предположим теперь, что матрица  $A$  не нормальна. Заменяя  $A$  на  $A - \frac{\operatorname{tr} A}{2}I$ , можно считать, что  $\operatorname{tr} A = 0$ . Если оба собственных значения матрицы  $A$  равны 0, то после замены  $A$  на  $U^*AU$  можно считать, что  $A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ . В таком случае

$$W(A) = \{az_1z_2 \mid z_1, z_2 \in \mathbb{C}, \quad |z_1|^2 + |z_2|^2 = 1\}.$$

Это множество представляет собой круг радиуса  $|a|$ .

Остаётся рассмотреть случай, когда матрица  $A$  не нормальна и её собственные значения равны  $\lambda$  и  $-\lambda$ , причём  $\lambda \neq 0$ . Заменяя  $A$  на  $\lambda^{-1}A$ , можно считать, что собственные значения равны 1 и  $-1$ . После замены  $A$  на  $U^*AU$  можно считать, что  $A = \begin{pmatrix} 1 & 2a \\ 0 & -1 \end{pmatrix}$ , где число  $a$  вещественное. Действительно, если  $A = \begin{pmatrix} 1 & z \\ 0 & -1 \end{pmatrix}$  и  $U = \operatorname{diag}(e^{i\varphi}, e^{-i\varphi})$ , то  $U^*AU = \begin{pmatrix} 1 & ze^{-2i\varphi} \\ 0 & -1 \end{pmatrix}$ .

Рассмотрим матрицу

$$B = \frac{A + A^*}{2} + \alpha \frac{A - A^*}{2}, \quad \text{где } \alpha = \frac{\sqrt{1 + a^2}}{a}.$$

Легко проверить, что оба собственных значения матрицы  $B$  равны 0 и  $\sqrt{\operatorname{tr} B^*B} = 2\sqrt{1 + a^2}$ . Поэтому можно выбрать унитарную матрицу  $U$  так, что  $U^*BU = \begin{pmatrix} 0 & 2\sqrt{1 + a^2} \\ 0 & 0 \end{pmatrix}$ . Как уже отмечалось, для такой матрицы множество  $W(B)$  представляет собой круг радиуса  $2\sqrt{1 + a^2}$ .

Если  $(Aw, w) = x + iy$ , где  $x, y \in \mathbb{R}$ , то  $(A^*w, w) = x - iy$ . Поэтому  $x + iy \in W(A)$  тогда и только тогда, когда  $x + i\alpha y \in W(B)$ . Следовательно, если  $W(B)$  представляет собой круг радиуса  $2\sqrt{1 + a^2}$ , то  $W(A)$  представляет собой эллипс с большой полуосью  $2\sqrt{1 + a^2}$  и малой полуосью  $2a$ .

Второе решение. Для матриц порядка 2 достаточно рассмотреть случай, когда  $A = \begin{pmatrix} \lambda_1 & z \\ 0 & \lambda_2 \end{pmatrix}$ , где  $\lambda_1, \lambda_2$  и  $z$  — некоторые комплексные числа. Пусть  $u = (u_1, u_2)$  — единичный вектор в  $\mathbb{C}^2$ . Тогда  $(Au, u) = \lambda_1|u_1|^2 + \lambda_2|u_2|^2 + z\bar{u}_1u_2$ . Фиксируем число  $t = |u_1|$ . Тогда  $u_1 = \sqrt{t}e^{i\varphi}$ ,  $u_2 = \sqrt{1-t}e^{i\psi}$  и  $(Au, u) = t\lambda_1 + (1-t)\lambda_2 + z\sqrt{t(1-t)}e^{i(\psi-\varphi)}$ . Поэтому множество  $W(A)$  получается следующим образом. На отрезке, соединяющем точки  $\lambda_1$  и  $\lambda_2$ , выбирается точка, делящая этот отрезок в отношении  $t : (1-t)$ , и строится окружность радиуса  $|z|\sqrt{t(1-t)}$  с центром в этой точке. Требуется доказать, что множество всех точек таких окружностей — это точки, расположенные внутри некоторого эллипса с фокусами  $\lambda_1$  и  $\lambda_2$ .

Достаточно рассмотреть случай, когда  $\lambda_1 = -1$  и  $\lambda_2 = 1$ . Пусть  $|z| = 2d$ . Докажем, что искомым эллипс задаётся уравнением  $d^2x^2 + (d^2 + 1)y^2 = d^2(d^2 + 1)$ . Будем рассуждать в обратном направлении: возьмём точку  $(x_0, y_0)$  на данном эллипсе, проведём через неё нормаль и вычислим расстояние от точки  $(x_0, y_0)$  до точки пересечения нормали с большой осью эллипса. Касательная к эллипсу в точке  $(x_0, y_0)$  задаётся уравнением  $d^2x_0x + (d^2 + 1)y_0y = d^2(d^2 + 1)$ , поэтому нормаль задаётся уравнением  $-(d^2 + 1)y_0x + d^2x_0y = -x_0y_0$ , где правая часть находится как значение левой части при  $x = x_0$  и  $y = y_0$ . Нормаль пересекает большую ось эллипса в точке  $(x_1, 0)$ , где  $x_1$  удовлетворяет уравнению

$-(d^2 + 1)y_0x_1 = -x_0y_0$ , т.е.  $x_1 = \frac{x_0}{d^2+1}$ . Если  $x_1 = 1 - 2t$ , то  $x_0 = (d^2 + 1)(1 - 2t)$  и  $(d^2 + 1)y_0 = d^2(d^2 + 1) - d^2x_0^2$ , т.е.  $y_0^2 = d^2 - d^2(d^2 + 1)(1 - 2t)^2$ . Несложные вычисления показывают, что квадрат расстояния от точки  $(x_1, 0)$  до точки  $(x_0, y_0)$  равен  $4d^2t(1 - t)$ , как и должно быть.

Наибольшее значение  $x_0$  равно  $\sqrt{d^2 + 1}$ ; ему соответствует  $x_1 = \frac{1}{\sqrt{d^2+1}}$ . Таким образом, когда центр окружности движется по отрезку  $\left[-\frac{1}{\sqrt{d^2+1}}, \frac{1}{\sqrt{d^2+1}}\right]$ , мы получаем требуемое множество. Остаётся доказать, что окружности с центрами в остальных точках отрезка  $[-1, 1]$  не добавляют ничего нового. Покажем, что все эти окружности содержатся внутри окружности, соответствующей  $x_0 = \sqrt{d^2 + 1}$ . Для этого достаточно проверить, что

$$1 - 2t + 2d\sqrt{t(1-t)} \leq \sqrt{d^2 + 1}$$

при  $\frac{\sqrt{d^2+1}-1}{2\sqrt{d^2+1}} \leq t \leq 1$ . Несложные вычисления показывают, что производная функции  $f(t) = 1 - 2t + 2d\sqrt{t(1-t)}$  обращается в нуль при  $t = \frac{\sqrt{d^2+1}+1}{2\sqrt{d^2+1}}$ . Кроме того,  $f(1) = 0$ , поэтому на указанном интервале  $f(t)$  монотонно убывает.

Ясно, что рассматриваемый эллипс вырождается в отрезок (или точку) тогда и только тогда, когда  $z = 0$ , т.е. матрица  $T$  диагональная. Согласно теореме Шура это эквивалентно тому, что матрица  $A$  нормальная.

**11.3.** Пусть  $u \in U$ . Поскольку оператор  $P$  эрмитов и  $Pu = u$ ,  $(PAu, u) = (Au, Pu) = (Au, u)$ .

**11.4.** Пусть  $a_1 = (Av_1, v_1)$  и  $a_2 = (Av_2, v_2)$ , где  $v_1$  и  $v_2$  — единичные векторы в пространстве  $V$ . Возьмём комплексное подпространство  $U = \langle v_1, v_2 \rangle$ , рассмотрим эрмитов проектор  $P$  на это подпространство и рассмотрим в  $U$  оператор  $\tilde{A} = (PA)|_U$ . Согласно задаче 11.2  $W(\tilde{A})$  — эллипс. Этот эллипс содержит точки  $a_1$  и  $a_2$ , поэтому он содержит весь отрезок с концами в этих точках. С другой стороны, согласно задаче 11.3  $W(\tilde{A}) \subset W(A)$ , поэтому  $W(A)$  тоже содержит этот отрезок. (Если  $\dim U = 1$ , то  $a_1 = a_2$  и доказывать нечего.)

**11.5.** Если  $e_1, \dots, e_n$  — ортонормированный базис, в котором диагональ имеет указанный вид, то  $\operatorname{tr} A = (Ae_1, e_1) \in W(A)$ .

Обратное утверждение докажем индукцией по размерности  $n$ . Если  $\operatorname{tr} A \in W(A)$ , то существует единичный вектор  $v$ , для которого  $(Av, v) = \operatorname{tr} A$ . Матрица оператора  $A$  в ортонормированном базисе  $e_1 = v, e_2, \dots, e_n$  имеет вид  $\begin{pmatrix} \operatorname{tr} A & B \\ C & D \end{pmatrix}$ , где  $D$  — матрица порядка  $n-1$  с нулевым следом. Если  $n = 2$ , то  $D = 0$ , и доказательство завершено. Пусть  $n > 2$ . Чтобы применить предположение индукции, нужно проверить, что  $0 \in W(D)$ . Пусть  $\lambda_1, \dots, \lambda_{n-1}$  — собственные значения матрицы  $D$ . Тогда

$$\frac{1}{n}(\lambda_1 + \dots + \lambda_{n-1}) = \frac{\operatorname{tr} D}{n-1} = 0,$$

т.е.  $0$  принадлежит выпуклой оболочке собственных значений матрицы  $D$ . Но все собственные значения матрицы  $D$  принадлежат  $W(D)$ , а кроме того,  $W(D)$  — выпуклое множество (теорема 8.11.1.1). Поэтому  $0 \in W(D)$ . По предположению индукции существует унитарная матрица  $U$ , для которой на диагонали матрицы  $U^*DU$  стоят нули. Матрица  $V = \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix}$  унитарная и главная диагональ матрицы  $V^* \begin{pmatrix} \operatorname{tr} A & B \\ C & D \end{pmatrix} V$  имеет требуемый вид.

**11.6.** Пусть  $A$  — комплексная (вещественная) матрица порядка  $n$ . Тогда след матрицы  $A - \frac{\operatorname{tr} A}{n}I$  равен  $0$ . Как видно из решения задачи 11.5, в таком случае  $0$  принадлежит числовому образу. Поэтому существует унитарная (ортогональная) матрица  $U$ , для которой на диагонали матрицы  $U^*AU - \frac{\operatorname{tr} A}{n}I$  стоят нули, т.е. на диагонали матрицы  $U^*AU$  стоят равные числа.

## Глава 9.

# Некоммутативная линейная алгебра

### 9.1. Матрицы с некоммутирующими элементами

По аналогии с матрицами, элементы которых принадлежат полю или коммутативному кольцу, можно рассматривать матрицы, элементы которых принадлежат некоммутативному телу. Произведение двух таких матриц определяется стандартным образом:  $AB = C$ , где  $c_{ij} = \sum_k a_{ik}b_{kj}$ .

#### 9.1.1. Векторные пространства над телами

Векторное пространство  $V$  над телом  $K$  определяется точно так же, как векторное пространство над полем, но теперь мы будем считать, что векторное пространство *правое*, т.е. векторы стоят слева, а элементы тела стоят справа:  $v\lambda$ ,  $\lambda \in K$  (говоря по-научному,  $V$  является правым модулем над  $K$ ). Ниже мы убедимся, что такое соглашение приводит к обычному правилу для преобразования координат вектора.

Понятия линейной независимости векторов и базиса вводятся обычным образом. Доказательство того, что любые два базиса векторного пространства  $V$  содержат одно и то же число элементов, тоже вполне стандартное, но удобнее использовать не те рассуждения, которые мы применили в коммутативном случае для доказательства теоремы 2.1.1.2.

**Теорема 9.1.1.1.** *Любые два базиса векторного пространства  $V$  содержат одно и то же число элементов.*



**Доказательство.** Пусть  $e_1, \dots, e_n$  — базис пространства  $V$ , а  $\varepsilon_1, \dots, \varepsilon_m$  — некоторый набор линейно независимых векторов пространства  $V$ . Достаточно доказать, что  $m \leq n$ . Представим вектор  $\varepsilon_1$  в виде линейной комбинации векторов базиса:

$$\varepsilon_1 = e_1\lambda_1 + \dots + e_n\lambda_n. \quad (9.1)$$

В этом выражении не все числа  $\lambda_1, \dots, \lambda_n$  равны нулю. Пусть для определённости  $\lambda_1 \neq 0$ . Тогда  $e_1 \in \langle \varepsilon_1, e_2, \dots, e_n \rangle$ . Кроме того, векторы  $\varepsilon_1, e_2, \dots, e_n$  линейно независимы. Действительно, предположим, что  $\varepsilon_1\mu_1 + e_2\mu_2 + \dots + e_n\mu_n = 0$ . Если  $\mu \neq 0$ , то мы можем представить  $\varepsilon_1$  в виде линейной комбинации векторов  $e_2, \dots, e_n$ . Вычитая это выражение из (9.1), получаем линейную зависимость между векторами  $e_1, \dots, e_n$ , чего не может быть. Следовательно,  $\mu_1 = 0$ , а тогда  $\mu_2 = \dots = \mu_n$ , поскольку векторы  $e_2, \dots, e_n$  линейно независимы.

Предположим теперь, что (после изменения нумерации векторов  $e_1, \dots, e_n$ ) мы получили базис  $\varepsilon_1, \dots, \varepsilon_k, e_{k+1}, \dots, e_n$ . Пусть

$$\varepsilon_{k+1} = \varepsilon_1\lambda_1 + \dots + \varepsilon_k\lambda_k + e_{k+1}\lambda_{k+1} + \dots + e_n\lambda_n.$$

Векторы  $\varepsilon_1, \dots, \varepsilon_k$  линейно независимы, поэтому хотя бы один из коэффициентов при  $e_{k+1}, \dots, e_n$  отличен от нуля. Для определённости будем считать, что  $\lambda_{k+1} \neq 0$ . Те же самые рассуждения, которые были проведены выше, показывают, что если заменить вектор  $e_{k+1}$  на  $\varepsilon_{k+1}$ , то мы снова получим базис. Предположим, что  $m \geq n$ . Тогда, повторяя такие замены, мы получим базис  $\varepsilon_1, \dots, \varepsilon_n$ , поэтому  $m \leq n$ .  $\square$

Отображение  $f: V^n \rightarrow W^m$  правых векторных пространств над телом  $K$  называют *линейным*, если для всех  $\lambda_1, \lambda_2 \in K$ ,  $v_1, v_2 \in V^n$  выполняется равенство  $f(v_1\lambda_1 + v_2\lambda_2) = f(v_1)\lambda_1 + f(v_2)\lambda_2$ . Линейная функция определяется аналогично.

Матрицу размера  $m \times n$  с элементами из тела  $K$  можно рассматривать как матрицу линейного отображения  $n$ -мерного пространства над  $K$  в  $m$ -мерное пространство над  $K$ . А именно, если в пространствах  $V^n$  и  $W^m$  заданы базисы  $e_1, \dots, e_n$  и  $\varepsilon_1, \dots, \varepsilon_m$ , то матрице  $A = (a_{ij})$  соответствует линейное отображение  $f: V^n \rightarrow W^m$ , которое на базисных векторах задаётся формулой  $f(e_j) = \sum_{i=1}^m \varepsilon_i a_{ij}$ . Тогда  $f(\sum_{j=1}^n e_j x_j) =$

$\sum_{j=1}^n \sum_{i=1}^m \varepsilon_i a_{ij} x_j$ . Поэтому отображение  $f$  переводит вектор  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  в век-

тор

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Правое векторное пространство над телом  $K$  можно рассматривать как *левое* векторное пространство, положив  $\lambda v = v\lambda^{-1}$  при  $\lambda \neq 0$ . В частности, если мы фиксируем в левом и правом пространствах один и тот же базис, то вектору с ненулевой координатой  $x_i$  в левом пространстве соответствует вектор с координатой  $x_i^{-1}$  в правом пространстве.

Обратите внимание, что мы не можем положить  $\lambda v = v\lambda$ , поскольку тогда мы получим  $\mu(\lambda v) = (v\lambda)\mu \neq v\mu\lambda$ . А при нашем соглашении мы получаем  $\mu(\lambda v) = (v\lambda^{-1})\mu^{-1} = v(\mu\lambda)^{-1}$ .

Для правого векторного пространства  $V$  двойственное пространство  $V^*$  имеет естественную структуру левого векторного пространства, поскольку мы можем положить  $\langle \lambda f, v \rangle = \lambda \langle f, v \rangle$ . В таком случае  $\langle \lambda f, v\mu \rangle = \lambda \langle f, v \rangle \mu$ .

Заметим, что если мы рассматриваем правые векторные пространства над телом, то естественно считать, что матрицы линейных отображений одного пространства в другое образуют левое векторное пространство. Действительно, тогда мы получаем, что  $\lambda Av = A\lambda^{-1}v = Av\lambda$ , т.е. умножение матрицы на  $\lambda$  соответствует умножению вектора на  $\lambda$ ; в коммутативном случае это свойство записывается так:  $A(\lambda v) = (\lambda A)v$ .

Если мы рассматриваем линейные отображения пространства  $V^n$  в себя (линейные операторы), то в качестве второго базиса  $\varepsilon_1, \dots, \varepsilon_m$  мы можем взять первый базис  $e_1, \dots, e_n$  и записать матрицу линейного оператора  $f$  относительно этого базиса. Положим  $e'_j = f(e_j) = \sum e_i a_{ij}$ . Квадратную матрицу  $A = (a_{ij})$  называют *обратимой* или *невырожденной*, если векторы  $e'_1, \dots, e'_n$  образуют базис. В таком случае  $e_j = \sum e'_i b_{ij}$  для некоторых чисел  $b_{ij}$  (эти числа определены однозначно). Поэтому

$$\begin{aligned} e_j &= \sum e'_i b_{ij} = \sum e_k a_{ki} b_{ij}, \\ e'_j &= \sum e_i a_{ij} = \sum e'_k b_{ki} a_{ij}. \end{aligned}$$

Из однозначности разложения вектора по базису получаем  $AB = I$  и  $BA = I$ .

Отметим, что в случае квадратных матриц из равенства  $AB = I$  следует равенство  $BA = I$ . Действительно, нам нужно лишь доказать, что векторы  $e'_j = \sum e_i a_{ij}$  образуют базис. При этом мы знаем, что  $\sum e'_i b_{ij} = \sum e_k a_{ki} b_{ij} = e_j$ , т.е. базисные векторы  $e_j$  линейно выражаются через векторы  $e'_j$ , которых столько же, сколько базисных. Но тогда векторы  $e'_j$  тоже линейно независимы.

Если  $A_1, \dots, A_n$  — строки матрицы  $A$ , то  $i$ -я строка матрицы  $BA$  равна  $b_{i1}A_1 + b_{i2}A_2 + \dots + b_{in}A_n$ . Таким образом, строки матрицы  $BA$  являются левыми линейными комбинациями строк матрицы  $A$ . Поэтому равенство  $BA = I$  эквивалентно тому, что  $n$  единичных векторов можно представить в виде левых линейных комбинаций векторов  $A_1, \dots, A_n$ . Это, в свою очередь, эквивалентно тому, что векторы  $A_1, \dots, A_n$  линейно независимы при умножении слева. Аналогично столбцы невырожденной матрицы линейно независимы при умножении справа.

Правая линейная комбинация столбцов матрицы  $A$  с коэффициентами  $\lambda_1, \dots, \lambda_n$  равна столбцу  $A\lambda$ , где  $\lambda = (\lambda_1, \dots, \lambda_n)^T$ . Левая линейная комбинация строк матрицы  $A$  с коэффициентами  $\mu_1, \dots, \mu_n$  равна строке  $\mu A$ , где  $\mu = (\mu_1, \dots, \mu_n)$ . Поэтому матрица  $A$  вырожденная тогда и только тогда, когда существует ненулевой вектор-столбец  $v$ , для которого  $Av = 0$ . Эквивалентное условие заключается в том, что существует ненулевая вектор-строка  $w$ , для которой  $wA = 0$ .

Невырожденные матрицы порядка  $n$  с элементами из тела  $K$  образуют группу относительно умножения. Эту группу обозначают  $GL(n, K)$  или  $GL_n(K)$ .

Для произвольной матрицы  $A$  (не обязательно квадратной) *левым строчным рангом* называют размерность пространства левых линейных комбинаций строк этой матрицы. Аналогично определяются правый строчный, левый столбцовый и правый столбцовый ранги.

Рассмотрим матрицу  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  с элементами из некоммутативного тела. Столбцы этой матрицы линейно зависимы справа тогда и только тогда, когда система уравнений

$$\begin{aligned} ax + by &= 0, \\ cx + dy &= 0 \end{aligned}$$

имеет ненулевое решение  $(x, y)$ . Будем предполагать, что все числа  $a, b, c, d$  отличны от нуля. Тогда можно выразить  $x$  из первого уравнения и подставить во второе. В результате получим  $(-ca^{-1}b + d)y = 0$ . Таким образом, столбцы рассматриваемой матрицы линейно зависимы справа тогда и только тогда, когда выполняется равенство  $ca^{-1}b - d = 0$ . Обратите внимание, что это равенство не полиномиальное.

Легко проверить, что линейная независимость слева строк рассматриваемой матрицы тоже эквивалентна равенству  $ca^{-1}b = d$ . А для линейной независимости слева столбцов (и линейной независимости справа строк) получаем другое соотношение:  $ba^{-1}c = d$ ; оно соответствует вырожденности транспонированной матрицы.

Выше было показано, что матрица порядка  $n$  невырождена тогда и

только тогда, когда её левый строчный ранг равен  $n$  (эквивалентное условие: правый столбцовый ранг равен  $n$ ). Отметим, что за невырожденность матрицы отвечают именно левый строчный и правый столбцовый ранг. Правый строчный и левый столбцовый ранг с невырожденностью матрицы не связаны. Действительно, если  $bc \neq cb$ , то вырожденная матрица  $\begin{pmatrix} 1 & b \\ c & cb \end{pmatrix}$  имеет правый строчный ранг 2, а невырожденная матрица  $\begin{pmatrix} 1 & b \\ c & bc \end{pmatrix}$  имеет правый строчный ранг 1.

**Теорема 9.1.1.2.** *Левый строчный ранг равен правому столбцовому, а правый строчный ранг равен левому столбцовому.*

**Доказательство.** Пусть  $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$  — матрица линейного отображения  $V \rightarrow W$ ,  $e_1, \dots, e_n$  — базис пространства  $V$ ,  $e_1^*, \dots, e_n^*$  — базис двойственного пространства  $V^*$ . Правый столбцовый ранг матрицы  $A$  равен размерности пространства  $U$  столбцов длины  $m$ , имеющих вид  $A_1x_1 + \dots + A_nx_n$ , где  $A_j = (a_{1j}, \dots, a_{mj})^T$  —  $j$ -й столбец матрицы  $A$ , а  $x_1, \dots, x_n$  — некоторые элементы тела. Пространство  $U$  можно представить и по-другому, исходя из строк матрицы  $A$ . А именно, для каждой строки  $(a_{i1}, \dots, a_{in})$  рассмотрим ковектор  $f_i = a_{i1}e_1^* + \dots + a_{in}e_n^*$ . Ясно, что  $f_i(\sum e_jx_j) = \sum a_{ij}x_j$ , поэтому

$$A_1x_1 + \dots + A_nx_n = (f_1(x), \dots, f_m(x))^T,$$

где  $x = \sum e_jx_j$ . Таким образом,  $U = \text{Im } f$ , где отображение  $f$  задано формулой  $f(x) = (f_1(x), \dots, f_m(x))^T$ . Ядро отображения  $f$  состоит из тех векторов пространства  $V$ , которые аннулируются пространством  $W = \langle f_1, \dots, f_m \rangle \subset V^*$ . Поэтому  $\dim W + \dim \text{Ker } f = \dim V$  (см. §2.1.8), а значит,  $\dim U = \dim V - \dim \text{Ker } f = \dim W$ . Но  $W$  — это пространство, порождённое левыми линейными комбинациями строк матрицы  $A$ .

Равенство правого строчного ранга и левого столбцового ранга матрицы  $A$  следует из равенства левого строчного ранга и правого столбцового ранга матрицы  $A^T$ .  $\square$

**Теорема 9.1.1.3.** *При умножении матрицы слева или справа на невырожденную матрицу её левый строчный (и правый столбцовый) ранг сохраняется, а правый строчный (и левый столбцовый) ранг может измениться.*

**Доказательство.** Согласно теореме 9.1.1.2 левый строчный ранг равен правому столбцовому рангу, поэтому достаточно проверить, что если  $P$  и  $Q$  — невырожденные матрицы, то матрица  $PA$  имеет такой же левый

строчный ранг, как и матрица  $A$ , а матрица  $AQ$  имеет такой же правый столбцовый ранг, как и матрица  $A$ . Строки матрицы  $PA$  являются левыми линейными комбинациями строк матрицы  $A$ , поэтому левый строчный ранг у матрицы  $PA$  не больше, чем у матрицы  $A$ . Аналогично левый строчный ранг у матрицы  $P^{-1}PA = A$  не больше, чем у матрицы  $PA$ . Следовательно, левые строчные ранги матриц  $PA$  и  $A$  равны. Для матрицы  $AQ$  доказательство аналогично.

Пример, когда правый строчный ранг матрицы при умножении на невырожденную матрицу не сохраняется, можно построить следующим образом. Напомним, что если  $bc \neq cb$ , то матрица  $\begin{pmatrix} 1 & b \\ c & bc \end{pmatrix}$  невырожденная и имеет правый строчный ранг 1. Поэтому при умножении этой матрицы на единичную матрицу (ранг которой во всех смыслах равен её порядку) мы получаем матрицу, правый строчный ранг которой равен 1, а не 2.  $\square$

Такие же рассуждения, как и при доказательстве теоремы 2.2.3.2, показывают, что для любой матрицы  $A$  размера  $m \times n$  можно выбрать невырожденную матрицу  $L_A$  (порядка  $m$ ) и невырожденную матрицу  $R_A$  (порядка  $n$ ) так, что  $L_A A R_A = I_a$  — единичная матрица некоторого порядка  $a$ , дополненная нулями до размеров матрицы  $A$ .

### 9.1.2. Элементарные преобразования матриц

Рассмотрим матрицы фиксированного порядка  $n$  с элементами из тела  $K$ . Пусть  $E_{ij}$  — матрица, у которой  $a_{ij} = 1$ , а все остальные элементы равны 0. Для всех  $i \neq j$  и  $\lambda \in K$  рассмотрим матрицы  $B_{ij}(\lambda) = I + \lambda E_{ij}$ . Такие матрицы будем называть *элементарными*, а умножение (справа или слева) на элементарную матрицу будем называть *элементарным преобразованием*. Матрица  $B_{ij}(\lambda)A$  получается из матрицы  $A$  прибавлением к  $i$ -й строке матрицы  $A$  её  $j$ -й строки, умноженной слева на  $\lambda$ , а матрица  $AB_{ij}(\lambda)$  получается из матрицы  $A$  прибавлением к  $j$ -му столбцу матрицы  $A$  её  $i$ -го столбца, умноженного справа на  $\lambda$ . Из этого, в частности, следует, что  $B_{ij}(\lambda)^{-1} = B_{ij}(-\lambda)$ . Поэтому произведения матриц вида  $B_{ij}(\lambda)$  (для всех  $i \neq j$  и  $\lambda \in K$ ) образуют группу; эту группу обозначают  $SL(n, K)$  или  $SL_n(K)$ . Следующая теорема показывает, в частности, что если  $K$  — поле, то  $SL(n, K)$  — группа матриц с определителем 1.

**Теорема 9.1.2.1.** Любую невырожденную матрицу  $A$  можно представить в виде  $A = BD(\mu)$ , где  $B \in SL(n, K)$  и  $D(\mu) = \text{diag}(1, \dots, 1, \mu)$ .

**Доказательство.** Матрица  $A$  невырождена, поэтому в первом столбце найдётся элемент  $a_{i1} \neq 0$ . Если  $a_{21} = 0$ , то мы прибавим ко второй стро-

ке  $i$ -ю строку и получим матрицу, у которой  $a_{21} \neq 0$ . Теперь умножим слева вторую строку на  $(1 - a_{11})a_{21}^{-1}$  и прибавим её к первой строке. В результате получим матрицу, у которой  $a_{11} = 1$ . Затем умножим (слева) первую строку на  $a_{i1}$  и вычтем её из  $i$ -й строки для всех  $i > 1$ . В результате получим невырожденную матрицу, у которой  $a_{11} = 1$  и  $a_{i1} = 0$  при  $i > 1$ .

Строки  $A_2, \dots, A_n$  полученной матрицы линейно независимы слева. Поэтому (если  $n \geq 3$ ) можно аналогично преобразовать второй столбец и получить матрицу, у которой  $a_{22} = 1$  и  $a_{i2} = 0$  при  $i > 2$ . Затем из первой строки можно вычесть вторую строку, умноженную слева на некоторое число, и получить матрицу с  $a_{12} = 0$ .

Такие операции можно повторять до тех пор, пока не получится матрица, которая отличается от единичной матрицы только последним столбцом. В этой матрице элемент  $a_{nn} = \mu$  отличен от нуля, потому что иначе последняя строка матрицы состояла бы из нулей, а такая матрица вырожденная. Вычитая из первых  $n - 1$  строк последнюю строку, умноженную слева на подходящие числа, получаем требуемую матрицу  $D(\mu)$ .  $\square$

**Теорема 9.1.2.2 (лемма Уайтхеда).** Пусть  $A, B \in \text{GL}(n, K)$ . Тогда матрицы  $AB \oplus I_n = \begin{pmatrix} AB & 0 \\ 0 & I_n \end{pmatrix}$  и  $BA \oplus I_n = \begin{pmatrix} BA & 0 \\ 0 & I_n \end{pmatrix}$  получают друг из друга посредством композиции элементарных преобразований одного типа (например, умножений справа на элементарные матрицы).

**Доказательство.** Нужно доказать, что матрица  $(BA)^{-1}AB \oplus I_n = A^{-1}B^{-1}AB \oplus I_n$  лежит в  $\text{SL}(2n, K)$ . Прежде всего заметим, что любая треугольная матрица с единичной диагональю представляется в виде произведения  $p$  элементарных матриц, где  $p$  — число ненулевых элементов вне диагонали. Поэтому для любой матрицы  $C \in \text{GL}(n, K)$  матрица

$$\begin{pmatrix} 0 & C \\ -C^{-1} & 0 \end{pmatrix} = \begin{pmatrix} I & C \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ -C^{-1} & I \end{pmatrix} \begin{pmatrix} I & C \\ 0 & I \end{pmatrix} \quad (9.1)$$

лежит в  $\text{SL}(2n, K)$ . Следовательно, матрица

$$\begin{pmatrix} C^{-1} & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} 0 & C \\ -C^{-1} & 0 \end{pmatrix}$$

тоже лежит в  $\text{SL}(2n, K)$ , поскольку обе матрицы в правой части имеют такой вид, как матрица в левой части (9.1). Наконец, матрица

$$\begin{pmatrix} A^{-1}B^{-1}AB & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A^{-1} & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} B^{-1} & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} AB & 0 \\ 0 & (AB)^{-1} \end{pmatrix}$$

лежит в  $\text{SL}(2n, K)$ .  $\square$

### 9.1.3. Нормальная форма Брюа

Назовём квадратную матрицу *верхней унитарной*, если все её диагональные элементы равны 1, а все элементы ниже диагонали равны 0. Назовём квадратную матрицу *нижней унитарной*, если все её диагональные элементы равны 1, а все элементы выше диагонали равны 0.

Умножение слева на нижнюю унитарную матрицу заключается в том, что левые кратные некоторых строк прибавляются к последующим строкам, а умножение справа на верхнюю унитарную матрицу заключается в том, что правые кратные некоторых столбцов прибавляются к последующим столбцам.

Любая верхняя (нижняя) унитарная матрица является произведением элементарных матриц, поэтому она лежит в группе  $SL(n, K)$ . Ясно также, что верхние (нижние) унитарные матрицы образуют группу относительно умножения.

**Теорема 9.1.3.1 (нормальная форма Брюа).** *Любую матрицу  $A$  размера  $t \times n$  можно представить в виде*

$$A = LMU, \quad (9.1)$$

где  $L$  — нижняя унитарная матрица порядка  $t$ ,  $U$  — верхняя унитарная матрица порядка  $n$ , а  $M$  — матрица размера  $t \times n$ , у которой в каждой строке и в каждом столбце есть не более одного ненулевого элемента. При этом матрица  $M$  однозначно определяется матрицей  $A$ .

**Доказательство.** Пусть первая ненулевая строка матрицы  $A$  имеет номер  $i$ , а первый ненулевой элемент этой строки — это  $a_{ij}$ . Прибавляя левые кратные  $i$ -й строки ко всем последующим строкам, мы можем добиться того, что в  $j$ -м столбце все элементы, кроме  $a_{ij}$ , станут равны 0. Эти операции соответствуют умножению слева на нижнюю унитарную матрицу. Затем, прибавляя правые кратные  $j$ -го столбца ко всем последующим столбцам, мы можем добиться того, что в  $i$ -й строке все элементы, кроме  $a_{ij}$ , станут равны 0. Эти операции соответствуют умножению справа на верхнюю унитарную матрицу. В результате мы получили матрицу, у которой в  $i$ -й строке и в  $j$ -м столбце есть только один ненулевой элемент  $a_{ij}$ , причём все строки выше этого элемента нулевые. Выберем первую ненулевую строку ниже  $i$ -й и повторим эти операции. После нескольких таких шагов мы приведём матрицу  $A$  к матрице  $M$  требуемого вида.

Докажем теперь, что матрица  $M$  однозначно задаётся матрицей  $A$ . Пусть есть два разных разложения (9.1):  $LMU = L'M'U'$ . Тогда  $M =$

$L_1 M' U_1$ , где  $L_1 = L^{-1} L'$  и  $M_1 = U' U^{-1}$  — унитреугольные матрицы (нижняя и верхняя). Пусть  $m'_{ij} \neq 0$ . Тогда в  $i$ -й строке и в  $j$ -м столбце матрицы  $M'$  все элементы, кроме  $m'_{ij}$ , нулевые. При умножении матрицы  $M'$  слева на нижнюю треугольную матрицу  $L_1$  к элементу  $m'_{ij}$  прибавляются левые кратные элементов  $j$ -го столбца, расположенных выше  $i$ -й строки. Все эти элементы равны 0, поэтому  $m'_{ij}$  не изменяется. Аналогично  $m'_{ij}$  не изменяется и при умножении справа на верхнюю треугольную матрицу  $M_1$ . Следовательно,  $m_{ij} = m'_{ij}$ . Таким образом, каждому ненулевому элементу  $m'_{ij}$  матрицы  $M'$  соответствует равный ему элемент  $m_{ij}$  матрицы  $M$ . Аналогично каждому ненулевому элементу  $m_{ij}$  матрицы  $M$  соответствует равный ему элемент  $m'_{ij}$  матрицы  $M'$ .  $\square$

Разложение (9.1) называют *разложением Брюа*.

**Замечание.** Ранг (в любом смысле) матрицы  $M$  из разложения Брюа равен количеству её ненулевых элементов. Матрицы  $L$  и  $U$  невырожденные. Поэтому согласно теореме 9.1.1.3 левый строчный (и правый столбцовый) ранг матрицы  $A$  равен количеству ненулевых элементов матрицы  $M$ .

#### 9.1.4. Коммутант группы $GL(n, K)$

Напомним, что *коммутант* группы  $G$  — это подгруппа  $G' \subset G$ , порождённая элементами вида  $aba^{-1}b^{-1}$ ,  $a, b \in G$ .

**Теорема 9.1.4.1.** Если  $n \geq 2$ , то матрица  $D(\mu) = \text{diag}(1, \dots, 1, \mu)$ , где  $\mu = aba^{-1}b^{-1}$  и  $a, b \in K$ , лежит в группе  $SL(n, K)$ .

**Доказательство.** Запишем последовательность элементарных преобразований (которые заключаются в прибавлении к одной строке левого кратного другой строки), переводящую единичную матрицу в матрицу  $D(\mu)$ :

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &\rightarrow \begin{pmatrix} 1 & 0 \\ a^{-1} & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -a \\ a^{-1} & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -a \\ a^{-1} & b^{-1} \end{pmatrix} \rightarrow \\ &\rightarrow \begin{pmatrix} aba^{-1} & 0 \\ a^{-1} & b^{-1} \end{pmatrix} \rightarrow \begin{pmatrix} aba^{-1} & 0 \\ 1 & b^{-1} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -\mu \\ 1 & b^{-1} \end{pmatrix} \rightarrow \\ &\rightarrow \begin{pmatrix} 0 & -\mu \\ 1 & \mu \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 1 & \mu \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}. \end{aligned}$$

Эти преобразования записаны для матриц порядка 2, но их можно записать и для матриц произвольного порядка, подразумевая, что в преобразованиях участвуют только две последние строки.  $\square$



**Теорема 9.1.4.2.** Для любого  $n \geq 2$  коммутантом группы  $GL(n, K)$  является группа  $SL(n, K)$ , за исключением того случая, когда  $n = 2$  и  $K = \mathbb{F}_2$ .

**Доказательство.** Докажем сначала, что коммутант группы  $GL(n, K)$  всегда содержится в  $SL(n, K)$ . Легко проверить, что если  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ , то  $DB_{ij}(\lambda)D^{-1} = B_{ij}(\lambda_i\lambda_j^{-1})$ . Поэтому для любой диагональной матрицы  $D$  и любой матрицы  $B \in SL(n, K)$  матрица  $DBD^{-1}$  тоже лежит в  $SL(n, K)$ , т.е.  $DB = B'D$ , где  $B' \in SL(n, K)$ .

Согласно теореме 9.1.2.1 любую невырожденную матрицу  $A$  можно представить в виде  $A = BD(\mu)$ , где  $B \in SL(n, K)$  и  $D(\mu) = \text{diag}(1, \dots, 1, \mu)$ . Поэтому

$$A_1^{-1}A_2^{-1}A_1A_2 = D_1^{-1}B_1^{-1}D_2^{-1}B_2^{-1}B_1D_1B_2D_2 = D_1^{-1}D_2^{-1}D_1D_2B,$$

где  $B \in SL(n, K)$ . Если  $D_1 = D(\mu_1)$  и  $D_2 = D(\mu_2)$ , то согласно теореме 9.1.4.1 матрица  $D_1^{-1}D_2^{-1}D_1D_2 = D(\mu_1^{-1}\mu_2^{-1}\mu_1\mu_2)$  лежит в  $SL(n, K)$ .

Докажем теперь, что если  $K \neq \mathbb{F}_2$  или  $n > 2$ , то любую матрицу из  $SL(n, K)$  можно представить в виде произведения коммутаторов. Для этого достаточно проверить, что любую матрицу  $B_{ij}(\lambda)$  можно представить в виде коммутатора; более того, можно ограничиться матрицей  $B_{12}(\lambda)$  порядка 2. Тожество

$$\begin{pmatrix} 1 & 0 \\ 0 & \mu^{-1} \end{pmatrix} \begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x(1-\mu) \\ 0 & 1 \end{pmatrix}$$

показывает, что нужно подобрать  $x$  и  $\mu \neq 0$  так, чтобы выполнялось равенство  $x(1-\mu) = \lambda$ . Если  $K \neq \mathbb{F}_2$ , то мы выбираем произвольный элемент  $\mu$ , отличный от 0 и 1, и полагаем  $x = \lambda(1-\mu)^{-1}$ .

Если же  $n \geq 3$ , то мы воспользуемся тем, что

$$\begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix}.$$

□

## Задачи

**1.1.** Пусть матрица  $B$  получена из матрицы  $A$  прибавлением к  $i$ -й строке  $j$ -й строки, умноженной слева на  $\lambda$ , причём  $i \neq j$ . Докажите, что матрица  $B^{-1}$  получена из матрицы  $A^{-1}$  вычитанием из  $j$ -го столбца  $i$ -го столбца, умноженного справа на  $\lambda$ .

**1.2.** Пусть тело  $K$  обладает следующим свойством: матрица  $A$  с элементами из тела  $K$  обратима тогда и только тогда, когда обратима матрица  $A^T$ . Докажите, что тогда тело  $K$  является полем.

**1.3.** Докажите, что матрица

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 1 & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

с элементами из любого некоммутативного кольца обратима.

**1.4.** Докажите, что треугольная матрица с элементами из некоммутативного тела невырожденная тогда и только тогда, когда все её диагональные элементы ненулевые.

**1.5.** Докажите, что если элемент  $x$  некоммутативного тела удовлетворяет уравнению  $x^n + a_{n-1}x^{n-1} + \dots + a_0$ , то матрица

$$\begin{pmatrix} x & -1 & 0 & \dots & 0 \\ 0 & x & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 \\ a_0 & a_1 & a_2 & \dots & x + a_{n-1} \end{pmatrix}$$

вырожденная.

**1.6.** Пусть  $A$  — матрица размера  $m \times n$ , где  $m < n$ . Докажите, что существует ненулевой вектор-столбец  $x$ , для которого  $Ax = 0$ .

**1.7.** Докажите, что матрица  $A$  регулярна слева (т.е. из равенства  $XA = 0$  всегда следует, что  $X = 0$ ) тогда и только тогда, когда у неё есть правая обратная матрица  $B$  (для которой  $AB = I$ ).

**1.8.** Левый строчный ранг матрицы  $A$  порядка  $n$  равен  $r$ . Докажите, что матрицу  $B$  порядка  $n$  можно выбрать так, чтобы её левый строчный ранг был бы равен  $n - r$  и матрица  $A + B$  была бы невырожденной.

## 9.2. Определители и собственные значения

### 9.2.1. Определитель Дьёдонне

Для матриц, элементы которых принадлежат некоммутативному телу, обычное определение определителя теряет смысл, потому что теперь

надо фиксировать порядок множителей в произведениях элементов, а сделать это естественным образом не удаётся.

Подход Дьёдонне [Di] к вычислению определителя матрицы с элементами из некоммутативного тела  $K$  основан на том, что матрицу  $A \in \mathrm{GL}(n, K)$  можно представить в виде произведения матриц из  $\mathrm{SL}(n, K)$  и диагональной матрицы. Такие представления бывают разные. Во-первых, есть представление  $A = BD(\mu)$ , где  $B \in \mathrm{SL}(n, K)$  и  $D(\mu) = \mathrm{diag}(1, \dots, 1, \mu)$  (теорема 9.1.2.1). Во-вторых, есть разложения Брюа  $A = LMU$ .

Естественно предположить, что определитель любой матрицы из  $\mathrm{SL}(n, K)$  равен 1, а потому определитель матрицы  $A$  зависит только от  $\mu$  в первом разложении (и от матрицы  $M$  во втором разложении). Но просто положить  $\det A = \mu$  мешает теорема 9.1.4.1, согласно которой матрица  $D(\mu) = \mathrm{diag}(1, \dots, 1, \mu)$ , где  $\mu = aba^{-1}b^{-1}$ , лежит в группе  $\mathrm{SL}(n, K)$ .

Поэтому мы будем считать, что  $\det A$  (для невырожденной матрицы  $A$ ) принимает значения в абелевой группе  $K^{\mathrm{ab}}$  (относительно умножения), которая строится по телу  $K$  следующим образом. Пусть  $K^\times$  — мультипликативная группа тела  $K$  (т.е. группа ненулевых элементов  $K$  относительно умножения), а  $K^{\times'}$  — её коммутант. Положим  $K^{\mathrm{ab}} = K^\times / K^{\times'}$ . Тогда каждому элементу  $a \in K$  соответствует его класс эквивалентности  $\bar{a} \in K^{\mathrm{ab}}$ .

**Замечание.** Обратите внимание, что элемент  $\overline{-1}$  не обязательно равен  $-1$ . Например, для тела кватернионов  $\overline{-1} = 1$  (см. теорему 9.3.7.1)

Определитель невырожденной матрицы  $A$  можно было бы определить как  $\bar{\mu}$ , где число  $\mu$  берётся из разложения  $A = BD(\mu)$ . Но при таком подходе нужно проверять, что число  $\bar{\mu}$  не зависит от выбора разложения  $A = BD(\mu)$ , поскольку это разложение не единственно. Удобнее задать определитель с помощью аксиом и построить его индукцией по порядку матрицы (из этих аксиом следует, в частности, что  $\det(BD(\mu)) = \bar{\mu}$ ). При построении определителя матрицы порядка  $n$  нам понадобятся свойства определителя матрицы порядка  $n - 1$ . Поэтому схема рассуждений будет такая. Сначала мы выведем некоторые следствия из аксиом, затем построим определитель матрицы порядка  $n$  на основе определителя матрицы порядка  $n - 1$  (используя при этом его свойства, доказанные перед этим), а в конце мы проверим, что определитель матрицы порядка  $n$  тоже удовлетворяет этим аксиомам.

Мы потребуем, чтобы *определитель Дьёдонне*  $\det A$  удовлетворял следующим аксиомам.

I. Если матрица  $A'$  получена из матрицы  $A$  умножением одной строки матрицы  $A$  слева на  $\mu$ , то  $\det A' = \bar{\mu} \det A$ .

II. Если матрица  $A'$  получена из матрицы  $A$  прибавлением одной строки к другой, то  $\det A' = \det A$ .

III. Определитель единичной матрицы равен 1.

Выведем теперь некоторые следствия из этих аксиом.

1) Если к строке  $A_i$  прибавить другую строку  $A_j$ , умноженную слева на  $\lambda$ , то определитель при этом не изменится.

Если  $\lambda = 0$ , то не изменяется даже сама матрица, поэтому будем считать, что  $\lambda \neq 0$ . Пусть матрица  $A'$  получается из матрицы  $A$  заменой строки  $A_j$  на  $\lambda A_j$ . Тогда  $\det A = \lambda^{-1} \det A'$ . Прибавив к строке  $\lambda A_j$  матрицы  $A'$  строку  $A_i$  и вынеся из строки  $\lambda A_j$  множитель  $\lambda$ , получаем требуемое.

2) Если строки  $A_i$  и  $A_j$  поменять местами, то определитель матрицы  $A$  умножится на  $\overline{-1}$ .

Заменяем строку  $A_i$  на  $A_i + A_j$  и вычтем эту строку из  $A_j$ . В результате строка  $A_j$  заменится на  $-A_i$ . Прибавим эту строку к строке с номером  $i$ . В результате строки  $A_i$  и  $A_j$  поменяются местами и одна из них изменит знак.

3) Если  $A = BD(\mu)$ , где  $B \in \text{SL}(n, K)$ , то  $\det A = \bar{\mu}$ .

Прежде всего заметим, что  $\det D(\mu) = \bar{\mu}$ . Действительно, если из последней строки вынести  $\mu$ , то останется единичная матрица, определитель которой равен 1. Ясно также, что умножение справа на матрицу  $B \in \text{SL}(n, K)$  не изменяет определитель.

4)  $\det(AB) = \det A \cdot \det B$ .

Пусть  $A = CD(\mu)$ , где  $C \in \text{SL}(n, K)$ . Матрица  $D(\mu)B$  получается из матрицы  $B$  умножением последней строки на  $\mu$ , поэтому  $\det(D(\mu)B) = \bar{\mu} \det B$ . При умножении матрицы  $D(\mu)B$  слева на матрицу  $C \in \text{SL}(n, K)$  определитель не изменяется. Поэтому

$$\det(AB) = \det(D(\mu)B) = \bar{\mu} \det B = \det A \cdot \det B.$$

5) Если к одному столбцу матрицы прибавить другой столбец, умноженный справа на  $\lambda$ , то определитель при этом не изменится.

Действительно, прибавление  $i$ -го столбца, умноженного на  $\lambda$ , к  $j$ -му столбцу соответствует замене матрицы  $A$  на матрицу  $AB_{ij}(\lambda)$ . Но  $\det(AB_{ij}(\lambda)) = \det A \det B_{ij}(\lambda) = \det A$ .

6) Если столбец матрицы умножить справа на  $\mu$ , то определитель умножится на  $\bar{\mu}$ .

Для последнего столбца это следует из равенства  $\det(AD(\mu)) = \det A \cdot \bar{\mu}$ . Для остальных столбцов оно будет доказано, если мы докажем, что при перестановке двух столбцов определитель умножается на  $\overline{-1}$ . Это свойство можно попробовать доказать точно так же, как мы доказывали аналогичное свойство для строк. Но в действительности это доказа-

тельство использует свойство б) для умножения на  $\overline{-1}$ . Поэтому мы поступим следующим образом. Правило перестановки последнего столбца с любым другим доказать можно, поскольку свойство б) для последнего столбца доказано. Используя перестановки последнего столбца с остальными столбцами, мы доказываем свойство б). А теперь уже можно доказать и правило перестановки для любых столбцов.

Перейдём теперь к построению определителя. Для матрицы  $A = (a)$  порядка 1 мы полагаем  $\det A = \bar{a}$ . Предположим теперь, что определитель матрицы порядка  $n - 1$  уже построен. Рассмотрим невырожденную матрицу  $A$  порядка  $n$ . Её строки  $A_1, \dots, A_n$  линейно независимы слева, поэтому существуют однозначно определённые элементы  $\lambda_1, \dots, \lambda_n$ , для которых  $\lambda_1 A_1 + \dots + \lambda_n A_n = (1, 0, \dots, 0)$ . Запишем каждую строку в виде  $A_i = (a_{i1}, B_i)$ , где  $B_i$  — строка длины  $n - 1$ . Тогда  $\sum_{i=1}^n \lambda_i a_i = 1$  и  $\sum_{i=1}^n \lambda_i B_i = 0$ .

Для каждого  $i = 1, \dots, n$  рассмотрим матрицу  $C_i$  порядка  $n - 1$ , которая составлена из всех строк  $B_1, \dots, B_n$ , кроме строки  $B_i$ . Выберем  $\lambda_i \neq 0$  (очевидно, что все элементы  $\lambda_i$  не могут быть равны нулю одновременно) и положим

$$\det A = \overline{(-1)^{i+1} \lambda_i^{-1}} \det C_i.$$

Прежде всего нужно проверить, что эта величина не зависит от выбора  $\lambda_i \neq 0$ . Предположим, что  $\lambda_i \neq 0$  и  $\lambda_j \neq 0$ , причём  $i \neq j$ . Пусть  $P$  и  $Q$  — матрицы, которые получаются из матрицы  $C_i$  заменой строки  $B_j$  на  $\lambda_j B_j$  и на  $B_i$  соответственно. Ясно, что  $\det C_i = \overline{\lambda_j^{-1}} \det P$ . Если в матрице  $P$  к строке  $\lambda_j B_j$  прибавить остальные строки, умноженные слева на соответствующие элементы  $\lambda_k$ , то вместо этой строки мы получим  $\sum_{k \neq i} \lambda_k B_k = -\lambda_i B_i$ . Следовательно,  $\det C_i = \overline{\lambda_j^{-1} (-\lambda_i)} \det Q$ . Но матрицу  $Q$  можно преобразовать в матрицу  $C_j$ , сделав  $|i - j| - 1$  перестановок пар строк. Поэтому

$$\det C_i = \overline{(-1)^{i-j-1} \lambda_j^{-1} (-\lambda_i)} \det C_j,$$

т.е.

$$\overline{(-1)^{i+1} \lambda_i^{-1}} \det C_i = \overline{(-1)^{j+1} \lambda_j^{-1}} \det C_j,$$

что и требовалось.

Теперь нужно проверить, что для построенного таким образом определителя выполняются аксиомы I, II и III.

I. Пусть строка  $A_i$  матрицы  $A$  заменяется на  $\mu A_i$ . Тогда для полученной матрицы  $A'$  элемент  $\lambda_i$  заменяется на  $\lambda_i \mu^{-1}$ , а все остальные элемен-

ты  $\lambda_k$  не изменяются. Если  $\lambda_i \neq 0$ , то матрица  $C_i$  не изменяется. Поэтому

$$\det A' = \overline{(-1)^{i+1}(\lambda_i \mu^{-1})^{-1}} \det C_i = \overline{\mu(-1)^{i+1} \lambda_i^{-1}} \det C_i = \bar{\mu} \det A.$$

II. Если строка  $A_i$  заменяется на  $A_i + A_j$ , то  $\lambda_j$  заменяется на  $\lambda_j - \lambda_i$ , а все остальные элементы  $\lambda_k$  не изменяются, потому что

$$\lambda_i(A_i + A_j) + (\lambda_j - \lambda_i)A_j = \lambda_i A_i + \lambda_j A_j.$$

Если  $\lambda_k \neq 0$  для некоторого  $k$ , отличного от  $i$  и  $j$ , то именно этот элемент  $\lambda_k$  мы применим для вычисления определителя. Ясно, что при этом  $\det C_k$  не изменяется, так как к одной строке прибавляется другая строка. Если  $\lambda_i \neq 0$ , то для вычисления определителя можно применить и этот элемент: ясно, что элемент  $\lambda_i$  и матрица  $C_i$  не изменяются. Осталось рассмотреть случай, когда  $\lambda_j \neq 0$ , а все остальные элементы  $\lambda_k$  равны нулю. В этом случае  $\lambda_j B_j = 0$ , а значит,  $B_j = 0$ . Строка  $B_i$  заменяется на  $B_i + B_j$ , поэтому матрица  $C_j$  не изменяется. Элемент  $\lambda_j$  заменяется на  $\lambda_j - \lambda_i$ ; он тоже не изменяется.

III. Для единичной матрицы  $\lambda_1 = 1$ , а все остальные  $\lambda_k$  равны 0. Матрица  $C_1$  единичная.

Вычислим для примера определитель Дьёдонне невырожденной матрицы  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . Если  $a \neq 0$ , то

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ ca^{-1} & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & d - ca^{-1}b \end{pmatrix} \begin{pmatrix} 1 & a^{-1}b \\ 0 & 1 \end{pmatrix},$$

поэтому  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \overline{ad - aca^{-1}b}$ . А в случае, когда  $a = 0$ , мы получаем

$$\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ db^{-1} & 1 \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

поэтому  $\det \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = \overline{-bc}$ .

**Теорема 9.2.1.1.** При  $n \geq 2$  определитель Дьёдонне матрицы  $A \in \text{GL}(n, K)$  равен 1 тогда и только тогда, когда  $A \in \text{SL}(n, K)$ .

**Доказательство.** Представим матрицу  $A$  в виде  $A = BD(\mu)$ , где  $B \in \text{SL}(n, K)$ . Тогда  $\det A = \bar{\mu}$ , поэтому  $\det A = 1$  тогда и только тогда, когда элемент  $\mu$  принадлежит коммутанту группы  $K^\times$ , т.е. он равен произведению коммутаторов. Но в таком случае  $D(\mu) \in \text{SL}(n, K)$  согласно теореме 9.1.4.1.  $\square$

**Теорема 9.2.1.2.** Пусть  $A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$ , где  $B$  и  $D$  — квадратные матрицы. Тогда  $\det A = \det B \cdot \det D$ .

**Доказательство.** Можно считать, что матрица  $B$  невырожденная, поскольку иначе матрица  $A$  вырожденная. Вычитая из строк матрицы  $C$  подходящие левые кратные строк матрицы  $B$ , мы можем преобразовать матрицу  $A$  к виду  $\begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$ . При таких преобразованиях определитель не изменяется. Умножая полученную матрицу на матрицу из группы  $SL$ , можно получить матрицу  $\begin{pmatrix} D(\mu) & 0 \\ 0 & D(\nu) \end{pmatrix}$ , где  $\bar{\mu} = \det B$  и  $\bar{\nu} = \det D$ . Вынося из соответствующих строк матрицы  $\begin{pmatrix} D(\mu) & 0 \\ 0 & D(\nu) \end{pmatrix}$  множители  $\mu$  и  $\nu$ , получаем требуемое равенство:  $\det A = \bar{\mu}\bar{\nu} = \det B \cdot \det D$ .  $\square$

**Замечание.** Для матрицы  $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$  тоже имеет место равенство  $\det A = \det B \cdot \det D$ . Доказывается оно аналогично.

### 9.2.2. Тождество Капелли

Тождество Капелли играет существенную роль в теории инвариантов. Это тождество связывает разные композиции дифференцирований и умножений на переменные, но формально его можно рассматривать как тождество для определителей матриц, элементы одной из которых некоммутирующие. Под определителем матрицы с некоммутирующими элементами здесь подразумевается следующее выражение:

$$\det \|a_{ij}\|_1^n = \sum_{\sigma} (-1)^{\sigma} a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}. \quad (9.1)$$

В некоммутативной ситуации существен порядок, в котором берутся множители. Здесь выбран порядок, соответствующий разложению по столбцу. Поскольку порядок множителей выбран одним из многих возможных способов, такой определитель не обладает, вообще говоря, никакими хорошими свойствами (например, определитель произведения матриц не обязательно равен произведению определителей), но в некоторых случаях бывает полезно рассматривать определители такого вида.

В тождестве Капелли участвуют композиции умножений на переменные  $x_{ij}$ , где  $i$  и  $j$  изменяются от 1 до  $n$ , и дифференцирований  $\frac{\partial}{\partial x_{ij}}$ . Для начала рассмотрим простейший нетривиальный случай  $n = 2$ . Чтобы упростить формулы, положим  $x_{k1} = x_k$  и  $x_{k2} = y_k$ . Пусть  $F = x_1^{a_1} x_2^{a_2} y_1^{b_1} y_2^{b_2}$ . Рассмотрим оператор

$$D_{ij}(F) = \sum_{k=1}^n x_{ki} \frac{\partial F}{\partial x_{kj}}.$$

Ясно, что

$$\begin{aligned} D_{12}D_{21}(F) &= \sum_k y_k \frac{\partial}{\partial x_k} \left( \sum_l x_l \frac{\partial F}{\partial y_l} \right) = \\ &= \sum_k y_k \frac{\partial F}{\partial y_k} + \sum_{k,l} y_k x_l \frac{\partial^2 F}{\partial y_l \partial x_k} = bF + \sum_{k,l} y_k x_l \frac{\partial^2 F}{\partial y_l \partial x_k}, \end{aligned}$$

где  $b = b_1 + b_2$ . Положим  $a = a_1 + a_2$  и рассмотрим выражение

$$abF = a \left( \sum_l y_l \frac{\partial F}{\partial y_l} \right) = \sum_{k,l} x_k y_l \frac{\partial^2 F}{\partial x_k \partial y_l}.$$

Сравнивая это равенство с предыдущим, получаем

$$(a+1)bF = D_{12}D_{21}(F) + (x_1y_2 - x_2y_1) \left( \frac{\partial^2 F}{\partial x_1 \partial y_2} - \frac{\partial^2 F}{\partial x_2 \partial y_1} \right).$$

Воспользовавшись тем, что  $(D_{11}+1)D_{22}(F) = (a+1)bF$ , полученное тождество можно записать в виде

$$\begin{vmatrix} D_{11}+1 & D_{12} \\ D_{21} & D_{22} \end{vmatrix} (F) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \cdot \begin{vmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial y_2} \end{vmatrix} (F).$$

В левой части стоит определитель матрицы с некоммутирующими элементами, и его нужно понимать именно так, как было указано выше.

В общем случае *тождество Капелли* имеет вид

$$\begin{vmatrix} D_{11}+n-1 & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22}+n-2 & \dots & D_{2n} \\ \dots & \dots & \dots & \dots \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{vmatrix} = \begin{vmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{nn} \end{vmatrix} \cdot \begin{vmatrix} \frac{\partial}{\partial x_{11}} & \dots & \frac{\partial}{\partial x_{1n}} \\ \dots & \dots & \dots \\ \frac{\partial}{\partial x_{n1}} & \dots & \frac{\partial}{\partial x_{nn}} \end{vmatrix}.$$

Чтобы доказать это тождество, рассмотрим (следуя [Fu]) вспомогательные операторы

$$\Delta_{i_1 j_1} \dots \Delta_{i_m j_m} F = \sum_{k_1, \dots, k_m=1}^n x_{k_1 i_1} \dots x_{k_m i_m} \frac{\partial^m F}{\partial x_{k_1 j_1} \dots \partial x_{k_m j_m}}.$$

Обратите внимание, что так обозначается оператор в целом, а не композиция операторов  $\Delta_{ij} = D_{ij}$ . Но при этом, как легко видеть, порядок членов в выражении  $\Delta_{i_1 j_1} \dots \Delta_{i_m j_m}$  несуществен. Поэтому можно рассмотреть матрицу  $\|\Delta_{ij}\|_1^n$  с коммутирующими элементами и формально записать её определитель:

$$|\Delta_{ij}|_1^n = \sum_{\sigma} (-1)^{\sigma} \Delta_{1\sigma(1)} \dots \Delta_{n\sigma(n)}.$$



Матрицу  $\|\Delta_{ij}\|_1^n$  можно представить в виде произведения матриц  $\|x_{ij}\|_1^n \cdot \|\partial/\partial x_{ij}\|_1^n$ . Поэтому

$$|\Delta_{ij}|_1^n = |x_{ij}|_1^n \cdot |\partial/\partial x_{ij}|_1^n.$$

Таким образом, нужно доказать тождество

$$\begin{vmatrix} D_{11} + n - 1 & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22} + n - 2 & \dots & D_{2n} \\ \dots & \dots & \dots & \dots \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{vmatrix} = \begin{vmatrix} \Delta_{11} & \Delta_{12} & \dots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \dots & \Delta_{2n} \\ \dots & \dots & \dots & \dots \\ \Delta_{n1} & \Delta_{n2} & \dots & \Delta_{nn} \end{vmatrix}.$$

Его можно рассматривать как формальное тождество для определителей двух матриц (у одной из которых элементы некоммутирующие); для доказательства этого тождества достаточно знать, что элементы данных матриц связаны следующими соотношениями:

$$D_{ij}D_{kl} = \begin{cases} D_{ij}\Delta_{kl} = \Delta_{ij}\Delta_{kl} & \text{при } j \neq k; \\ \Delta_{ij}\Delta_{kl} + D_{il} & \text{при } j = k; \end{cases}$$

и вообще, если  $j$  отлично от  $a_1, \dots, a_r$ , то

$$D_{ij}\Delta_{a_1b_1} \dots \Delta_{a_rb_r} = \Delta_{ij}\Delta_{a_1b_1} \dots \Delta_{a_rb_r}, \quad (9.2)$$

а если  $j$  совпадает ровно с одним из чисел  $a_1, \dots, a_r$  (а именно, с  $a_k$ ), то

$$D_{ij}\Delta_{a_1b_1} \dots \Delta_{a_rb_r} = \Delta_{ij}\Delta_{a_1b_1} \dots \Delta_{a_rb_r} + \Delta_{a_1b_1} \dots \Delta_{ib_k} \dots \Delta_{a_rb_r}, \quad (9.3)$$

где  $\Delta_{ib_k}$  стоит на месте члена  $\Delta_{a_kb_k}$ . (Кроме того,  $D_{ij} = \Delta_{ij}$ , но следует помнить, что  $\Delta_{i_1j_1} \dots \Delta_{i_rj_r}$  — это не произведение отдельных множителей, а обозначение для некоторого оператора, и мы не можем заменить сразу несколько членов  $\Delta_{ij}$  на соответствующие  $D_{ij}$ ).

Докажем индукцией по  $m$ , что все соответственные миноры порядка  $m$ , составленные из последних  $m$  столбцов рассматриваемых матриц, равны. При  $m = 1$  это очевидно. Докажем теперь, что если это утверждение верно для миноров порядка  $m$ , то оно верно и для миноров порядка  $m + 1$ . Напомним, что выражение (9.1) соответствует разложению по столбцу. Поэтому, разложив рассматриваемые миноры порядка  $m + 1$  по первому столбцу, мы получим, что их равенство эквивалентно тому, что все миноры порядка  $m + 1$  матрицы

$$\begin{pmatrix} D_{1k} - \Delta_{1k} & D_{1,k+1} & \dots & D_{1n} \\ D_{2k} - \Delta_{2k} & D_{2,k+1} & \dots & D_{2n} \\ \dots & \dots & \dots & \dots \\ D_{kk} - \Delta_{kk} + m & D_{k,k+1} & \dots & D_{kn} \\ \dots & \dots & \dots & \dots \\ D_{nk} - \Delta_{nk} & D_{n,k+1} & \dots & D_{nn} \end{pmatrix},$$

где  $k = n - m$ , равны нулю.

Разложим рассматриваемый минор по первому столбцу. Если этот минор не содержит строку исходной матрицы с номером  $k$ , то применима формула (9.2), поэтому разложение по первому столбцу представляет собой сумму нулевых слагаемых. Рассмотрим теперь случай, когда минор содержит строки исходной матрицы с номерами  $i_1, \dots, i_{m+1}$ , причём  $i_l = k$  для некоторого  $l$ . В этом случае разложение по первому столбцу имеет вид

$$(-1)^l(D_{kk} - \Delta_{kk} + m)A_{i_l} + \sum_{i \neq l} (-1)^r(D_{i,k} - \Delta_{i,k})A_{i_r},$$

где  $A_{i_1}, \dots, A_{i_{m+1}}$  — миноры, которые получаются из рассматриваемого минора вычёркиванием первого столбца и строк  $i_1, \dots, i_{m+1}$  соответственно. Ясно, что  $(D_{kk} - \Delta_{kk})A_{i_l} = 0$ , поскольку в этой ситуации снова применима формула (9.2). А для вычисления  $(D_{i,k} - \Delta_{i,k})A_{i_r}$  нужно применить формулу (9.3). Эта формула показывает, что  $(D_{i,k} - \Delta_{i,k})A_{i_r}$  — это минор, который получается следующим образом. Возьмём матрицу, образованную последними  $m$  столбцами и строками с номерами  $i_1, \dots, i_{m+1}$ ; вычёркнем строку  $i_r$  и запишем строку  $i_r$  на месте строки  $k = i_l$ . По-другому эту операцию можно описать так: мы берём минор  $A_{i_l}$  и переставляем строку  $i_r$  так, чтобы она оказалась между строками  $i_l - 1$  и  $i_l + 1$ ; такой минор отличается от минора  $A_{i_l}$  знаком  $(-1)^{l-r+1}$ . Следовательно,  $(D_{i,k} - \Delta_{i,k})A_{i_r} = (-1)^{l-r+1}A_{i_l}$  и

$$(-1)^l m A_{i_l} + \sum_{i \neq l} (-1)^r (D_{i,k} - \Delta_{i,k}) A_{i_r} = (-1)^l m A_{i_l} + \sum_{i \neq l} (-1)^{l+1} A_{i_l} = 0.$$

### 9.2.3. Собственные значения

Для матрицы  $A$  над телом  $K$  можно определить разные виды собственных значений. Элемент  $\lambda \in K$  называют *правым собственным значением*, если  $Ax = x\lambda$  для некоторого ненулевого вектор-столбца  $x$ ; этот вектор называют *правым собственным*.

Элемент  $\lambda \in K$  называют *левым<sup>1</sup> собственным значением*, если  $Ax = \lambda x$  для некоторого ненулевого вектор-столбца  $x$ ; этот вектор называют *левым собственным*.

<sup>1</sup>Иногда левым собственным значением называют элемент  $\lambda \in K$ , для которого существует ненулевая вектор-строка  $x$ , обладающая тем свойством, что  $xA = \lambda x$ . Свойства так определённых левых собственных значений во многом близки свойствам правых собственных значений, в отличие от тех левых собственных значений, которые рассматриваются в этой книге.

**Теорема 9.2.3.1.** *Элемент  $\lambda \in K$  является левым собственным значением матрицы  $A$  тогда и только тогда, когда матрица  $A - \lambda I$  вырожденная.*

**Доказательство.** Матрица  $A - \lambda I$  вырожденная тогда и только тогда, когда  $(A - \lambda I)v = 0$  для некоторого ненулевого вектора  $v$  (см. с. 475). Равенство  $(A - \lambda I)v = 0$  эквивалентно равенству  $Av = \lambda v$ .  $\square$

Понятие правого собственного значения в некоторых отношениях более естественно, чем понятие левого собственного значения, потому что по нашему соглашению элементы тела стоят справа, а векторы стоят слева. Правые собственные значения возникают, например, в следующей ситуации. Над полем диагонализируемость матрицы  $A$  означает, что существует такая невырожденная матрица  $X$ , что  $X^{-1}AX = D$ , где  $D$  — диагональная матрица, у которой на диагонали стоят собственные значения матрицы  $A$ . Умножив полученное равенство справа на матрицу  $X$ , перейдём к равенству  $AX = XD$ . Расписав это равенство по отдельным столбцам, получим  $Ax_i = x_i\lambda_i$ ; здесь мы получаем именно правые собственные значения.

В некоммутативном случае, если элемент  $\lambda$  является правым собственным значением, то и любой сопряжённый с ним элемент  $a^{-1}\lambda a$  тоже является собственным значением. Действительно, если  $Av = \lambda v$ , то  $Ava = v\lambda a = va(a^{-1}\lambda a)$ .

Для левых собственных значений аналогичное свойство не выполняется: если элемент  $\lambda$  является левым собственным значением, то сопряжённый с ним элемент может и не быть собственным значением. Рассмотрим в качестве примера кватернионную матрицу  $\begin{pmatrix} 0 & i \\ j & 0 \end{pmatrix}$ . Элемент  $\lambda$  является левым собственным значением, если существуют кватернионы  $p$  и  $q$  (не равные нулю одновременно), для которых выполняются равенства  $iq = \lambda p$  и  $jp = \lambda q$ . Из этой системы уравнений получаем  $i\lambda^{-1}j = \lambda$ , т.е.  $-i\lambda^{-1} = \lambda j$ . В частности,  $\|\lambda\| = 1$ , поэтому для  $\lambda = a + bi + cj + dk$  полученное уравнение запишется так:  $-ai - b + ck - dj = aj + bk - c - di$ . Следовательно,  $a = d = 0$  и  $b = c$ . Вспоминая, что  $\|\lambda\| = 1$ , получаем  $\lambda = \pm \frac{i+j}{\sqrt{2}}$ . Но согласно следствию теоремы 8.1.4.1 любые два чисто мнимых кватерниона одной длины сопряжены, а мы здесь получаем только два кватерниона из класса сопряжённости.

В коммутативном случае матрицы  $A$  и  $X^{-1}AX$  имеют одинаковые собственные значения. В некоммутативном случае для правых собственных значений это свойство сохраняется. Действительно, если  $Av = \lambda v$ , то  $X^{-1}AX(X^{-1}v) = (X^{-1}v)\lambda$ .

**Теорема 9.2.3.2.** *Собственные векторы  $v_1, \dots, v_r$ , соответствующие попарно не сопряжённым правым собственным значениям  $\lambda_1, \dots, \lambda_r$ , линейно*

независимы.

**Доказательство.** Предположим, что векторы  $v_1, \dots, v_r$  линейно зависимы. Выбрав среди этих векторов наименьшее множество линейно зависимых векторов и изменив при необходимости их нумерацию, можно считать, что

$$v_1 = v_2 a_2 + \dots + v_r a_r. \quad (9.1)$$

По определению собственного вектора  $v_i \neq 0$ . Из минимальности выбранного множества векторов следует, что  $a_i \neq 0$ . Домножим равенство (9.1) справа на собственное значение  $\lambda_1$ . В результате получим:

$$v_1 \lambda_1 = v_2 a_2 \lambda_1 + \dots + v_r a_r \lambda_1. \quad (9.2)$$

С другой стороны,

$$v_1 \lambda_1 = A v_1 = A v_2 a_2 + \dots + A v_r a_r = v_2 \lambda_2 a_2 + \dots + v_r \lambda_r a_r. \quad (9.3)$$

Сравнивая равенства (9.2) и (9.3), получаем

$$v_2(a_2 \lambda_1 - \lambda_2 a_2) + \dots + v_r(a_r \lambda_1 - \lambda_r a_r)$$

Векторы  $v_2, \dots, v_r$  линейно независимы, поэтому  $\lambda_1 = a_2^{-1} \lambda_2 a_2$ . Приходим к противоречию.  $\square$

### Задачи

**2.1.** Докажите, что левыми собственными значениями треугольной матрицы  $A$  являются её диагональные элементы и только они.

**2.2.** Докажите что кватернионная матрица  $A = \begin{pmatrix} i & j \\ -j & i \end{pmatrix}$  нильпотентна, но у неё есть ненулевое левое собственное значение.

б) Докажите что кватернионная матрица  $B = \begin{pmatrix} j & i \\ i & -j \end{pmatrix}$  не нильпотентна, но её единственное левое значение равно 0.

## 9.3. Кватернионные пространства

Кватернионное пространство  $\mathbb{H}^n$  — это множество наборов из  $n$  кватернионов  $x = (x_1, \dots, x_n)$ , причём  $xq = (x_1 q, \dots, x_n q)$ . Если выбрать базисные базисные векторы  $e_m = (0, \dots, 0, 1, 0, \dots, 0)$ , где 1 стоит на  $m$ -м месте, то векторы из  $\mathbb{H}^n$  можно представить в виде линейных комбинаций  $e_1 x_1 + \dots + e_n x_n$ .

### 9.3.1. Кватернионная структура

По аналогии с комплексной структурой на  $\mathbb{R}^{2n}$  можно определить кватернионную структуру на  $\mathbb{C}^{2n}$ . Мы будем исходить из кватернионного пространства  $\mathbb{H}^n$ . *Кватернионная структура* — это отображение  $R_j: \mathbb{H}^n \rightarrow \mathbb{H}^n$ , которое заключается в умножении справа на  $j$ . Обратите внимание, что  $R_j(vq) = vqj$ , а  $R_j(v)q = vjq$ , поэтому отображение  $R_j$  не является линейным над  $\mathbb{H}$  отображением кватернионного пространства.

Каждый кватернион  $a + bi + cj + dk$  можно представить в виде  $z + wj$ , где  $z = a + bi$  и  $w = c + di$  — комплексные числа. Поэтому каждый вектор кватернионного пространства  $\mathbb{H}^n$  можно представить в виде  $\sum e_m z_m + \sum e_m w_m j = \sum e_m z_m + \sum e_m j \bar{w}_m$ . Последнее выражение побуждает изменить обозначения и считать, что вектор  $v$  кватернионного пространства записан в виде  $v = \sum e_m z_m + \sum e_m j w_m$ . Тогда  $R_j(v) = \sum e_m z_m j + \sum e_m j w_m j = \sum e_m j \bar{z}_m - \sum e_m \bar{w}_m$ .

Выберем в  $\mathbb{C}^{2n}$  базис  $e_1, \dots, e_n, e_1 j, \dots, e_n j$ . При естественном отождествлении пространства  $\mathbb{C}^{2n}$  с кватернионным пространством  $\mathbb{H}^n$  с базисом  $e_1, \dots, e_n$  кватернионная структура — это отображение

$$R_j(z_1, \dots, z_n, w_1, \dots, w_n) = (-\bar{w}_1, \dots, -\bar{w}_n, \bar{z}_1, \dots, \bar{z}_n).$$

Отметим ещё раз, что это отображение не является линейным над  $\mathbb{C}$ . Но линейное над  $\mathbb{C}$  отображение  $A: \mathbb{C}^{2n} \rightarrow \mathbb{C}^{2n}$  является линейным над  $\mathbb{H}$  отображением пространства  $\mathbb{H}^n$  тогда и только тогда, когда оно коммутирует с кватернионной структурой, т.е.  $AR_j = R_j A$ .

Если мы запишем матрицу линейного над  $\mathbb{H}$  отображения пространства  $\mathbb{H}^n$  в виде  $A + jB$ , где  $A$  и  $B$  — комплексные матрицы, то матрица того же самого отображения, рассматриваемого как отображение пространства  $\mathbb{C}^{2n}$ , имеет вид  $\begin{pmatrix} A & -\bar{B} \\ B & A \end{pmatrix}$ . Действительно,

$$\begin{aligned} (A + jB)e_p &= \sum e_q a_{qp} + \sum e_q j b_{qp}, \\ (A + jB)e_p j &= \sum e_q a_{qp} j + \sum e_q j b_{qp} j = \sum e_q j \bar{a}_{qp} - \sum e_q \bar{b}_{qp}. \end{aligned}$$

Это выражение можно получить и по-другому. В пространстве  $\mathbb{C}^{2n}$  отображение  $R_j$  можно записать в виде  $R_j(v) = \bar{J}v$ , где  $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$ . Поэтому если мы запишем соотношение  $R_j X = X R_j$  для комплексной матрицы  $X = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$ , то получим, что  $\bar{J}Xv = X\bar{J}v$ , для любого  $v \in \mathbb{C}^{2n}$ . Учитывая, что  $X\bar{J}v = \bar{X}Jv$ , получаем, что комплексная матрица  $X$  порядка  $2n$  соответствует кватернионно линейному отображению тогда и только тогда, когда  $JX = \bar{X}J$ . Это соотношение показывает, что  $C = -\bar{B}$  и  $D = \bar{A}$ .

### 9.3.2. Гиперэрмитовы матрицы

Квадратную кватернионную матрицу  $A = \|a_{ij}\|_1^n$  называют *гиперэрмитовой*, если  $a_{ij} = \overline{a_{ji}}$ , т.е.  $A = A^*$ , где  $A^*$  — кватернионно сопряжённая матрица.

Каждой гиперэрмитовой матрице  $A = \|a_{ij}\|_1^n$  можно сопоставить *гиперэрмитову полулинейную форму*

$$A(x, y) = \sum_{i,j=1}^n \bar{x}_i a_{ij} y_j$$

на пространстве  $\mathbb{H}^n$ . Гиперэрмитова форма  $A(x, y)$  обладает следующими свойствами:

- 1) она аддитивна как по аргументу  $x$ , так и по аргументу  $y$ ;
- 2)  $A(x, y\lambda) = A(x, y)\lambda$ , т.е. форма  $A$  линейна по второму аргументу;
- 3)  $A(x, y) = \overline{A(y, x)}$ .

Если  $A$  — матрица гиперэрмитовой формы, записанной относительно некоторого базиса,  $C$  — матрица перехода от этого базиса к другому базису, а  $A'$  — матрица той же самой гиперэрмитовой формы относительно нового базиса, то  $A' = C^* A C$ .

Легко проверить, что для любой гиперэрмитовой матрицы  $A$  и любой матрицы  $C$  того же порядка матрица  $C^* A C$  гиперэрмитова. В частности, для любой квадратной кватернионной матрицы  $C$  матрица  $C^* C$  гиперэрмитова.

Гиперэрмитову полулинейную форму  $A(x, y)$  называют *положительно определённой*, если  $A(x, x) > 0$  для любого ненулевого вектора  $x$ .

Кватернионное векторное пространство с фиксированной положительно определённой гиперэрмитовой формой называют *гиперэрмитовым пространством*.

Для любого кватернионно линейного оператора  $X: V \rightarrow V$  можно рассмотреть *гиперэрмитово сопряжённый* оператор  $X^*: V \rightarrow V$ , который задаётся тем свойством, что  $(Xv, w) = (v, X^*w)$  для всех векторов  $v$  и  $w$ .

Если в гиперэрмитовом пространстве  $V$  фиксирован ортонормированный базис, то оператор  $X$  самосопряжённый тогда и только тогда, когда его матрица относительно этого базиса гиперэрмитова.

**Теорема 9.3.2.1.** *Матрица овеществления гиперэрмитова оператора симметрическая.*

**Доказательство.** Запишем матрицу гиперэрмитова оператора в виде  $A + jB$ , где  $A$  и  $B$  — комплексные матрицы. Гиперэрмитовость означает, что  $\overline{A + jB} = A^T + jB^T$ . Ясно, что  $\overline{A + jB} = \bar{A} + \bar{j}\bar{B} = \bar{A} - j\bar{B}$ . Таким образом,  $A^T = \bar{A}$  и  $B^T = -\bar{B}$ . Следовательно,

$$\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix} = \begin{pmatrix} A^T & B^T \\ -\bar{B}^T & \bar{A}^T \end{pmatrix} = \begin{pmatrix} \bar{A} & -B \\ \bar{B} & A \end{pmatrix} = \overline{\begin{pmatrix} A & -B \\ B & \bar{A} \end{pmatrix}}.$$

Итак, если мы рассматриваем гиперэрмитов оператор в пространстве  $\mathbb{H}^n$  как оператор в пространстве  $\mathbb{C}^{2n}$ , то этот оператор является эрмитовым. Поэтому овеществление этого оператора является симметрическим (задача 7.1).  $\square$

В случае комплексных матриц собственные значения любой эрмитовой матрицы вещественные. Для левых собственных значений гиперэрмитовых матриц аналогичное утверждение неверно. Например,

$$\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \begin{pmatrix} 1 \\ k \end{pmatrix} = (1 - j) \begin{pmatrix} 1 \\ k \end{pmatrix}.$$

Тем не менее, правые собственные значения гиперэрмитовой матрицы всегда вещественны.

**Теорема 9.3.2.2.** *Все правые собственные значения гиперэрмитовой матрицы  $A$  вещественны.*

**Доказательство.** Пусть  $Av = v\lambda$  для некоторого ненулевого вектора  $v$ . Тогда, с одной стороны,  $(v, Av) = (v, v\lambda) = (v, v)\lambda$ , а с другой стороны,  $(v, Av) = \overline{(Av, v)} = \overline{(v, A^*v)} = \overline{(v, Av)} = \overline{(v, v\lambda)} = (v, v)\bar{\lambda}$ , поскольку  $(v, v)$  — вещественное число. Кроме того,  $(v, v) \neq 0$ , поэтому  $\bar{\lambda} = \lambda$ .  $\square$

**Теорема 9.3.2.3.** *Правые собственные векторы гиперэрмитовой матрицы  $A$ , соответствующие разным правым собственным значениям, ортогональны.*

**Доказательство.** Пусть  $Av_1 = v_1\lambda_1$  и  $Av_2 = v_2\lambda_2$  для некоторых ненулевых векторов  $v_1$  и  $v_2$ . Согласно теореме 9.3.2.2 числа  $\lambda_1$  и  $\lambda_2$  вещественные. Поэтому, с одной стороны,  $(Av_1, v_2) = (v_1, A^*v_2) = (v_1, Av_2) = \lambda_2(v_1, v_2)$ , а с другой стороны,  $(Av_1, v_2) = (v_2, Av_1) = \lambda_1(v_2, v_1) = \lambda_1(v_1, v_2)$ . Таким образом, если  $\lambda_1 \neq \lambda_2$ , то  $(v_1, v_2) = 0$ .  $\square$

**Теорема 9.3.2.4.** *Для любого самосопряжённого оператора  $A$  в гиперэрмитовом пространстве можно выбрать ортонормированный базис так, что матрица оператора относительно этого базиса будет диагональной.*

**Доказательство.** Если подпространство  $W$  инвариантно относительно самосопряжённого оператора  $A$ , то подпространство  $W^\perp$  тоже инвариантно относительно  $A$ . Действительно, пусть  $w \in W$  и  $v \in W^\perp$ . Тогда  $(Av, w) = (v, A^*w) = (v, Aw) = 0$ .

Ясно также, что у оператора  $A$  есть хотя бы один собственный вектор, поскольку его овереществление является симметрическим оператором (теорема 9.3.2.1). Дальнейшее доказательство проводится по индукции, точно так же, как в комплексном случае (теорема 3.1.3.2).  $\square$

### 9.3.3. Унитарные матрицы

Кватернионный вектор-столбец  $u$  называют *единичным*, если  $u^*u = 1$ .

Кватернионные вектор-столбцы  $u$  и  $v$  называют *ортгоналичными*, если  $u^*v = 0$ . Это равенство эквивалентно равенству  $v^*u = 0$ , поскольку  $\bar{u}_1 v_1 + \dots + \bar{u}_n v_n = \bar{v}_1 u_1 + \dots + \bar{v}_n u_n$ .

Мы будем говорить, что кватернионные вектор-столбцы  $u_1, u_2, \dots, u_n$  образуют ортонормированную систему, если все эти векторы единичные и  $u_i^*u_j = 0$  при  $i \neq j$ .

Кватернионную квадратную матрицу  $U$  называют *унитарной*, если  $U^*U = I$ . Это означает, что столбцы этой матрицы образуют ортонормированную систему. Из равенства  $U^*U = I$  следует равенство  $U^* = U^{-1}$ , а из него следует равенство  $UU^* = I$  (здесь мы пользуемся тем, что левая обратная матрица совпадает с правой обратной). Таким образом, ортонормированность столбцов матрицы эквивалентна ортонормированности её строк.

**Теорема 9.3.3.1.** Пусть  $u_1$  — единичный кватернионный вектор-столбец длины  $n$ . Тогда можно выбрать  $n - 1$  единичных кватернионных вектор-столбцов  $u_2, \dots, u_n$  длины  $n$  так, что векторы  $u_1, u_2, \dots, u_n$  образуют ортонормированную систему.

**Доказательство.** Прежде всего заметим, что если  $A$  — кватернионная<sup>2</sup> матрица размера  $m \times n$ , где  $m < n$ , то уравнение  $Ax = 0$  имеет ненулевое решение. Действительно, пусть  $A = A_1 + jA_2$  и  $x = x_1 + jx_2$ , где  $A_1$  и  $A_2$  — комплексные матрицы, а  $x_1$  и  $x_2$  — комплексные векторы. Тогда уравнение  $Ax = 0$  можно записать в виде  $\begin{pmatrix} A_1 & -A_2 \\ A_2 & \bar{A}_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$ . Это уравнение имеет ненулевое комплексное решение, поскольку  $2m < 2n$ .

Умножив ненулевое решение уравнения  $Ax = 0$  на подходящее действительное число, можно считать, что это решение является единичным вектором. В качестве  $u_2$  выберем единичный вектор, являющийся

<sup>2</sup>В действительности это верно для матриц над любым телом: см. задачу 1.6.



решением уравнения  $u_1^*x = 0$ . В качестве  $u_3$  выберем единичный вектор, являющийся решением уравнения  $U_2^*x = 0$ , где  $U_2$  — матрица составленная из столбцов  $u_1$  и  $u_2$ . Остальные векторы выбираются аналогично.  $\square$

Теорему 9.3.3.1 можно переформулировать так: для любого единичного кватернионного вектора существует унитарная кватернионная матрица, первым столбцом которой является этот вектор.

Квадратную кватернионную матрицу  $A$  называют *нормальной*, если  $AA^* = A^*A$ . Ясно, что любая унитарная кватернионная матрица нормальна. Любая гиперэрмитова матрица тоже нормальна.

### 9.3.4. Правые собственные значения

Для гиперэрмитовых матриц существование собственных значений (левых и правых) доказывается легко. Действительно, овеществление гиперэрмитовой матрицы является симметрической матрицей (теорема 9.3.2.1), поэтому у неё есть вещественные собственные значения (каждое вещественное собственное значение является одновременно правым и левым). Описание правых собственных значений кватернионной матрицы несложно получить, рассмотрев её комплексификацию.

Напомним, что если какой-то элемент тела является правым собственным значением матрицы, то и все элементы, сопряжённые с ним, тоже являются правыми собственными значениями. Класс кватернионов, сопряжённых с кватернионом  $a + q$ , где  $a$  — действительное число, а  $q$  — чисто мнимый кватернион, состоит из всех кватернионов вида  $a + r$ , где  $r$  — такой чисто мнимый кватернион, что  $\|r\| = \|q\|$  (следствие теоремы 8.1.4.1). Таким образом, если кватернион вещественный, то его класс сопряжённости состоит из одного элемента, а если кватернион не вещественный то в его классе эквивалентности есть ровно два комплексных числа, причём эти числа комплексно сопряжённые. Поэтому, чтобы получить полное описание правых собственных значений кватернионной матрицы, достаточно найти её комплексные правые собственные значения.

**Теорема 9.3.4.1.** *Любая квадратная кватернионная матрица имеет правое собственное значение.*

**Доказательство.** Запишем кватернионную матрицу в виде  $A + jB$ , где  $A$  и  $B$  — комплексные матрицы, а кватернионный вектор запишем в виде  $x + jy$ , где  $x$  и  $y$  — комплексные векторы. Ясно, что  $(A + jB)(x + jy) =$

$(Ax - \bar{B}y) + j(Bx + \bar{A}y)$ . Кроме того, для любого комплексного числа  $\lambda$  получаем  $(x + jy)\lambda = \lambda x + j\lambda y$ . Поэтому равенство  $(A + jB)(x + jy) = (x + jy)\lambda$  эквивалентно равенству

$$\begin{pmatrix} A & -\bar{B} \\ B & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}.$$

Таким образом, комплексные правые собственные значения кватернионной матрицы  $A + jB$  — это в точности собственные значения комплексной матрицы  $\begin{pmatrix} A & -\bar{B} \\ B & A \end{pmatrix}$ .  $\square$

Собственные значения комплексной матрицы  $X$  порядка  $2n$ , соответствующей кватернионной матрице порядка  $n$ , разбиваются на пары комплексно сопряжённых чисел. Действительно,  $\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ , поэтому

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}^{-1} \begin{pmatrix} A & -\bar{B} \\ B & A \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \begin{pmatrix} \bar{A} & -B \\ \bar{B} & A \end{pmatrix} = \overline{\begin{pmatrix} A & -\bar{B} \\ B & A \end{pmatrix}}.$$

Таким образом, матрица  $X$  подобна матрице  $\bar{X}$ . Поэтому  $|X - \bar{\lambda}I| = |\bar{X} - \bar{\lambda}I| = |\overline{X - \lambda I}|$ , а значит,  $\bar{\lambda}$  является собственным значением матрицы  $X$  тогда и только тогда, когда  $\lambda$  является собственным значением  $X$ .

Вещественные собственные значения матрицы  $X$  тоже встречаются парами. Действительно, рассмотрим вместо матрицы  $X$  матрицу  $X + itI$ . Эта матрица тоже подобна комплексно сопряжённой с ней матрице, поэтому при малом  $t$  в малой окрестности вещественного собственного значения матрицы  $X$  мы получаем пару комплексно сопряжённых собственных значений матрицы  $X + itI$ . При  $t \rightarrow 0$  эта пара комплексно сопряжённых собственных значений превращается в пару вещественных собственных значений матрицы  $X$ .

Каждая пара комплексно сопряжённых чисел соответствует одному классу сопряжённости кватернионов, поэтому для кватернионной матрицы порядка  $n$  мы получаем  $n$  классов сопряжённости правых собственных значений.

### 9.3.5. Разложение Шура

Для кватернионных матриц тоже есть аналог разложения Шура (для комплексных матриц — это теорема 3.7.1.1).

**Теорема 9.3.5.1.** Любую квадратную кватернионную матрицу  $A$  можно представить в виде  $A = UTU^*$ , где  $U$  — унитарная,  $T$  — треугольная матрица, причём матрица  $A$  нормальна тогда и только тогда, когда  $T$  — диагональная матрица.

**Доказательство.** Применим индукцию по порядку матрицы  $A$ . Согласно теореме 9.3.4.1 матрица  $A$  имеет правое собственное значение, т.е. можно выбрать кватернион  $\lambda$  и ненулевой кватернионный вектор-столбец  $x$  так, что  $Ax = \lambda x$ . Можно считать, что кватернионный вектор-столбец  $x$  единичный. Согласно теореме 9.3.3.1 существует унитарная кватернионная матрица  $W$ , первым столбцом которой служит вектор  $x$ . Тогда

$$W^*AW = \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

По предположению индукции существует такая унитарная кватернионная матрица  $V$ , что матрица  $V^*A_1V$  треугольная. Тогда матрица  $U = W \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$  искомая.

Доказательство того, что матрица  $A$  нормальная тогда и только тогда, когда матрица  $T$  треугольная, ничем не отличается от доказательства в комплексном случае.  $\square$

**Теорема 9.3.5.2.** Пусть  $A$  — верхняя треугольная кватернионная матрица. Тогда любой диагональный элемент этой матрицы является её правым собственным значением.

**Доказательство.** Пусть  $A$  — верхняя треугольная кватернионная матрица порядка  $n$  с диагональными элементами  $\lambda_1, \dots, \lambda_n$ . Применим индукцию по  $n$ . При  $n = 1$  доказывать нечего. Если  $n > 1$ , то матрицу  $A$  можно записать в виде  $A = \begin{pmatrix} \lambda_1 & a \\ 0 & A_1 \end{pmatrix}$ , где  $A_1$  — матрица порядка  $n - 1$  и  $a = (a_2, \dots, a_n)$ . Ясно, что  $\lambda_1$  — правое собственное значение матрицы  $A$ , поскольку  $A(1, 0, \dots, 0)^T = (1, 0, \dots, 0)^T \lambda_1$ . По предположению индукции матрица  $A_1$  имеет собственные значения  $\lambda_2, \dots, \lambda_n$ . Поэтому достаточно доказать, что если  $\lambda$  — один из кватернионов  $\lambda_2, \dots, \lambda_n$ , причём он не сопряжён с  $\lambda_1$ , то он является собственным значением матрицы  $A$ .

Предположим, что  $A_1 y = y \lambda$ , где  $y = (y_2, \dots, y_n)^T$  — ненулевой кватернионный вектор. Пусть  $x$  — произвольный кватернион. Тогда

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda & a \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_1 x + ay \\ A_1 y \end{pmatrix} = \begin{pmatrix} \lambda_1 x + ay \\ y \lambda \end{pmatrix}.$$

Остаётся подобрать кватернион  $x$  так, что  $\lambda_1 x + ay = x \lambda$ , где  $ay = a_2 y_2 + \dots + a_n y_n = b$  — некоторый кватернион.

Отображение  $x \mapsto \lambda x - x \lambda_1$  является линейным отображением пространства кватернионов, которое мы рассматриваем как линейное пространство над  $\mathbb{R}$ . Докажем, что это отображение невырожденное. Предположим, что  $\lambda x - x \lambda_1 = 0$  для некоторого  $x \neq 0$ . Тогда  $\lambda = x \lambda_1 x^{-1}$ .

Но этого не может быть, поскольку кватернионы  $\lambda$  и  $\lambda_1$  не сопряжены. Таким образом, рассматриваемое отображение невырожденное, поэтому для любого кватерниона  $b$  можно выбрать кватернион  $x$  так, что  $\lambda x - x \lambda_1 = b$ .  $\square$

### 9.3.6. Определитель Штуди

Напомним, что если мы запишем матрицу линейного над  $\mathbb{H}$  отображения пространства  $\mathbb{H}^n$  в виде  $A + jB$ , где  $A$  и  $B$  — комплексные матрицы, то матрица того же самого отображения, рассматриваемого как отображение пространства  $\mathbb{C}^{2n}$ , имеет вид  $\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix}$ .

*Определитель Штуди* кватернионной матрицы  $A + jB$  — это обычный определитель соответствующей ей комплексной матрицы  $\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix}$ . Определитель Штуди впервые появился в работе [St2]. Определитель Штуди мы будем обозначать  $\text{Sdet}$ .

Обратите внимание, что определитель Штуди — квадратичная (а не линейная) функция от элементов кватернионной матрицы (точнее говоря, от их комплексных составляющих).

Напомним, что комплексная матрица  $X = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$  соответствует кватернионно линейному отображению тогда и только тогда, когда  $JX = \bar{X}J$ , где  $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$  (см. §9.3.1). В таком случае  $\det(X) = \det(\bar{X})$ , поскольку  $\det(J) \neq 0$ , а значит, число  $\text{Sdet}(X)$  вещественное. Более того, число  $\text{Sdet}(X)$  неотрицательное, т.е. определитель Штуди принимает только неотрицательные значения. Действительно, согласно теореме 9.3.9.1 пространство невырожденных кватернионных матриц линейно связно, поэтому его образ при отображении  $\text{Sdet}$  линейно связан. Ясно, что образ единичной матрицы — это 1, поэтому образ содержится в множестве положительных чисел.

### 9.3.7. Определитель Дьёдонне

Напомним, что определитель Дьёдонне для невырожденных матриц с элементами из тела  $K$  принимает значения в  $K^{\text{ab}} = K^\times / K^{\times'}$ , где  $K^{\times'}$  — коммутант тела  $K$ . Поэтому прежде всего надо вычислить коммутант тела кватернионов.

*Коммутатором* двух ненулевых кватернионов  $p$  и  $q$  называют кватернион  $pqr^{-1}q^{-1}$ .

**Теорема 9.3.7.1.** *Кватернион  $z$  является коммутатором двух кватернионов тогда и только тогда, когда  $\|z\| = 1$ .*

**Доказательство.** [As] Ясно, что  $\|pqp^{-1}q^{-1}\| = 1$ , поэтому нужно лишь проверить, что любой кватернион  $z$  единичной длины можно представить в виде  $z = pqp^{-1}q^{-1}$ . Пусть  $z = \cos \varphi + r \sin \varphi$ , где  $r$  — чисто мнимый кватернион единичной длины. Согласно следствию теоремы 8.1.4.1 кватернион  $r$  можно представить в виде  $r = xix^{-1}$ , где  $x$  — некоторый кватернион. Таким образом,  $z = xux^{-1}$ , где  $u = \cos \varphi + i \sin \varphi$ . Пусть  $w = \cos(\varphi/2) + i \sin(\varphi/2)$ . Тогда  $u = w^2$ . Если чисто мнимые кватернионы  $a$  и  $b$  единичной длины лежат в плоскости, ортогональной  $i$ , то  $ab = -(a, b) + a \times b = -\cos \psi + i \sin \psi$ , где  $\psi$  — угол между векторами  $a$  и  $b$ . Поэтому  $w$  можно представить в виде  $w = ab$ , где  $a$  и  $b$  — чисто мнимые кватернионы единичной длины. Для таких кватернионов  $a^{-1} = -a$  и  $b^{-1} = -b$ , поэтому

$$\begin{aligned} z &= xux^{-1} = xw^2x^{-1} = xababx^{-1} = xab(-a)(-b)x^{-1} = xaba^{-1}b^{-1}x^{-1} = \\ &= (xax^{-1})(xbx^{-1})(xax^{-1})^{-1}(xbx^{-1})^{-1}, \end{aligned}$$

и мы получаем требуемое представление.  $\square$

Таким образом, для тела кватернионов  $\mathbb{H}$  мы получаем, что  $\mathbb{H}^{\text{ab}} = \mathbb{H}^{\times}/\mathbb{H}^{\times'}$  — группа положительных чисел относительно умножения, т.е. определитель Дьёдонне принимает положительные значения.

**Теорема 9.3.7.2.** *Определитель Дьёдонне кватернионной матрицы  $(q)$  порядка 1 равен  $\|q\|$ .*

**Доказательство.** Определитель Дьёдонне кватернионной матрицы  $(q)$  порядка 1 равен  $\bar{q}$ , где  $\bar{q}$  — класс кватерниона  $q \in \mathbb{H}$  в  $\mathbb{H}^{\text{ab}} = \mathbb{H}^{\times}/\mathbb{H}^{\times'}$ . Согласно теореме 9.3.7.1 кватернион  $q^{-1}\|q\|$  лежит в  $\mathbb{H}^{\times'}$ , поэтому кватернион  $q$  лежит в том же классе, что и  $q(q^{-1}\|q\|) = \|q\|$ . Кроме того, согласно той же теореме кватернионы с разными нормами лежат в разных классах.  $\square$

**Теорема 9.3.7.3.** *Определитель Дьёдонне  $\det A$  невырожденной кватернионной матрицы и её определитель Штуди  $\text{Sdet } A$  связаны соотношением  $\text{Sdet } A = (\det A)^2$ .*

**Доказательство.** При замене матрицы  $A$  на матрицу  $U^{-1}AU$  ни определитель Дьёдонне, ни определитель Штуди не изменяются. Поэтому согласно теореме 9.3.5.1 достаточно рассмотреть случай, когда матрица  $A$  треугольная. Определитель Дьёдонне треугольной матрицы равен модулю произведения диагональных элементов. Поэтому остаётся проверить, что определитель Штуди треугольной матрицы равен произведению квадратов модулей её диагональных элементов. Рассмотрим для

простоты случай матрицы порядка 3 (общий случай разбирается аналогично). В этом случае нужно доказать, что определитель матрицы вида

$$\begin{pmatrix} a_1 & * & * & -\bar{b}_1 & * & * \\ 0 & a_2 & * & 0 & -\bar{b}_2 & * \\ 0 & 0 & a_3 & 0 & 0 & -\bar{b}_3 \\ b_1 & * & * & \bar{a}_1 & * & * \\ 0 & b_2 & * & 0 & \bar{a}_2 & * \\ 0 & 0 & b_3 & 0 & 0 & \bar{a}_3 \end{pmatrix}$$

равен  $\prod_{i=1}^3 (|a_i|^2 + |b_i|^2)$ . Будем приводить эту матрицу к треугольному виду. Вычтем из третьей строки первую строку, умноженную на  $b_1/a_1$ . При этом элемент  $\bar{a}_1$  заменится на  $\bar{a}_1 + \frac{b_1 \bar{b}_1}{a_1}$ . Первый элемент третьей строки стал равен нулю. Сделаем теперь равным нулю второй элемент этой строки. Для этого вычтем из неё вторую строку, умноженную на соответствующий коэффициент. Заметим, что при этом элемент  $\bar{a}_1 + \frac{b_1 \bar{b}_1}{a_1}$  не изменяется, поскольку из него вычитается 0. Затем сделаем равным нулю третий элемент третьей строки. После этого перейдём к четвёртой строке и т.д. В результате получим, что определитель равен

$$\prod_{i=1}^3 a_i \left( \bar{a}_i + \frac{b_i \bar{b}_i}{a_i} \right) = \prod_{i=1}^3 (|a_i|^2 + |b_i|^2).$$

□

### 9.3.8. Определитель Мура

*Определитель Мура*  $\text{Mdet } A$  гиперэрмитовой кватернионной матрицы  $A$  был введён в работе [Мо4]. Пусть  $\sigma$  — перестановка  $n$  чисел. Представим её в виде произведения непересекающихся циклов и запишем каждый цикл так, чтобы он начинался с наименьшего числа. Наконец, упорядочим циклы так, чтобы их первые числа были расположены в порядке убывания:

$$\sigma = (n_{11} \dots n_{1l_1})(n_{21} \dots n_{2l_2}) \dots (n_{r1} \dots n_{rl_r}),$$

где для каждого  $i$  при всех  $j > 1$  выполняется неравенство  $n_{i1} < n_{ij}$  и, кроме того,  $n_{11} > n_{21} > \dots > n_{r1}$ . Тогда для гиперэрмитовой кватернионной матрицы  $A$  мы полагаем

$$\text{Mdet } A = \sum_{\sigma \in S_n} (-1)^\sigma a_{n_{11}n_{12}} \dots a_{n_{1l_1}n_{11}} a_{n_{21}n_{22}} \dots a_{n_{rl_r}n_{r1}}.$$

Приведём без доказательства некоторые свойства определителя Мура.

- 1) Для гиперэрмитовой кватернионной матрицы  $A$  число  $\text{Mdet } A$  вещественное.
- 2) На пространстве всех гиперэрмитовых матриц порядка  $n$  существует единственный однородный многочлен  $P$  степени  $n$ , который обладает следующими двумя свойствами: на единичной матрице равен 1 и для любой гиперэрмитовой матрицы  $A$  имеет место равенство  $P^4(A) = \det A_{\mathbb{R}}$ , где  $A_{\mathbb{R}}$  — овеществление кватернионной матрицы  $A$  (вещественная матрица порядка  $4n$ ). Этот многочлен и есть определитель Мура.
- 3) Для гиперэрмитовой кватернионной матрицы  $A$  имеет место равенство  $|\text{Mdet } A| = \det A$ , где  $\det$  — определитель Дьёдонне.
- 4) Для любой невырожденной кватернионной матрицы  $A$  имеет место равенство  $\text{Sdet } A = \text{Mdet } AA^*$ .
- 5) Определитель Мура любой комплексной эрмитовой матрицы, рассматриваемой как гиперэрмитова матрица, равен её обычному определителю.
- 6) Для любой гиперэрмитовой матрицы  $A$  и любой кватернионной матрицы  $C$  того же порядка имеет место равенство

$$\text{Mdet}(C^*AC) = \text{Mdet } A \cdot \text{Mdet}(C^*C).$$

### 9.3.9. Топология пространства $\text{GL}(n, \mathbb{H})$

**Теорема 9.3.9.1.** *Пространство невырожденных кватернионных матриц линейно связно, т.е. любые две матрицы  $A, B \in \text{GL}(n, \mathbb{H})$  можно соединить непрерывным путём, состоящим из невырожденных кватернионных матриц.*

**Доказательство.** Согласно теореме 9.1.2.1 любую невырожденную квадратную матрицу  $A$  можно представить в виде  $A = BD(\mu)$ , где  $B$  — произведение элементарных матриц  $B_{ij}(\lambda)$ , а  $D = \text{diag}(1, \dots, 1, \mu)$ . Путь  $A(t) = \prod B_{ij}(t\lambda)D(\mu)$  соединяет матрицу  $A = A(1)$  с матрицей  $A(0) = \text{diag}(1, \dots, 1, \mu)$ . Остаётся заметить, что пространство ненулевых кватернионов, которому принадлежит кватернион  $\mu$ , линейно связно.  $\square$

Для доказательства существования левых собственных значений нам понадобятся некоторые свойства гомотопических групп пространства  $GL(n, \mathbb{H})$  (теоремы 9.3.9.2 и 9.3.9.3). Необходимые для их доказательств сведения из топологии можно найти, например, в [П3] и в [П4].

Пространство  $GL(n, \mathbb{H})$  некомпактно. Вместо него удобнее рассмотреть гомотопически эквивалентное ему компактное подпространство  $Sp(n) \subset GL(n, \mathbb{H})$ , которое определяется следующим образом. Пусть  $(x, y) = \sum x_i \bar{y}_i$  — гиперэрмитово скалярное произведение. Пространство  $Sp(n)$  состоит из матриц  $U \in GL(n, \mathbb{H})$ , для которых  $(Ux, Uy) = (x, y)$  для всех  $x, y \in \mathbb{H}^n$ . Пространства  $Sp(n)$  и  $GL(n, \mathbb{H})$  гомотопически эквивалентны. Действительно,  $GL(n, \mathbb{H})$  можно отождествить с пространством всех базисов в  $\mathbb{H}^n$ , а  $Sp(n)$  можно отождествить с пространством всех ортонормированных базисов в  $\mathbb{H}^n$ . Используя ортогонализацию Грама-Шмидта, можно доказать, что тождественное отображение  $GL(n, \mathbb{H}) \rightarrow GL(n, \mathbb{H})$  гомотопно отображению  $GL(n, \mathbb{H}) \rightarrow Sp(n)$ , ограничение которого на  $Sp(n)$  тождественно. Из этого следует, что пространства  $Sp(n)$  и  $GL(n, \mathbb{H})$  гомотопически эквивалентны.

Для любого натурального  $k \leq n$  можно рассмотреть *многообразие Штифеля*  $V_k(\mathbb{H}^n)$ , точками которого являются ортонормированные наборы из  $k$  векторов в  $\mathbb{H}^n$ . Существует локально тривиальное расслоение  $SP(n) \rightarrow V_k(\mathbb{H}^n)$  со слоем  $SP(n-k)$ . Это расслоение устроено следующим образом. Возьмём ортонормированный набор из  $n$  векторов и отбросим в нём все векторы, кроме первых  $k$  векторов. Совпадающие наборы  $k$  векторов получаются в том случае, когда в качестве последних  $n-k$  векторов берётся ортонормированный базис ортогонального дополнения к подпространству, натянутому на данные  $k$  векторов. Все такие ортонормированные базисы вместе образуют пространство  $SP(n-k)$ .

Непосредственно из определения видно, что  $V_n(\mathbb{H}^n) = Sp(n)$  и  $V_1(\mathbb{H}^n) = S^{4n-1}$  — единичная сфера в  $\mathbb{H}^n$ .

**Теорема 9.3.9.2.** *Первая ненулевая гомотопическая группа пространства  $GL(n, \mathbb{H})$  — это группа  $\pi_3(GL(n, \mathbb{H})) = \mathbb{Z}$ .*

**Доказательство.** Запишем гомотопическую точную последовательность расслоения  $SP(n+1) \rightarrow V_{n+1}(\mathbb{H}^{n+1})$  со слоем  $SP(n)$ :

$$\dots \rightarrow \pi_{i+1}(S^{4n+3}) \rightarrow \pi_i(Sp(n)) \rightarrow \pi_i(Sp(n+1)) \rightarrow \pi_i(S^{4n+3}) \rightarrow \dots$$

Если  $i \leq 4n+1$ , то  $\pi_{i+1}(S^{4n+3}) = \pi_i(S^{4n+3}) = 0$ , поэтому  $\pi_i(Sp(n)) \cong \pi_i(Sp(n+1))$ . При  $i = 3$  неравенство  $i \leq 4n+1$  выполняется для всех  $n$ , поэтому  $\pi_3(Sp(n)) \cong \pi_3(Sp(1)) = \pi_3(S^3) = \mathbb{Z}$ .  $\square$



**Теорема 9.3.9.3.** *Отображение  $S^3 \rightarrow \mathrm{GL}(n, \mathbb{H})$ , при котором  $\lambda \in S^3 \subset \mathbb{H}$  отображается в матрицу  $\lambda I$ , представляет элемент  $n$  группы  $\pi_3(\mathrm{GL}(n, \mathbb{H})) = \mathbb{Z}$ .*

**Доказательство.** Если проследить за отображениями в точных последовательностях из доказательства теоремы 9.3.9.2, то можно установить, что образующей группы  $\pi_3(\mathrm{GL}(n, \mathbb{H}))$  служит отображение  $S^3 \rightarrow \mathrm{GL}(n, \mathbb{H})$ , при котором  $\lambda \in S^3 \subset \mathbb{H}$  отображается в матрицу  $\mathrm{diag}(\lambda, 1, \dots, 1)$ . Отображение  $S^3 \rightarrow S^3$ , заданное формулой  $\lambda \mapsto \lambda^n$ , имеет степень  $n$ , поэтому отображение  $S^3 \rightarrow \mathrm{GL}(n, \mathbb{H})$ , при котором  $\lambda \in S^3 \subset \mathbb{H}$  отображается в матрицу  $\mathrm{diag}(\lambda^n, 1, \dots, 1)$ , представляет элемент  $n$  группы  $\pi_3(\mathrm{GL}(n, \mathbb{H})) = \mathbb{Z}$ . Остаётся доказать, что отображения  $\lambda \mapsto \mathrm{diag}(\lambda^n, 1, \dots, 1)$  и  $\lambda \mapsto \mathrm{diag}(\lambda, \dots, \lambda)$  гомотопны. Для этого достаточно проверить, что отображения  $\lambda \mapsto \mathrm{diag}(\lambda^{k+1}, 1)$  и  $\lambda \mapsto \mathrm{diag}(\lambda^k, \lambda)$  гомотопны, а затем за несколько шагов построить гомотопию между требуемыми отображениями.

Гомотопия, связывающая отображения  $\lambda \mapsto \mathrm{diag}(\lambda^{k+1}, 1)$  и  $\lambda \mapsto \mathrm{diag}(\lambda^k, \lambda)$ , задаётся формулой

$$f_t(\lambda) = \begin{pmatrix} \lambda^k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix},$$

где  $0 \leq t \leq \pi/2$ .  $\square$

### 9.3.10. Левые собственные значения

Для кватернионных матриц общего вида существование левых собственных значений доказывается более сложно, чем существование правых собственных значений. Это доказательство использует достаточно серьёзную топологическую теорему 9.3.9.3.

**Теорема 9.3.10.1.** *Любая (квадратная) кватернионная матрица имеет левое собственное значение.*

**Доказательство.** [Wo] Предположим, что у кватернионной матрицы  $A$  нет левых собственных значений. Согласно теореме 9.2.3.1 это означает, что для всех  $\lambda \in \mathbb{H}$  матрица  $\lambda I - A$  невырожденная. В частности, матрица  $A$  невырожденная. В дальнейшем мы будем рассматривать только кватернионы единичной длины, т.е. будем считать, что  $\lambda \in S^3 \subset \mathbb{H}$ . Семейство отображений  $f_t(\lambda) = \lambda I - tA$ ,  $0 \leq t \leq 1$ , представляет собой гомотопию в классе отображений  $S^3 \rightarrow \mathrm{GL}(n, \mathbb{H})$ , связывающую отображение  $f_1(\lambda) = \lambda I - A$  с отображением  $f_0(\lambda) = \lambda I$ . С другой стороны, семейство отображений  $g_t(\lambda) = t\lambda I - A$ ,  $0 \leq t \leq 1$ , представляет

собой гомотопию в классе отображений  $S^3 \rightarrow \mathrm{GL}(n, \mathbb{H})$ , связывающую постоянное отображение  $g_0(\lambda) = -A$  с отображением  $g_1(\lambda) = \lambda I - A$ . Таким образом, постоянное отображение  $S^3 \rightarrow -A \in \mathrm{GL}(n, \mathbb{H})$  гомотопно отображению  $S^3 \rightarrow \mathrm{GL}(n, \mathbb{H})$ , заданному формулой  $\lambda \mapsto \lambda I$ . Согласно теореме 9.3.9.3 последнее отображение представляет элемент  $n$  группы  $\pi_3(\mathrm{GL}(n, \mathbb{H})) = \mathbb{Z}$ , а постоянное отображение, очевидно, представляет нулевой элемент этой группы. Полученное противоречие показывает, что у матрицы  $A$  есть левое собственное значение.  $\square$

## Задачи

**3.1.** Пусть  $X = \begin{pmatrix} A & -\bar{B} \\ B & A \end{pmatrix}$  — комплексная матрица, соответствующая гиперэрмитовой матрице  $A + jB$ , и  $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$ . Докажите, что матрица  $JX$  кососимметрическая.

## Решения

**1.1.** Пусть  $e_1, \dots, e_n$  — строки матрицы  $A$ ,  $\varepsilon_1, \dots, \varepsilon_n$  — столбцы матрицы  $A^{-1}$ ,  $e'_1, \dots, e'_n$  — строки матрицы  $B$ ,  $\varepsilon'_1, \dots, \varepsilon'_n$  — столбцы матрицы  $B^{-1}$ . Тогда

$$e'_p \varepsilon'_q = \begin{cases} e_p \varepsilon_q, & \text{если } p \neq i, q \neq j; \\ (e_i + \lambda e_j) \varepsilon_q, & \text{если } p = i, q \neq j; \\ e_p (\varepsilon_j - \varepsilon_i \lambda), & \text{если } p \neq i, q = j; \\ (e_i + \lambda e_j) (\varepsilon_j - \varepsilon_i \lambda), & \text{если } p = i, q = j. \end{cases}$$

Если  $q \neq j$ , то  $(e_i + \lambda e_j) \varepsilon_q = e_i \varepsilon_q$ . Если  $p \neq i$ , то  $e_p (\varepsilon_j - \varepsilon_i \lambda) = e_p \varepsilon_j$ . Если  $p = i$  и  $q = j$ , то  $e'_p \varepsilon'_q = (e_i + \lambda e_j) (\varepsilon_j - \varepsilon_i \lambda) = \lambda - \lambda = 0$ .

**1.2.** Легко проверить, что матрица  $A = \begin{pmatrix} 1 & b \\ c & d \end{pmatrix}$  обратима тогда и только тогда, когда  $d \neq cb$ . Поэтому матрица  $A^T = \begin{pmatrix} 1 & c \\ b & d \end{pmatrix}$  обратима тогда и только тогда, когда  $d \neq bc$ . Условия  $d \neq cb$  и  $d \neq bc$  эквивалентны тогда и только тогда, когда тело  $K$  коммутативно.

**1.3.** Легко проверить, что матрица

$$B = \begin{pmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & 1 & b_{23} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

где

$$b_{ij} = \sum_{i \leq k_1 < k_2 < \dots < k_m < n} (-1)^{m+1} a_{ik_1} a_{k_1+1, k_2} \dots a_{k_m+1, n},$$

является обратной для матрицы  $A$  (при  $m = 0$  мы получаем слагаемое  $-a_{ij}$ ). Действительно,

$$\sum_{i < p < j} a_{ip} b_{pj} + b_{ij} + a_{ij} = \sum_{i < p < j} b_{ip} a_{pj} + b_{ij} + a_{ij} = 0.$$

**1.4.** Пусть  $A = \|a_{ij}\|_1^n$  — верхняя треугольная матрица, т.е.  $a_{ij} = 0$  при  $i > j$ . Предположим, что все диагональные элементы этой матрицы ненулевые. Равенство  $Ax = 0$ , где  $x = (x_1, \dots, x_n)^T$ , эквивалентно системе уравнений

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0; \\ a_{22}x_2 + \dots + a_{2n}x_n &= 0; \\ &\dots \\ a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n &= 0; \\ a_{nn}x_n &= 0. \end{aligned}$$

Из последнего уравнения получаем  $x_n = 0$ . Затем из предпоследнего уравнения получаем  $x_{n-1} = 0$  и т.д. Для нижней треугольной матрицы рассуждения аналогичны, но в этом случае мы последовательно получаем  $x_1 = 0$ ,  $x_2 = 0$  и т.д.

Предположим теперь, что  $a_{kk} = 0$ , причём  $a_{ii} \neq 0$  при  $i < k$ . Пусть  $x = (x_1, \dots, x_n)^T$ . Достаточно найти ненулевое решение системы  $Ax = 0$ , у которого  $x_{k+1} = x_{k+2} = \dots = x_n$ . В таком случае система уравнений запишется в виде

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1k}x_k &= 0; \\ a_{22}x_2 + \dots + a_{2k}x_k &= 0; \\ &\dots \\ a_{k-1,k-1}x_{k-1} + a_{k-1,k}x_k &= 0; \\ a_{kk}x_k &= 0. \end{aligned}$$

Положим  $x_k = 1$ . Поскольку  $a_{k-1,k-1} \neq 0$ , из уравнения  $a_{k-1,k-1}x_{k-1} + a_{k-1,k} = 0$  находим  $x_{k-1}$  и т.д.

**1.5.** Требуется доказать, что столбцы данной матрицы линейно зависимы справа. Умножим столбец с номером  $k$  справа на  $x^{k-1}$ . Легко проверить, что сумма полученных столбцов является столбцом с нулевыми элементами.

**1.6.** Воспользуемся обозначениями из доказательства теоремы 9.1.1.2. Пространство вектор-столбцов  $x$ , для которых выполняется равенство  $Ax = 0$ , — это пространство  $\text{Ker } f$ . Напомним, что  $\dim \text{Ker } f = \dim V - \dim W$ , где  $\dim V = n$  и  $\dim W \leq m$ . Поэтому  $\dim V - \dim W \geq n - m > 0$ , т.е.  $\dim \text{Ker } f > 0$ .

**1.7.** Воспользуемся разложением Брюа  $A = LMU$ . Матрица  $A$  регулярна слева тогда и только тогда когда матрица  $M$  регулярна слева. У матрицы  $A$  есть правая обратная матрица тогда и только тогда, когда у матрицы  $M$  есть правая обратная матрица. А для матрицы  $M$  оба условия эквивалентны тому, что в каждой её строке есть ненулевой элемент.

**1.8.** С помощью разложения Брюа матрицу  $A$  можно представить в виде  $A = LDP U$ , где  $L$  и  $U$  — невырожденные (треугольные) матрицы,  $D = \text{diag}(d_1, \dots, d_r, 0, \dots, 0)$ , а  $P$  — матрица перестановки. В качестве требуемой матрицы  $B$  тогда можно взять матрицу  $LD_1 P U$ , где  $D_1 = \text{diag}(\underbrace{0, \dots, 0}_r, \underbrace{1, \dots, 1}_{n-r})$ .

**2.1.** Согласно теореме 9.2.3.1 нужно доказать, что матрица  $A - \lambda I$  вырожденная тогда и только тогда, когда  $\lambda$  — это один из диагональных элементов матрицы  $A$ . Ясно, что матрица  $A - \lambda I$  треугольная с элементами  $a_{ii} - \lambda$  на диагонали. Согласно задаче 1.4 эта матрица вырожденная тогда и только тогда, когда  $a_{ii} - \lambda = 0$  для некоторого  $i$ .

**2.2.** а) Легко проверить, что  $A^2 = 0$ , поэтому матрица  $A$  нильпотентна. Покажем, что  $\lambda = i - 1$  является левым собственным значением матрицы  $A$ . Согласно теореме 9.2.3.1 это эквивалентно тому, что матрица  $A - \lambda I = \begin{pmatrix} 1 & j \\ -j & 1 \end{pmatrix}$  вырожденная. Остаётся заметить, что матрица  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  вырожденная тогда и только тогда, когда  $d = ca^{-1}b$ .

б) Легко проверить, что  $B^3 = -4B$ , поэтому матрица  $B$  не может быть нильпотентной. Матрица  $B - \lambda I = \begin{pmatrix} j-\lambda & i \\ i & -j-\lambda \end{pmatrix}$  вырожденная тогда и только тогда, когда  $-j - \lambda = i(j - \lambda)^{-1}i$ . Умножая это равенство справа на  $i(j - \lambda)$ , получаем  $(j + \lambda)i(j - \lambda) = i$ , т.е.  $\lambda k + k\lambda = \lambda i\lambda$ . Если  $\lambda \neq 0$ , то, умножая это равенство слева и справа на  $\mu = \lambda^{-1}$ , получаем  $k\mu + \mu k = i$ . Но если  $\mu = a + bi + cj + dk$ , то  $k\mu + \mu k = -2d + 2ak$ ; это число не может быть равно  $i$ .

**3.1.** Ясно, что  $J^T = -J$ , поэтому  $(JX)^T = -X^T J = -J\overline{X^T}$ . Далее,  $\overline{X^T} = X$ , поскольку матрица  $X$  соответствует гиперэрмитовой матрице. Следовательно,  $(JX)^T = -JX$ .

# Книги, рекомендуемые для изучения линейной алгебры

- Беллман Р. *Введение в теорию матриц*, М.: Наука, 1969.
- Гантмахер Ф. Р. *Теория матриц*, М.: Наука, 1988.
- Гельфанд И. М. *Лекции по линейной алгебре*, М.: Наука, 1966.
- Голуб Дж., Ван Лоун Ч., *Матричные вычисления*, М.: Мир, 1999.
- Кострикин А. И., Манин Ю. И. *Линейная алгебра и геометрия*, М.: Наука, 1986.
- Мальцев А. И. *Основы линейной алгебры*, М.: Наука, 1975.
- Маркус М., Минк Х. *Обзор по теории матриц и матричных неравенств*, М.: Наука, 1972.
- Постников М. М. *Лекции по геометрии. Семестр II. Линейная алгебра*, М.: Наука, 1986.
- Постников М. М. *Лекции по геометрии. Семестр V. Группы и алгебры Ли* М.: Наука, 1982.
- Проскураков И. В. *Сборник задач по линейной алгебре*, М.: Наука, 1984.
- Халмош П. *Конечномерные векторные пространства*, М.: Физматгиз, 1963.
- Хорн Р., Джонсон Ч. *Матричный анализ*, М.: Мир, 1989.
- Шилов Г. Е. *Математический анализ. Конечномерные линейные пространства*, М.: Наука, 1969.
- Crowe M. J. *A history of vector analysis*, Notre Dame; London: Univ. of Notre Dame press, 1967.
- Greub W. H. *Linear algebra*, Berlin e.a.: Springer-Verlag, 1967.
- Greub W. H. *Multilinear algebra*, Berlin e.a.: Springer-Verlag, 1967.
- Muir T., Metzler W. H. *A treatise on the theory of determinants* New York, 1930.

# Цитированная литература

- [Ар] Артин Э., *Геометрическая алгебра*, М.: Наука, 1969.
- [Б1] Банах С., *Теория линейных операций*, Ижевск: РХД, 2001.
- [Б2] Белов А. Я., *Об одной задаче комбинаторной геометрии*, УМН **47** вып. 3 (1992), 151–152.
- [Б3] Богоявленский О. И., *Опрокидывающиеся солитоны*, М.: Наука, 1991.
- [Г1] Гершгорин С. А., *Über die Abgrenzung der Eigenwerte einer Matrix*, Известия АН СССР. Сер. физ.-мат. **7** (1931), 749–754.
- [Г2] Гренандер У., Сегё Г., *Тёплицевы формы и их приложения*, М.: ИЛ, 1961.
- [Г3] Григорьев Д. Ю., *Алгебраическая сложность вычисления семейства билинейных форм*, Ж. ВМ и МФ. **19** (1979), 563–580.
- [Д] Дэй М. М., *Нормированные линейные пространства*, М.: ИЛ, 1961.
- [Еп] Епифанов Г. В., *Универсальность сечений кубов*, Матем. заметки **3** (1967), 93–95.
- [И] Иохвидов И. С., *Ганкелевы и тёплицевы матрицы и формы*, М.: Наука, 1974.
- [Ко] Коваленко С. П., *Вычисление угла между подпространствами евклидова пространства*, Мат. методы анализа динамич. систем №7 (1983), 24–27.
- [Ма] Манаков С. В., *Заметка об интегрировании уравнений Эйлера динамики  $n$ -мерного твердого тела*, Функц. анализ и его приложения **10** №4 (1976), 93–94.

- [Ни] Нисневич Л. Б., Брызгалов В. И. *Об одной задаче  $n$ -мерной геометрии*, УМН **8** вып. 4 (1953), 169–172.
- [П1] Прасолов В. В., Тихомиров В. М. *Геометрия*, М.: МЦНМО, 2007.
- [П2] Прасолов В. В., *Многочлены*, М.: МЦНМО, 1999.
- [П3] Прасолов В. В., *Элементы комбинаторной и дифференциальной топологии*, М.: МЦНМО, 2004.
- [П4] Прасолов В. В., *Элементы теории гомологий*, М.: МЦНМО, 2006.
- [Т] *Теория алгебр Ли. Топология групп Ли. Семинар «Софус Ли»*, М.: ИЛ, 1962.
- [Х] Хамфрис Дж., *Введение в теорию алгебр Ли и их представлений*, М.: МЦНМО, 2003.
- [Ч] Чеботарев Н. Г., *Математическая автобиография*, УМН **III** вып. 4 (1948), 3–66.
- [Ш] Шестопал В. Е., *Решение матричного уравнения  $AX - XB = C$* , Матем. заметки **19** (1976), 449–451.
- [Ai] Aitken A. C. *A note on trace-differentiation and the  $\Omega$ -operator*, Proc. Edinburgh Math. Soc. **10** (1953), 1–4.
- [Ad] Adams J. F., *Vector fields on spheres*, Ann. Math. **75** (1962), 603–632.
- [Al] Alesker S., *Non-commutative linear algebra and plurisubharmonic functions of quaternionic variables*, preprint math.CV/0104209.
- [Af] Afriat S. N., *On the latent vectors and characteristic values of products of pairs of symmetric idempotents*, Quart. J. Math. **7** (1956), 76–78.
- [Al] Albert A. A., *On the orthogonal equivalence of sets of real symmetric matrices*, J. Math. Mech. **7** (1958), 219–235.
- [Ar] Arf C., *Untersuchungen über quadratische Formen in Körpern der Charakteristik 2*, J. Reine Angew. Math. **183** (1941), 148–167.
- [As] Aslaksen H., *Quaternionic determinants*, Math. Intelligencer **15** (1996), №3, 57–65.
- [Au1] Au-Yeung Y., *A theorem on a mapping from a sphere to the circle and the simultaneous diagonalization of two Hermitian matrices*, Proc. Amer. Math. Soc. **20** (1969), 545–548.

- [Au2] Au-Yeung Y., *A simple proof of the convexity of the field of values defined by two Hermitian forms*, Aequat. Math. **12** (1975), 82–83.
- [Au3] Aupetit B., *An improvement of Kaplansky's lemma on locally algebraic operators*, Studia Math. **88** (1988), 275–278.
- [Ba1] Barnett S., *Matrices in control theory*, London: Van Nostrand Reinhold, 1971.
- [Ba2] Barnett S., *A note on the Bezoutian matrix*, SIAM J. Appl. Math. **22** (1972), 84–86.
- [Be1] Bell H. E., *Gershgorin's theorem and the zeros of polynomials*, Amer. Math. Monthly **72** (1965), 292–295.
- [Be2] Bellman R., *Notes on matrix theory — IV*, Amer. Math. Monthly **62** (1955), 172–173.
- [Be3] Bellman R., Hoffman A., *On a theorem of Ostrowski and Taussky*, Arch. Math. **5** (1954), 123–127.
- [Be4] Bernau S. J., Wojciechowski P. J., *Images of bilinear mappings into  $\mathbb{R}^3$* , Proc. Amer. Math. Soc. **124** (1996), 3605–3612.
- [Bi] Bilyeu R. G., *Metric definition of the linear structure*, Proc. Amer. Math. Soc. **25** (1970), 205–206.
- [Br1] Brenner J. L., *Matrices of quaternions*, Pacific J. Math. **1** (1951), 329–335.
- [Br2] Brickman L., *On the field of values of a matrix*, Proc. Amer. Math. Soc. **12** (1961), 61–66.
- [Bu] Burgermeister P.-F., *Classification des représentation de la double flèche*, L'Enseignement Math. **32** (1986), 199–210.
- [Ca1] Calabi E., *Linear systems of real quadratic forms*, Proc. Amer. Math. Soc. **15** (1964), 844–846.
- [Ca2] Carlitz L., *A determinant*, Amer. Math. Monthly **64** (1957), 186–188.
- [Ch1] Chevalley C., *A new kind of relationship between matrices*, Amer. J. Math. **65** (1943), 521–531.
- [Ch2] Chan N. N., Kim-Hung Li, *Diagonal elements and eigenvalues of a real symmetric matrix*, J. Math. Anal. Appl. **91** (1983), 622–626.



- [Co] Cohn P. M., *Algebra*, vol. 3, 2nd edition, John Wiley & Sons, 1991.
- [Cu] Cullen C. G., *A note on convergent matrices*, Amer. Math. Monthly **72**, (1965), 1006–1007.
- [Da] Dancis J., *A quantitative formulation of Sylvester law of inertia, III*, Linear Algebra and Appl. **80** (1986), 141–158.
- [Di] Dieudonné J., *Les déterminants sur un corp non-commutatif*, Bull. Soc. Math. France **71** (1943), 27–45.
- [Dj1] Djoković D. Ž., *On the Hadamard product of matrices*, Math. Zeitschrift **86** (1964), 395.
- [Dj2] Djoković D. Ž., *Product of two involutions*, Arch. Math. **18** (1967), 582–584.
- [Dj3] Djoković D. Ž., *A determinantal inequality for projectors in a unitary space*, Proc Amer. Math. Soc. **27** (1971), 19–23.
- [Do] Dodgson C. L., *Condensation of determinants*, Proc. Royal Soc. London **15** (1866), 150–155.
- [Dr1] Drazin M. A., Haynsworth E. V., *Criteria for the reality of matrix eigenvalues*, Math. Zeitschrift **78** (1962), 449–452.
- [Dr2] Drazin M. A., Dungey J. W., Greunderg K. W., *Some theorems on commutative matrices*, J. London Math. Soc. **26** (1951), 221–228.
- [Du] Dunkl C. F., Williams K. S., *A simple norm inequality*, Amer. Math. Monthly **71** (1964), 53–54.
- [Ea] Eastwood M., Penrose R., *Drawing with complex numbers*, preprint math/0001097.
- [Ed] Edelman A., Kostlan E., *The road from Kac's matrix to Kac's random polynomials*, preprint.
- [Ev1] Evans R. J., Isaacs I. M., *Generalized Vandermonde determinants and roots of unity of prime order*, Proc. Amer. Math. Soc. **58** (1976), 51–54.
- [Ev2] Everitt W. N., *A note on positive definite matrices*, Proc. Glasgow Math. Ass. **3** (1958), 173–175.
- [Fa1] Faulhaber J., *Academia algebrae*, Augsburg, 1631.

- [Fa2] Farahat H. K., Ledermann W., *Matrices with prescribed characteristic polynomials*, Proc. Edinburgh Math. Soc. **11** (1958–1959), 143–146.
- [Fe] Fergus G., *A note on matrices with zero trace*, Amer. Math. Monthly **73** (1966), 630–631.
- [Fi1] Fiedler M., *Remarks on the Schur complements*, Linear Algebra and Appl. **39** (1981), 189–195.
- [Fi2] Fillmore P. A., *On the similarity and the diagonal of a matrix*, Amer. Math. Monthly **76** (1969), 167–169.
- [Fl1] Flanders H., *On spaces of linear transformations with bound rank*, J. London Math. Soc. **37** (1962), 10–16.
- [Fl2] Flanders H., Wimmer H. K., *On matrix equations  $AX - XB = C$  and  $AX - YB = C$* , SIAM J. Appl. Math. **32** (1977), 707–710.
- [Fr1] Frank W. M., *A bound on determinants*, Proc. Amer. Math. Soc. **16** (1965), 360–363.
- [Fr2] Franck P., *Sur la meilleure approximation d'une matrice donnée par une matrice singulière*, C. r. Acad. Sci. Paris **253** (1961), 1297–1298.
- [Fr3] Friedland S., *Matrices with prescribed off-diagonal elements*, Israel J. Math. **11** (1972), 184–189.
- [Fu] Fulton W., Harris J., *Representation theory*, Berlin e.a.: Springer-Verlag, 1991.
- [Ga] Gage M. E., *A note on skew-Hopf fibration*, Proc. Amer. Math. Soc. **93** (1985), 145–150.
- [Ge] Gessel I., Viennot G., *Binomial determinants, paths, and hook length formulae*, Adv. Math. **58** (1985), 300–321.
- [Gi] Gibson P. M., *Matrix commutators over algebraically closed field*, Proc. Amer. Math. Soc. **52** (1975), 30–32.
- [Go] Goddard L. S., Schneider H., *Pairs of matrices with a non-zero commutator*, Proc. Cambr. Phil. Soc. **51** (1955), 551–553.
- [Gr1] Green C., *A multiple exchange property for bases*, Proc. Amer. Math. Soc. **39** (1973), 45–50.

- [Gr2] Green C. *Another exchange property for bases*, Proc. Amer. Math. Soc. **46** (1974), 155–156.
- [Gr3] Greenberg M. J. *Note on the Cayley–Hamilton theorem*, Amer. Math. Monthly **91** (1984), 193–195.
- [Gr4] Grosof M. S., Taiani G., *Vandermonde strikes again*, Amer. Math. Monthly **100** (1993), 575–577.
- [Gr5] Grünbaum B., *How many triangles?*, Geombinatorics **8** №1 (1998), 154–159.
- [Ha1] Halmos P. R., *Numerical ranges and normal dilations*, Acta Sci. Math. **25** (1964), 1–5.
- [Ha2] Hausdorff F., *Der Wertvorrat einer Bilinearform*, Math. Zeitschrift **3** (1919), 314–316.
- [Ha3] Haynsworth E. V., *Applications of an inequality for the Schur complement*, Proc. Amer. Math. Soc. **24** (1970), 512–516.
- [Ho] Hoffman A. J., Wielandt H. W., *The variation of the spectrum of a normal matrix*, Duke Math. J. **20** (1953), 37–39.
- [Hs] Hsu P. L., *On symmetric orthogonal and skew-symmetric matrices*, Proc. Edinburgh Math. Soc. **10** (1953), 37–44.
- [Hu] Hu-yun Sh., *Estimation of the eigenvalues of  $AB$  for  $A > 0$ ,  $B > 0$* , Linear Algebra and Appl. **73** (1986), 147–150.
- [J1] Jacob H. G., *Another proof of the rational decomposition theorem*, Amer. Math. Monthly **80** (1973), 1131–1134.
- [J2] Jordan P., von Neumann J., *On inner products in linear metric spaces*, Ann. Math. **36** (1935), 719–723.
- [Ka1] Karlin S., *Total positivity V. I*, Stanford University Press, 1968.
- [Ka2] Kahane J., *Grassmann algebras for proving a theorem on Pfaffians*, Linear Algebra and Appl. **4** (1971), 129–139.
- [Ke] Kelly L. M., Smiley D. M., Smiley M. F., *Two dimensional spaces are quadrilateral spaces*, Amer. Math. Monthly **72** (1965), 753–754.
- [Kl1] Kleinecke D. C., *On operator commutators*, Proc. Amer. Math. Soc. **8** (1957), 535–536.

- [Kl2] Klinger A., *The Vandermonde matrix*, Amer. Math. Monthly **74** (1967), 571–574.
- [Kr1] Krattenthaler C., *Advanced determinant calculus*, preprint math/9902004.
- [Kr2] Krusemeyer M., *Why does Wronskian work?*, Amer. Math. Monthly **95** (1988), 46–49.
- [Li] Li C. K., *A simple proof of the elliptical range theorem*, Proc. Amer. Math. Soc. **124** (1996), 1985–1986.
- [La] Lanczos C., *Linear systems in self-adjoint form*, Amer. Math. Monthly **65** (1958), 665–679.
- [Le] Ledermann W., *A note on skew-symmetric determinants*, Proc. Edinburgh Math. Soc. **36** (1993), 335–338.
- [Ma1] Majindar K. N., *On simultaneous hermitian congruence transformations of matrices*, Amer. Math. Monthly **70** (1963), 842–844.
- [Ma2] Marcus M., Minck H., *On two theorems of Frobenius*, Pacific J. Math. (1975), 149–151.
- [Ma3] Marcus M., Moyls B. N., *Linear transformations on algebras of matrices*, Canad. J. Math. **11** (1959), 61–66.
- [Ma4] Marcus M., Purves R., *Linear transformations on algebras of matrices: the invariance of the elementary symmetric functions*, Canad. J. Math. **11** (1959), 383–396.
- [Ma5] Massey W. S., *Cross product of vectors in higher dimensional euclidean spaces*, Amer. Math. Monthly **90** (1983), 697–701.
- [Me] Merris R., *Equality of decomposable symmetrized tensors*, Canad. J. Math. **27** (1975), 1022–1024.
- [Mi1] Mirzakhani M., *A simple proof of a theorem of Schur*, Amer. Math. Monthly **105** (1998), 260–262.
- [Mi2] Mirsky L., *An inequality for positive definite matrices*, Amer. Math. Monthly **62** (1955), 428–430.
- [Mi3] Mirsky L., *On generalization of Hadamard's determinantal inequality due to Szasz*, Arch. Math. **8** (1957), 274–275.

- [Mi4] Mirsky L., *A trace inequality of John von Neuman*, Monatsh. Math. **79** (1975), 303–306.
- [Mi5] Mitchell O. H., *Note on determinants of powers*, Amer. J. Math. **4** (1881), 341–344.
- [Mo1] Mohr E., *Einfacher Beweis des verallgemeinerter Determinantensatzes von Sylvester nebst einer Verschärfung*, Math. Nachr. **10** (1953), 257–260.
- [Mo2] Mond B., *A matrix version of Rennie's generalization of Kantorovich's inequality*, Proc. Amer. Math. Soc. **16** (1965), 1131.
- [Mo3] Motzkin T. S., *Determinants whose elements have equal norm*, Proc. Amer. Math. Soc. **11** (1960), 871–872.
- [Mo4] Moore E. H., *On determinants of an Hermitian matrix of quaternionic elements*, Bull. Amer. Math. Soc. **28** (1922), 161–162.
- [Mo5] Moore E. H., *General Analysis — Part I*, Mem. Amer. Phil. Soc. **1** (1935), 197.
- [Ne1] Newcomb R. W., *On the simultaneous diagonalization of two semi-definite matrices*, Quart. Appl. Math. **19** (1961), 144–146.
- [Ne2] Newman D. J., *Another proof of the minimax theorem*, Proc. Amer. Math. Soc. **11** (1960), 692–693.
- [Os] Ostrowski A. M., *On Schur's complement*, J. Comb. Theory **14** (1973), 319–322.
- [Pe1] Peitgen H. O., *On the Lefschetz number for iterates of continuous mappings*, Proc. Amer. Math. Soc. **54** (1976), 441–444.
- [Pe2] Penrose R. A., *A generalized inverse for matrices*, Proc. Cambr. Phil. Soc. **51** (1955), 406–413.
- [Pr] Proctor R. A., *Equivalence of the combinatorial and the classical definition of Schur functions*, J. Combinatorial Theory, Series A **51** (1989), 135–137.
- [Ra1] Ramakrishnan A., *A matrix decomposition theorem*, J. Math. Anal. Appl. **40** (1972), 36–38.
- [Ra2] Rado R., *Note on generalized inverses for matrices*, Proc. Cambr. Phil. Soc. **52**, 600–601.

- [Rh] de Rham. G., *Sur un théorème de Stieltjes relatif a certaines matrices*, Publ. Inst. Math. Acad. Serbe Sci. **4** (1952), 133–134.
- [Ro1] Robinson H., *An alternative to the Plücker relations*, Proc. Amer. Math. Soc. **66** (1977), 237–240.
- [Ro2] Roth W. E., *The equations  $AX - YB = C$  and  $AX - XB = C$  in matrices*, Proc. Amer. Math. Soc. **3** (1952), 392–396.
- [Sa] Samelson H., Thrall R. M., Wesler O., *A partition theorem for Euclidean  $n$ -space*, Proc. Amer. Math. Soc. **9** (1958), 805–807.
- [Sc1] Scharlau W., *Quadratic and Hermitian forms*, Berlin e.a.: Springer-Verlag, 1985.
- [Sc2] Scharnhorst K., *Angles in complex vector spaces*, preprint math/9904077.
- [Sc3] Scherk P., *On sets of vectors*, Math. Scand. **16** (1965), 38–40.
- [Sc4] Schur I., *Zur Theorie der vertauschbaren Matrizen*, J. reine angew. Math. **130** (1905), 66–76.
- [Sc5] Schwert H., *Direct proof of Lanczos' decomposition theorem*, Amer. Math. Monthly **67** (1960), 855–860.
- [Sh1] Shannon R. W., *Simplicial cells in arrangements of hyperplanes*, Geometriae Dedicata **8** (1979), 178–187.
- [Sh2] Shapiro H., *Commutators which commute with one factor*, Pacific J. Math. **181** (1997), 323–336.
- [Si] Šidák Z., *O počtu kladných prvků v močinách nezáporné matice*, Časopis pěst. mat. **89** (1964), 28–30.
- [Sm] Smiley M. F. *Matrix commutators*, Canadian J. Math. **13**, (1961), 353–355.
- [St1] Strassen V. *Gaussian elimination is not optimal*, Num. Math. **13** (1969), 354–356. Русский перевод: Киб. сб. **7** (1970), 366–370.
- [St2] Study R., *Zur Theorie der linearen Gleichungen*, Acta Math. **42** (1920), 1–61.
- [Ta] Taussky O., Zassenhaus H., *On the similarity of transformation between a matrix and its transpose*, Pacific J. Math. **9** (1959), 893–896.

- [To] Toeplitz O., *Das algebraische Analogon zu einem Satze von Fejér*, Math. Zeitschrift **2** (1918), 187–197.
- [Vä1] Väliäho H., *An elementary approach to the Jordan form of a matrix*, Amer. Math. Monthly **93** (1986), 30–32.
- [Vä2] Väliäho H., *Note on pencils of matrices*, Int. J. Control **45** (1987), 1487–1488.
- [Wa] Warwick K., *Using the Cayley-Hamilton theorem with  $N$ -partitioned Matrices*, IEEE Trans. Aut. Control. **28** (1983), 1127–1128.
- [Wi1] Wilf H. S., *A combinatorial determinant*, preprint math/9809120.
- [Wi2] Wimmer H. K., *The matrix equation  $X - AXB = C$  and an analogue of Roth's theorem*, Linear Algebra and Appl. **109** (1988), 145–147.
- [Wo] Wood R. M. W., *Quaternionic eigenvalues*, Bull. London. Math. Soc. **17** (1985), 137–138.
- [Y1] Yang Y., *A matrix trace inequality*, J. Math. Anal. Appl. **133** (1988), 573–574.
- [Y2] Yokonuma T., *Tensor spaces and exterior algebra*, Amer. Math. Soc., 1992.
- [Za] Zassenhaus H., *A remark on a paper of O. Taussky*, J. Math. Mech. **10** (1961), 179–180.
- [Ze1] Zeilberger D., *Reverend Charles to the aid of major Percy and Fields Medalist Enrico*, Amer. Math. Monthly **103** (1996), 501–502.
- [Ze2] Zeilberger D., *Dodgson's determinant-evaluation rule proved by two-timing men and women*, preprint math/9808079.
- [Zh1] Zhan X., *Matrix inequalities*, Lecture Notes in Math, V. 1790, Springer, 2002.
- [Zh2] Zhang F., *Quaternions and matrices of quaternions*, Linear Algebra and Appl. **251** (1997), 21–57.
- [Zh3] Zhang F., *Jordan canonical form of a partitioned complex matrix and its application to real quaternion matrices*, Commun. in Algebra **29(6)** (2001), 2363–2375.

# Предметный указатель

$J$ -кососимметрическая матри-  
ца, 257

$J$ -ортогональная матрица, 257

$k$ -линейное  
отображение, 277

$S$ -функция, 72

абелева

алгебра Ли, 378

группа, 120

конечно порождённая,  
121

свободная, 121

абелевой

группы

базис, 121

ранг, 121

абсолютные

барицентрические

координаты, 110

Адамара

неравенство, 322, 323

произведение, 340

алгебр Ли

гомоморфизм, 378

изоморфизм, 378

алгебра

внешняя, 286

Грассмана, 286

Клиффорда, 404

Кэли, 398

Ли, 377

абелева, 378

алгебраическая, 384

коммутативная, 378

матричная, 378, 394

нильпотентная, 378

полупростая, 385

простая, 385

разрешимая, 378

октав, 398

алгебраическая

алгебра

Ли, 384

алгебраическое

дополнение, 44

замыкание, 384

алгебры

представление, 406

удвоение, 392

алгебры Ли

представление, 382

альтернатива

Фредгольма, 114

альтернирование, 285

аннулирующий

многочлен, 211, 212

аннулятор, 107

антиголоморфное

подпространство, 162

антилинейное

отображение, 160



- антисимметризация, 285
- Артина
  - плоскость, 424
  - пространство, 424
- Арфа
  - инвариант, 415
  - теорема, 415
- ассоциированная матрица, 49
- ассоциированное отображение, 313
- аффинное
  - отображение, 109
  - подпространство, 109
  - пространство, 109
- Базена
  - тождество, 41
- базис, 101
  - абелевой группы, 121
  - жорданов, 200
  - ортогональный, 133
  - ортонормированный, 133, 155
  - симплектический, 413
- базисный минор, 42
- базисы
  - одинаково ориентированные, 119
  - противоположно ориентированные, 119
- барицентрические
  - координаты, 110
  - абсолютные, 110
- Безу
  - матрица, 419
- безутиана, 419
- Бернсайда
  - теорема, 407
- Бернулли
  - числа, 70
- билинейная
  - симметрическая функция, 240
  - форма, 278
  - кососимметрическая, 252
- Бине—Коши
  - формула, 43, 295
- биномиальный определитель, 35
- Биркгофа
  - теорема, 349
- Брюа
  - нормальная форма, 476
  - разложение, 477
- валентность
  - ковариантная, 280
  - контравариантная, 280
  - общая, 280
- Вандермонда
  - матрица
  - обобщённая, 52
  - опредетель, 24
- Вейля
  - неравенство, 351
- вектор, 100
  - изотропный, 424
  - ковариантный, 102
  - контравариантный, 102
  - положительный, 341
  - собственный, 189
  - левый, 487
  - правый, 487
  - экстремальный, 342
- векторное
  - поле
  - линейное, 402

- произведение, 293, 394, 400, 402
- пространство, 100
  - левое, 471
  - правое, 469
- векторы
  - линейно зависимые, 100
  - ортогональные, 423
- верхняя унитарная матрица, 476
- весовая функция, 147
- вещественная форма, 153
  - часть кватерниона, 393
- вещественно-комплексное пространство, 154
- вещественное подпространство, 153
- вещественный кватернион, 393
- Витта теорема, 424
- внешняя алгебра, 286
- Вольтерра уравнения, 443
- вполне вещественное подпространство, 162
- неотрицательная матрица, 52
- положительная матрица, 52
- приводимое представление, 406
- вронскиан, 437
- Вронского определитель, 437
- выпуклая линейная комбинация, 127
- Гамильтона—Кэли теорема, 212
- ганеклева форма, 430
- ганкелева матрица, 430
- Гаусса разложение, 227
- Герглоца—Крейна теорема, 432
- Гершгорина круги, 330
- Гёльдера неравенство, 173
- гиперэрмитова матрица, 491
  - форма, 491
    - положительно определённая, 491
- гиперэрмитово пространство, 491
  - сопряжённый оператор, 491
- главный минор, 42
- голоморфное подпространство, 161
- гомоморфизм алгебр Ли, 378
- Гофмана—Виландта теорема, 352
- Грама матрица, 135, 241
  - опредетель, 135
  - разложение, 227
- Грама—Шмидта

- ортогонализация, 134
- Грассмана
  - алгебра, 286
- Грина
  - матрица, 51
- группа
  - абелева, 120
- дважды
  - стохастическая матрица, 348
- двойственная
  - норма, 172
- двойственное
  - пространство, 102
- двойственный
  - набор, 261
  - оператор, 102
- делители
  - элементарные, 229
- детерминант, 19
- Джамбелли
  - формула, 75
- диагоналируемый
  - оператор, 189
- диагональ
  - матрицы, 349
- диаграмма
  - Юнга, 72, 261
- диаметр, 168
- дискриминант, 67, 421
- дифференцирование, 368
- длина
  - вектора, 133
  - разбиения, 71
- доминирующая
  - диагональ, 127
- дополнение
  - алгебраическое, 44
  - ортогональное, 107, 423
  - по Шуру, 58, 338
- Дьёдонне
  - определитель, 480
- евклидова
  - норма, 336
- единичная
  - сфера, 166
- единичный
  - шар, 166
- жорданов
  - базис, 200
- жорданова
  - клетка, 200
  - матрица, 200, 367
- закон
  - инерции, 243
- знакоопределённая
  - квадратичная форма, 241
  - симметрическая матрица, 241
- значение
  - собственное, 189
- значения
  - сингулярные, 329
  - собственные, 119
  - характеристические, 119
- идеал, 378
- идемпотент, 262
- изометрическое
  - отображение, 170
- изометричные
  - пространства, 170, 423
- изометрия, 423
- изоморфизм
  - алгебр Ли, 378
- изотропный
  - вектор, 424

- инвариант
  - Арфа, 415
- инвариантное
  - подпространство, 406
- инвариантные
  - множители, 229
- инволюция, 267
- инерции
  - закон, 243
- интеграл
  - первый, 441
- интерполяционный
  - многочлен
    - Лагранжа, 25
- Йордана—фон Ноймана
  - теорема, 165
- каноническая
  - ориентация, 154
  - форма
    - Фробениуса, 216
- Канторовича
  - неравенство, 327
- Капелли
  - тождество, 485
- Картана
  - критерий, 386
- Каца
  - матрица, 196
- квадратичная
  - форма, 119, 240, 410
  - знакоопределённая, 241
  - невырожденная, 411—413
  - неотрицательно определённая, 241
- квадратичная форма
  - отрицательно
    - определённая, 241
  - положительно
    - определённая, 241
- квадратичной
  - формы
    - матрица, 240
- квадратичные
  - формы
    - эквивалентные, 411
- кватернион, 393
  - вещественный, 393
  - чисто
    - мнимый, 393
- кватерниона
  - вещественная
    - часть, 393
  - мнимая
    - часть, 393
- кватернионная
  - структура, 490
  - унитарная
    - матрица, 493
- кватернионные
  - ортогональные
    - векторы, 493
- кватернионный
  - единичный
    - вектор, 493
- Киллинга
  - форма, 386
- классификация
  - пар линейных отображений, 205
- клетка
  - жорданова, 200
  - циклическая, 216
- Клиффорда
  - алгебра, 404
- ковариантная
  - валентность, 280
- ковариантность, 119
- ковариантный
  - вектор, 102
- ковектор, 102

- коммутант, 477
- коммутативная
  - алгебра Ли, 378
- коммутатор, 368, 377, 497
- коммутирующие
  - матрицы, 361, 451
- комплексификация
  - оператора, 153
  - пространства, 153
- комплексная
  - структура, 157, 402
- комплексно
  - сопряжённое
    - пространство, 160
- комплексное
  - проективное
    - пространство, 161
- компоненты
  - тензора, 280
- конечно порождённая
  - абелева
    - группа, 121
- конечномерное
  - пространство, 101
- континуанта, 32
- контравариантная
  - валентность, 280
- контравариантность, 119
- контравариантный
  - вектор, 102
- конус, 124
- координаты, 119
  - барицентрические, 110
  - абсолютные, 110
  - тензора, 280
- коразмерность, 102
- косая
  - функция
    - Шура, 73
- кососимметрическая
  - билинейная
    - форма, 252
- матрица, 36, 251
- форма, 278
- функция, 287
- кососимметрический
  - многочлен, 67
  - оператор, 253
  - тензор, 285
- косоэрмитов
  - оператор, 156
- Коши
  - матрица, 52, 57
  - определитель, 27
  - теорема, 331, 332
- коядро, 117
- Крамера
  - правило, 23
- критерий
  - Картана, 386
  - Сильвестра, 242
- Кронекера
  - произведение, 281
  - теорема, 206
- Кронекера–Капелли
  - теорема, 114
- кронекеровское
  - произведение, 281
- круги
  - Гершгорина, 330
- Куранта–Фишера
  - теорема, 245
- кэлеров
  - угол, 162
- Кэли
  - алгебра, 398
  - преобразование, 255
- Лагранжа
  - интерполяционный
    - многочлен, 25
  - теорема, 244

- Лакса  
уравнение, 441
- Лапласа  
теорема, 45
- левое  
векторное  
пространство, 471  
собственное  
значение, 487
- левый  
собственный  
вектор, 487
- Лежандра  
многочлены, 149
- лексикографический  
порядок, 290
- лемма  
Уайтхеда, 475
- Ли  
алгебра, 377  
абелева, 378  
коммутативная, 378  
матричная, 378  
теорема, 381  
факторалгебра, 378
- линейная  
комбинация  
выпуклая, 127  
функция, 102
- линейно  
зависимые  
векторы, 100  
функции, 437  
независимые  
элементы, 121
- линейное  
векторное  
поле, 402  
отображение, 470  
пространство, 100
- Льюиса Кэрролла  
тождество, 48, 50
- Мазура—Улама  
теорема, 170
- матриц  
произведение, 23
- матрица  
квадратичной  
формы, 240
- матрица  
 $J$ -кососимметрическая, 257  
 $J$ -ортогональная, 257  
ассоциированная, 49  
Безу, 419  
Вандермонда  
обобщённая, 52  
верхняя  
унитреугольная, 476  
вполне  
неотрицательная, 52  
положительная, 52  
ганкелева, 430  
гиперэрмитова, 491  
Грама, 135, 241  
Грина, 51  
дважды  
стохастическая, 348  
жорданова, 200, 367  
Каца, 196  
комплексная  
кососимметрическая, 156  
симметрическая, 156  
кососимметрическая, 36, 251  
Коши, 52, 57  
невыврожденная, 471  
неотрицательная, 341  
нижняя  
унитреугольная, 476  
нильпотентная, 259  
нормальная, 494

- обобщённая
  - обратная, 426
- обратимая, 471
- обратная, 45
- ортогональная, 133, 254
- перестановки, 211
- положительная, 341
- примитивная, 345
- присоединённая, 45
- разложимая, 342
- с доминирующей диагона-  
лью, 127
- Сильвестра, 417
- симметрической
  - билинейной функции,  
240
- сопровождающая, 28, 193,  
334
- сопряжённая, 156
- стохастическая
  - по строкам, 353
- тёплицева, 430
- трёхдиагональная, 32
- унитарная
  - кватернионная, 493
- Фробениуса, 28
- циклическая, 28
- элементарная, 474
- Якоби, 32, 51, 191
- матрица-единица, 228
- матрицы
  - диагональ, 349
  - коммутирующие, 361, 451
  - одновременно
    - триангулируемые, 370
  - перестановочные, 361
  - подобные, 199
  - ранг, 42
  - расширение, 450
  - сопряжённые, 199
  - экспонента, 434
- матричная
  - алгебра
    - Ли, 394
- метрика, 168
  - Фубини–Штуди, 161
- метрическое
  - пространство, 168
- минимум, 147
- минимальный
  - многочлен, 211, 212, 409
- Минковского
  - неравенство, 164
- минор, 42
  - базисный, 42
  - главный, 42
  - угловой, 42
- Митчелла
  - теорема, 53
- мнимая
  - часть
    - кватерниона, 393
- многообразие
  - Штифеля, 501
- многочлен
  - аннулирующий, 211, 212
  - вспомогательный, 29
  - кососимметрический, 67
  - минимальный, 211, 212, 409
  - симметрический, 63
  - характеристический, 119,  
189, 294
- многочлены
  - Лежандра, 149
  - ортогональные, 147
  - Чебышева, 152
  - Эрмита, 150
- множители
  - инвариантные, 229
- мономиальный
  - симметрический
    - многочлен, 64

мономорфизм, 117

Мура

определитель, 499

Мура—Пенроуза

теорема, 426

набор

двойственный, 261

Наполеона

теорема, 30

обобщённая, 31

невыврожденная

квадратичная

форма, 411–413

матрица, 471

форма, 278

невыврожденное

пространство, 423

скалярное

произведение, 423

неотрицательная

матрица, 341

неотрицательно

определённая

квадратичная форма, 241

симметрическая матрица, 241

непрерывное

отображение, 167

неприводимое

представление, 382, 406

неравенство

Адамара, 322, 323

Вейля, 351

Гёльдера, 173

Канторовича, 327

Минковского, 164

Оппенгейма, 340

Сасса, 323

Сильвестра, 128

треугольника, 164, 168

Фробениуса, 128

Шура, 327

нижняя

унитреугольная

матрица, 476

нильпотентная

алгебра Ли, 378

матрица, 259

норма, 164

двойственная, 172

евклидова, 336

операторная, 335

спектральная, 335

нормализатор, 380

нормальная

матрица, 494

форма

Брюа, 476

Смита, 227

нормальный

оператор, 157, 257

нормированное

пространство, 164

нормы

эквивалентные, 168

Ньютона

формулы, 65

обёртывающее

подпространство, 299

обобщённая

матрица

Вандермонда, 52

обратная

матрица, 426

теорема

Наполеона, 31

обобщённое

тождество

Сильвестра, 50

образ, 112



- обратимая
  - матрица, 471
- обратная
  - матрица, 45
- общая
  - валентность, 280
- овеществление, 154
- одновременно
  - триангулируемые матрицы, 370
- октав
  - алгебра, 398
- оператор
  - гиперэрмитово сопряжённый, 491
  - двойственный, 102
  - диагонализированный, 189
  - изменяющий ориентацию, 119
  - кососимметрический, 253
  - косоэрмитов, 156
  - нормальный, 157, 257
  - полупростой, 189
  - сжимающий, 224
  - сопряжённый, 102, 156
  - сохраняющий ориентацию, 119
  - унипотентный, 204
  - унитарный, 156
  - эрмитов, 156
- операторная
  - норма, 335
- Оппенгейма
  - неравенство, 340
- определитель, 19
  - биномиальный, 35
  - Вандермонда, 24
  - Вронского, 437
  - Грама, 135
  - Дьёдонне, 480
  - Коши, 27
  - Мура, 499
  - системы линейных уравнений, 24
  - Смита, 33
  - Штуди, 497
- определителя
  - разложение по строке, 21
- ориентация, 119
  - каноническая, 154
- ортогонализация
  - Грама—Шмидта, 134
- ортогональная
  - матрица, 133, 254
  - проекция, 135
- ортогональное
  - дополнение, 107, 423
  - преобразование, 119
- ортогональные
  - векторы, 423
  - многочлены, 147
  - эрмитовы проекторы, 263
- ортогональный
  - базис, 133
- ортонормированный
  - базис, 133, 155
- отображение
  - $k$ -линейное, 277
  - антилинейное, 160
  - ассоциированное, 313
  - аффинное, 109
  - изометрическое, 170
  - линейное, 470
  - непрерывное, 167
  - полилинейное, 277
  - Ходжа, 291
- отрицательно
  - определённая квадратичная форма, 241

- симметрическая матрица, 241
- пара
  - отображений
  - разложимая, 205
- параллелограмма
  - тождество, 165
- первый
  - интеграл, 441
- перестановки
  - матрица, 211
- перестановочные
  - матрицы, 361
- плоскость
  - Артина, 424
- Плюккера
  - соотношения, 299, 301
- подобные
  - матрицы, 199
- подпространство, 100
  - антиголоморфное, 162
  - аффинное, 109
  - вещественное, 153
  - вполне
    - вещественное, 162
  - голоморфное, 161
  - инвариантное, 406
- подразбиение, 73
- полилинейное
  - отображение, 277
- полный однородный
  - симметрический
  - многочлен, 64
- положительная
  - матрица, 341
- положительно
  - определённая
    - квадратичная форма, 241
    - симметрическая матрица, 241
  - эрмитова форма, 241
- положительный
  - вектор, 341
- полупростая
  - алгебра
    - Ли, 385
- полупростой
  - оператор, 189
- поля
  - характеристика, 408
- полярное
  - разложение, 222
- порядок
  - лексикографический, 290
- последовательность
  - точная, 116
- правило
  - Крамера, 23
- правое
  - векторное
    - пространство, 469
  - собственное
    - значение, 487
- правый
  - собственный
    - вектор, 487
- представление
  - алгебры, 406
  - алгебры Ли, 382
  - вполне
    - приводимое, 406
    - неприводимое, 382, 406
    - присоединённое, 368, 382
    - точное, 382
- представления
  - эквивалентные, 382
- преобразование
  - Кэли, 255
  - ортогональное, 119
  - элементарное, 474
- примитивная

матрица, 345  
присоединённая  
матрица, 45  
присоединённое  
представление, 368, 382  
проективное  
пространство  
комплексное, 161  
проектор, 262, 338  
проекторы  
эрмитовы  
ортогональные, 263  
проекция, 116, 262  
ортогональная, 135  
произведение  
Адамара, 340  
векторное, 293, 394, 400, 402  
Кронекера, 281  
кронекеровское, 281  
матриц, 23  
скалярное, 133  
тензорное, 278  
эрмитово, 155  
производящая  
функция, 63  
простая  
алгебра  
Ли, 385  
пространства  
изометричные, 170, 423  
пространство  
Артина, 424  
аффинное, 109  
векторное, 100  
вещественно-комплексное, 154  
гиперэрмитово, 491  
двойственное, 102  
комплексно сопряжённое, 160

конечномерное, 101  
линейное, 100  
метрическое, 168  
невыврожденное, 423  
нормированное, 164  
со скалярным произведе-  
нием, 165  
сопряжённое, 102  
эрмитово, 155  
Птолемея  
теорема  
обобщённая, 133  
пфаффиан, 62, 296  
радиус  
спектральный, 335  
разбиение, 71  
сопряжённое, 72  
разбиения  
длина, 71  
разложение  
Брюа, 477  
Гаусса, 227  
Грама, 227  
определителя  
по строке, 21  
полярное, 222  
Фиттинга, 210  
Холецкого, 227  
циклическое, 216  
Шура, 224, 495  
разложимая  
матрица, 342  
пара отображений, 205  
разложимый  
кососимметрический  
тензор, 298  
симметрический  
тензор, 298  
размерность, 101

аффинного пространства,  
     109  
 разрешимая  
     алгебра Ли, 378  
 ранг, 305  
     абелевой  
         группы, 121  
     левый  
         столбцовый, 472  
         строчный, 472  
     матрицы, 42  
     правый  
         столбцовый, 472  
         строчный, 472  
 расстояние, 168  
 расширение  
     матрицы, 450  
 результат, 417  
 реплика, 374  
 решётка, 120  
  
 Сасса  
     неравенство, 323  
 свёртка  
     тензора, 281  
 свободная  
     абелева  
         группа, 121  
 сечения, 128  
 сжимающий  
     оператор, 224  
 сигнатура, 243  
 Сильвестра  
     критерий, 242  
     матрица, 417  
     неравенство, 128  
     тождество, 50  
         обобщённое, 50  
 симметризация, 285  
 симметрическая  
     билинейная функция, 240

матрица  
     знакоопределённая, 241  
     неотрицательно опреде-  
         лённая, 241  
     отрицательно определён-  
         ная, 241  
     положительно опреде-  
         лённая, 241  
 форма, 278  
 симметрический  
     многочлен, 63  
     мономиальный, 64  
     полный однородный, 64  
     элементарный, 63  
 тензор, 285  
 симметрической  
     билинейной функции  
         матрица, 240  
 симметрия, 136  
 симплектический  
     базис, 413  
 сингулярные  
     значения, 329  
 скалярное  
     произведение, 133  
     невырожденное, 423  
 след, 188, 279  
 Смита  
     нормальная форма, 227  
     определитель, 33  
     теорема, 228  
 собственное  
     значение, 189  
     левое, 487  
     правое, 487  
 собственные  
     значения, 119  
     числа, 119  
 собственный  
     вектор, 189  
 соотношения

- Плюккера, 299, 301
- сопровождающая
  - матрица, 28, 193, 334
- сопряжение, 154
- сопряжённая
  - матрица, 156
- сопряжённое
  - пространство, 102
  - разбиение, 72
- сопряжённые
  - матрицы, 199
- сопряжённый
  - оператор, 102, 156
- спектр, 189
- спектральная
  - норма, 335
- спектральный
  - радиус, 335
- степенная
  - сумма, 64
- Стильгеса—де Рама
  - теорема, 347
- стохастическая
  - по строкам
  - матрица, 353
- структура
  - кватернионная, 490
  - комплексная, 157, 402
- сумма
  - степенная, 64
- сфера
  - единичная, 166
- таблица, 78
- тензор, 280
  - кососимметрический, 285
  - разложимый, 298
  - симметрический, 285
- тензора
  - компоненты, 280
  - координаты, 280
- свёртка, 281
- тензорное
  - произведение, 278
  - отображений, 281
- теорема
  - Арфа, 415
  - Бернсайда, 407
  - Биркгофа, 349
  - Витта, 424
  - Гамильтона—Кэли, 212
  - Герглоца—Крейна, 432
  - Гофмана—Виландта, 352
  - Йордана—фон Ноймана, 165
  - Коши, 331, 332
  - Кронекера, 206
  - Кронекера—Капелли, 114
  - Куранта—Фишера, 245
  - Лагранжа, 244
  - Лапласа, 45
  - Ли, 381
  - Мазура—Улама, 170
  - Митчелла, 53
  - Мура—Пенроуза, 426
  - Наполеона, 30
  - обобщённая, 31
  - о минимаксе, 147
  - об альтернативе, 147
  - Птолемея
    - обобщённая, 133
  - Смита, 228
  - Стильгеса—де Рама, 347
  - Тёплица—Хаусдорфа, 448
  - Фаульгабера—Якоби, 68
  - Фробениуса—Кенига, 349
  - Хэйнсворт, 61
  - Чеботарёва, 55
  - Шура, 340, 366
  - Энгеля, 380
  - эргодическая, 266
  - Якоби, 242

- Тёплица–Хаусдорфа  
   теорема, 448  
 тёплицева  
   матрица, 430  
   форма, 431  
 тип  
   тензора, 280  
 Тоды  
   цепочка, 441  
 тождество  
   Базена, 41  
   Капелли, 485  
   Льюиса Кэрролла, 48, 50  
   параллелограмма, 165  
   Сильвестра, 50  
   обобщённое, 50  
   Якоби, 47, 368, 377  
   Якоби–Труди, 73  
 точка  
   аффинного пространства,  
     109  
 точная  
   последовательность, 116  
 точное  
   представление, 382  
 треугольника  
   неравенство, 164, 168  
 трёхдиагональная  
   матрица, 32  
 Уайтхеда  
   лемма, 475  
 угловой  
   минор, 42  
 угол, 138  
   кэлеров, 162  
   эрмитов, 161  
 удвоение  
   алгебры, 392  
 унитарный  
   оператор, 156  
 унитарная  
   матрица, 476  
 уравнение  
   Лакса, 441  
   Эйлера, 441  
 уравнения  
   Вольтерра, 443  
 факторалгебра  
   Ли, 378  
 факторпространство, 116  
 Фаульгабера–Якоби  
   теорема, 68  
 Фибоначчи  
   числа, 41  
 Фиттинга  
   разложение, 210  
 форма  
   билинейная, 278  
   вещественная, 153  
   ганеклева, 430  
   гиперэрмитова, 491  
   квадратичная, 119, 240, 410  
   невырожденная, 411–413  
   Киллинга, 386  
   кососимметрическая, 278  
   невырожденная, 278  
   положительно определённая  
   гиперэрмитова, 491  
   представляет  
   элемент, 411  
   симметрическая, 278  
   тёплицева, 431  
 формула  
   Бине–Коши, 43, 295  
   Джамбелли, 75  
 формулы  
   Ньютона, 65

- формы
  - квадратичные
  - эквивалентные, 411
- Фредгольма
  - альтернатива, 114
- Фробениуса
  - каноническая форма, 216
  - матрица, 28
  - неравенство, 128
- Фробениуса–Кенига
  - теорема, 349
- Фубини–Штуди
  - метрика, 161
- функции
  - линейно
  - зависимые, 437
- функция
  - весовая, 147
  - кососимметрическая, 287
  - линейная, 102
  - производящая, 63
  - Шура, 72, 199
    - косая, 73
  - Эйлера, 33
- характеристика
  - поля, 408
- характеристические
  - значения, 119
  - числа, 119
- характеристический
  - многочлен, 119, 189, 294
- Ходжа
  - отображение, 291
- Холецкого
  - разложение, 227
- Хэйнсворт
  - теорема, 61
- центр, 378
  - масс, 110
- цепочка
  - Тоды, 441
- циклическая
  - клетка, 216
  - матрица, 28
- циклическое
  - разложение, 216
- циркулянт, 29, 193
- Чеботарёва
  - теорема, 55
- Чебышева
  - многочлены, 152
- числа
  - Бернулли, 70
  - собственные, 119
  - Фибоначчи, 41
  - характеристические, 119
- числовой
  - образ
    - оператора, 448
- чисто
  - мнимый
    - кватернион, 393
- шар
  - единичный, 166
- Штифеля
  - многообразие, 501
- Штуди
  - определитель, 497
- Шура
  - дополнение, 58, 338
  - неравенство, 327
  - разложение, 224, 495
  - теорема, 340, 366
  - функция, 72, 199
- Эйлера
  - уравнение, 441
  - функция, 33

эквивалентные  
    квадратичные  
        формы, 411  
    нормы, 168  
    представления, 382  
экспонента  
    матрицы, 434  
экстремальный  
    вектор, 342  
элементарная  
    матрица, 474  
элементарное  
    преобразование, 474  
элементарные  
    делители, 229  
элементарный  
    симметрический  
        многочлен, 63  
элементы  
    линейно  
        независимые, 121  
Энгеля  
    теорема, 380  
эпиморфизм, 117  
эргодическая  
    теорема, 266  
Эрмита  
    многочлены, 150  
эрмитов  
    оператор, 156  
    угол, 161  
эрмитова  
    форма  
        положительно определённая, 241  
эрмитово  
    произведение, 155  
    пространство, 155  
Юнга  
    диаграмма, 72, 261

ядро, 112  
Якоби  
    матрица, 32, 51, 191  
    теорема, 242  
    тождество, 47, 368, 377  
Якоби—Труди  
    тождество, 73