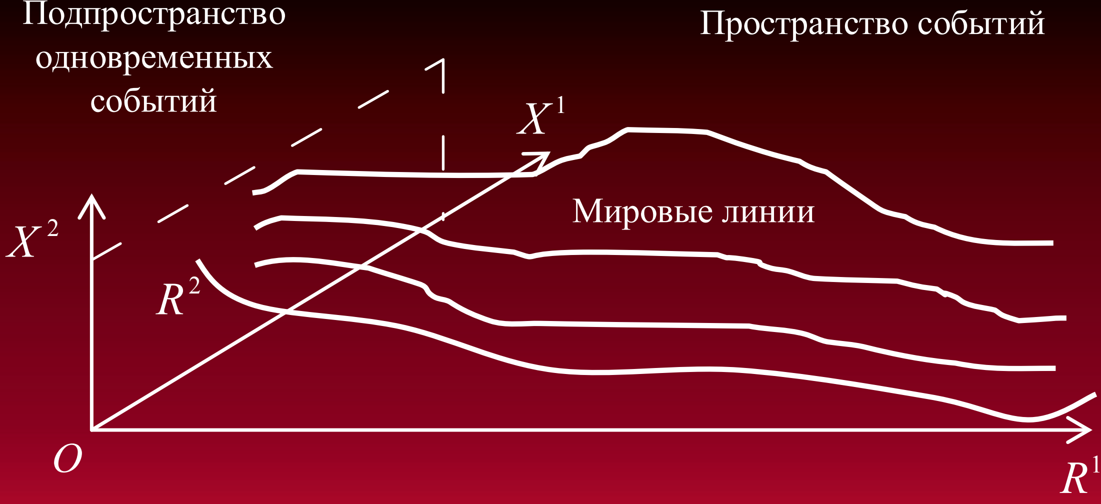
Уральский федеральный университет
Физико-технологический институт

В. Б. СУРНЕВ

ВЫСШАЯ МАТЕМАТИКА ДЛЯ ФИЗИКОВ. ЛИНЕЙНАЯ АЛГЕБРА

Учебное пособие



Министерство науки и высшего образования
Российской Федерации

Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина

В. Б. Сурнев

ВЫСШАЯ МАТЕМАТИКА ДЛЯ ФИЗИКОВ. ЛИНЕЙНАЯ АЛГЕБРА

Учебное пособие

Рекомендовано методическим советом
Уральского федерального университета
для студентов вуза, обучающихся
по направлению подготовки
14.03.02 «Ядерные физика и технологии»

Екатеринбург
Издательство Уральского университета

УДК 512.64(075.8)
ББЛ 22.14я73
С90

Рецензенты:

кафедра информатики ФГБОУ ВО УГГУ, завкафедрой канд. техн. наук, доцент *А. В. Дружинин*;

д-р физ.-мат. наук, вед. науч. сотр. института геофизики им. Ю. П. Булашевича УрО РАН *А. Ф. Шестаков*

Научный редактор — проф., д-р физ.-мат. наук *В. К. Першин*

Сурнев, В. Б.

С90 Высшая математика для физиков. Линейная алгебра : учебное пособие / В. Б. Сурнев ; М-во науки и высш. образования РФ. — Екатеринбург : Изд-во Урал. ун-та, — 436 с.

ISBN 978-5-7996-3148-2

Пособие охватывает все разделы дисциплины «Линейная алгебра», предусмотренные государственным стандартом направления подготовки бакалавров 14.03.02 «Ядерные физика и технологии». В каждой главе пособия подробно излагаются теоретические сведения и приводятся необходимые примеры решения типовых задач.

Предназначено для студентов направления подготовки 14.03.02 «Ядерные физика и технологии», а также для других направлений подготовки и специальностей физического профиля высших учебных заведений.

Библиогр.: 44 назв. Рис. 39.

УДК 512.64(075.8)
ББЛ 22.14я73

Учебное издание

Сурнев Виктор Борисович

ВЫСШАЯ МАТЕМАТИКА ДЛЯ ФИЗИКОВ. ЛИНЕЙНАЯ АЛГЕБРА

ISBN 978-5-7996-3148-2

Оглавление

Предисловие

Ни одна отрасль знания не может претендовать на право называться «научной», пока не сформулирует свои базовые понятия на языке математики.

Известно высказывание И. Канта: «В каждой естественной науке заключено столько истины, сколько в ней есть математики». Что касается связи физики и математики, то бытует мнение, что эти две науки неразрывно связаны и являются частями одной и той же физико-математической науки. Вынесенная в эпиграф сентенция является некоторым расширением приведённых мнений на другие науки и отражает личный взгляд автора на роль математики в научном прогрессе и её связь с другими науками.

Многолетний опыт работы в области теоретической геофизики и последующие более чем 25 лет преподавательской деятельности убедили автора в том, что многие понятия и выводы физики и техники становятся намного более понятными, если их преподавание основано на математике. Более того, в процессе преподавания физики некоторые её понятия и вовсе невозможно чётко сформулировать без применения математического языка. В качестве примера достаточно привести понятие сплошной гетерогенной среды, которое на словах можно сформулировать, например, так: «Гетерогенная среда — это среда, содержащая хаотически распределённые неоднородности с гладкими и резкими границами, заключённые в определённом объёме, ...». Однако, такое «определение» ни в коей мере не раскрывает физическое содержание понятия «гетерогенная среда», а только затуманивает и так довольно сложное понятие. Если же использовать математический язык, то определение гетерогенной среды можно выразить краткой и прозрачной фразой: «Гетерогенная среда — это сплошная среда, материальные параметры которой являются кусочно-непрерывными функциями, определёнными на подмножествах трёхмерного евклидова пространства».

Конечно, это определение понятно человеку, изучившему в достаточной мере математику. Кроме этого, требуются предварительное разъяснение других понятий, например, что такое материальный параметр среды. Тем не менее, можно утверждать, что определение физического понятия на языке математики резко упрощает понимание его сути. Таких примеров можно привести многие десятки и сотни. Таким образом, следует сделать вывод о том, что вынесенная в эпиграф сентенция отражает суть проблемы, а достаточно глубокое изучение математики совершенно необходимо студентам-физикам.

Вниманию читателя предлагается первая часть давно задуманного автором издания серии книг под общим названием «**Высшая математика для физиков**». Предполагается, что книги серии будут включать в себя материал лекций и практических занятий, которые автор проводил в течение 10 лет студентам специальности «Прикладная математика» Уральского государственного горного университета (УГГУ), а в последние несколько лет проводит для студентов направления подготовки «Ядерные физика и технологии» физико-технологического института (ФТИ) Уральского федерального университета (УрФУ), а также студентам направления подготовки «Технология геологической разведки» и специальности «Горное дело» УГГУ.

Путь к реализации задуманного был достаточно долгим и нелёгким. На этом пути автором за последние 20 лет было подготовлено и издано восемь учебных пособий [31–38], из которых три издания получили гриф УМО по образованию в области прикладной математики и управления качеством, а также несколько более мелких брошюр. Последним изданием на этом нелёгком пути стала книга автора «Математическое моделирование. Непрерывные детерминированные модели», получившая гриф УМО и изданная в 2013 году издательством Уральского государственного горного университета [38]. Последовавший за этим изданием пятилетний перерыв позволил автору восстановить силы, переосмыслить подходы к преподаванию математики для математиков-прикладников, физиков и инженеров-исследователей и подготовиться к написанию задуманной серии.

Несколько слов о целесообразности и своевременности такого издания. Если физик-экспериментатор или инженер-эксплуатационщик могут довольствоваться не слишком обширными и достаточно формальными познаниями в математике, то физик-теоретик, инженер-конструктор, инженер-исследователь и тем более специалист по математическому моделированию в области физики и техники должны не только обладать достаточно обширными познаниями в области математики, но и уметь творчески применять их для решения различных задач. Поэтому изучение достаточно полного курса высшей математики совершенно необходимо для будущих физиков-теоретиков и специалистов в области математического моделирования.

В недалёком прошлом было издано несколько математических курсов для подготовки студентов-физиков, например книги [7, 10, 24–29, 41]. Некоторые из этих книг написаны физиками «на физическом уровне строгости» [10] и вряд ли доставят читателю, нацеленному на работу в области теоретической физики и математического моделирования, необходимый объём математических знаний на соответствующем уровне строгости. Другие книги, например [40, 42], написаны математиками-теоретиками на уровне строгости, по моему мнению, избыточном для вчерашнего школьника, в силу чего малодоступны для регулярного изучения современному российскому студенту первого и второго курсов.

Наиболее близко по уровню строгости и доступности для специалистов по математическому моделированию и физике вообще написан известный курс высшей математики академика В. И. Смирнова [24–29]. Однако слишком большой объём, недостаточное количество или даже отсутствие физических примеров, архаичность изложения и терминологии в определённой степени осложняет использование данного издания для регулярного изучения математики названной категорией студентов и молодых специалистов. К тому же за прошедшие почти 60 лет со времени первого издания курса В. И. Смирнова значительно изменились подходы к преподаванию математических дисциплин, терминология и даже обозначения [16, 17].

К сказанному выше следует добавить, что за последние примерно 20 лет по объективным и субъективным причинам в арсенал специалиста по теоретической физике и математическому моделированию кроме общих разделов высшей математики, таких как «Линейная алгебра», «Математический анализ», «Обыкновенные дифференциальные уравнения», «Уравнения математической физики», «Теория функций комплексной переменной», «Теория вероятностей и математическая статистика», добавился и ряд других разделов, которые раньше преподавались в качестве специальных курсов. Назовём только такие разделы, как «Теория обобщённых функций», «Дифференциальная геометрия», «Интегральные уравнения и системы», «Функциональный анализ», «Вариационное исчисление» и некоторые другие разделы математики.

Кроме этого, вследствие далеко не самых лучших изменений ФГОС, выразившихся в разделении уровней подготовки (бакалавры и магистры), весьма нечётком понятии «компетенция», фактическом исключении из ФГОС понятия дисциплины и цикла дисциплин, произошло резкое сокращение времени, отпущенного на изучение математических и естественно-научных дисциплин бакалаврами.

Естественно, что подобные изменения в структуре образовательного процесса влекут за собой существенные изменения в структуре собственно курса математики, выражающиеся в компоновке материала, а именно: в последовательности изложения тем, в уровне математической строгости изложения и, наконец, в переработке порядка и уровня изложения внутри самих тем. Цель таких изменений — не теряя в строгости и полноте изложения, максимально сократить время на нужный объём подаваемых на аудиторных занятиях сведений. В качестве примера можно привести упор на теорию линейных операторов в линейной алгебре или совместное изложение дифференциального исчисления функций одной и нескольких независимых переменных в математическом анализе.

Кроме этого, в общий курс математики для физиков совершенно необходимо ввести дополнения, позволяющие студенту воочию увидеть необходимость изучения математических методов с точки зрения их применения в физике и показать, что физика и математика являются двумя сторонами одной физико-математической науки. Иными словами, студент должен проникнуться мыслью о том, что изучать физику без математики невозможно.

Стоит отметить, что введение физических дополнений в общий курс математики наталкивается на ряд существенных трудностей, связанных в первую очередь с тем, что во многих вузах и на факультетах физического профиля общая физика на первом семестре не преподаётся. С одной стороны, такая организация учебного процесса может быть обоснована как раз недостаточной математической подготовкой вчерашнего школьника, а с другой стороны, такое положение дел приводит к тому, что на первом семестре студент лишён «физической поддержки» при изучении математики. Если при изложении математического анализа, преподавание которого может быть основано на физических понятиях, отчасти известных ещё из школы, указанную трудность можно как-то нивелировать, то при изложении линейной алгебры вводимые понятия настолько новы для вчерашнего школьника, что указанная трудность «встаёт в полный рост». Например, проиллюстрировать первокурснику с точки зрения пока ещё неизвестной ему физики понятие квадратичной формы крайне затруднительно, так как впервые понятие квадратичной формы в полной мере проявится только при изучении механики, а именно при изучении кинетической энергии системы материальных точек в рамках механики Ньютона.

Предполагаемая серия книг является попыткой достаточно строгого, но не избыточно строгого, изложения общего курса высшей математики, ориентированного на подготовку специалистов в области физики и математического моделирования. Задуманная серия должна будет включать в себя книги, носящие следующие названия: «Линейная алгебра», «Дифференциальное и интегральное исчисление», «Дифференциальная геометрия и векторный анализ», «Обыкновенные дифференциальные уравнения и системы», «Теория функций комплексной переменной и уравнения математической физики», «Интегральные уравнения и элементы математического моделирования». Следует отметить, что кроме первых двух все названия условные и могут измениться в процессе издания.

Открывающая предполагаемую серию первая книга «**Линейная алгебра**» трактует входящую в учебный план физических и математических направлений подготовки одноимённую дисциплину. По своему содержанию и методике изложения данная книга опирается на учебное пособие автора [32], которое много лет использовалось для преподавания линейной алгебры студентам специальности «Прикладная математика» в Уральском государственном горном университете на факультете геологии и геофизики.

Естественно, что содержание книги подверглось значительной переработке и существенно дополнено. Во-первых, изменены, часто с целью упрощения, доказательства некоторых утверждений и теорем. Во-вторых, добавлено сравнительно большое число примеров с решением. В-третьих, исправлены опечатки и вычислительные ошибки, закравшиеся в примеры в книге [32] в достаточно большом числе. В-четвёртых, в книгу по возможности добавлено некоторое число сведений, иллюстрирующих применение теоретических положений линейной алгебры в физике. И наконец, самое существенное изменение состоит в том, что изменён формат книги, а именно, материал перегруппирован и размещён в семи главах в отличие от пяти глав книги [32]. По техническим причинам из книги пришлось исключить параграфы с практическими занятиями и заданиями для самостоятельной работы. Эти параграфы будут оформлены отдельно в виде практикума по линейной алгебре.

В силу описанных выше особенностей, сопровождающих преподавание высшей математики вообще и линейной алгебры в частности в вузе физического профиля, все имеющиеся в небольшом числе физические приложения и иллюстрации вынесены в конец книги, составляют содержание седьмой главы и доступны для освоения при параллельном изучении курса общей физики. Студентам, не изучающим общую физику в первом семестре, придётся отложить изучение физических приложений и до поры до времени довольствоваться изучением линейной алгебры в абстрактной форме. Впрочем, повторюсь, в данной книге таких приложений не слишком много.

В первой главе изложены самые начальные сведения из общей алгебры, которые являются своеобразным терминологическим фундаментом для всего дальнейшего изложения. В процессе изучения первой главы читатель получит достаточно полные сведения из наивной теории множеств и краткие сведения из теории алгебраических систем. Более подробно рассмотрены две алгебраические системы — поле действительных чисел и поле комплексных чисел.

Во второй главе подробно излагается теория конечномерных векторных пространств, которые являются как бы «ареной», на которой функционируют все вводимые далее объекты. Подробно изучается трёхмерное собственно евклидово пространство, являющееся основой аналитической геометрии. Понятие векторного пространства вводится в самом общем виде, достаточно полно изучается структура векторного пространства. В этой же главе изучаются аффинные и евклидовы пространства, имеющие основное значение для геометрических основ физики.

В третьей главе определяется понятие линейного оператора и подробно изучается структура линейных операторов в конечномерных пространствах. Рассматриваются элементы алгебры линейных операторов. Подробно изучаются достаточно сложные вопросы теории преобразования систем координат. Вводится понятие ранга оператора и его матрицы.

В четвёртой главе изучается задача на собственные значения линейного оператора и рассматриваются смежные вопросы теории собственных и инвариантных подпространств. Достаточно подробно изучаются сопряжённые, самосопряжённые и ортогональные операторы.

В пятой главе излагаются некоторые сведения из геометрии векторных пространств, в частности, строится решение некоторых типичных задач из геометрии многомерных собственно евклидовых пространств, изучается теория квадратичных форм. Завершается глава теорией поверхностей второго порядка в трёхмерном собственно евклидовом пространстве. По существу, пятая глава является квинтэссенцией предыдущих трёх глав.

В шестой главе рассматриваются некоторые алгебраические построения, не вошедшие в основной курс: кольцо многочленов, общая теория определителей, билинейные формы и их связь с квадратичными формами, каноническая (жорданова) форма матрицы линейного оператора, метрика в евклидовых пространствах.

В седьмой главе приведены некоторые простейшие приложения линейной алгебры к физике. Примеры взяты из классической и небесной механики, из теории электрических цепей и из квантовой механики. К сожалению, пришлось опустить возможные приложения в космологии.

Автор остался верен системе обозначений, принятой в книге [32]. Для векторов использованы обычные обозначения со стрелкой, операторы обозначаются большой буквой со «шляпкой». Для вектор-столбцов и вектор-строк в координатном пространстве использованы «bracket» обозначения П. А. М. Дирака [8], например, ket-вектор $|y\rangle$ и bra-вектор $\langle x|$ соответственно. Символ «*» используется для обозначения абстрактной алгебраической операции. Для конкретной операции, например сложения или умножения, используются символы «+», «×», и так далее.

Для координат векторов и элементов матриц применены обозначения с верхними и нижними индексами, что, возможно, несколько непривычно. Дело в том, что векторы имеют координаты двух типов — *контравариантные* и *ковариантные*: контравариантные координаты являются *первичными* и появляются в аффинном пространстве — для них используются обозначения с верхними индексами; ковариантные координаты появляются только в евклидовых пространствах [28]. Кроме этого, оправданием использования обозначений с верхними и нижними индексами служит не очевидное сразу удобство таких обозначений. Для записи различных утверждений и выводов широко используются простейшие логические обозначения.

Как было сказано выше, материал книги разбит на семь глав, которые разбиты в свою очередь на параграфы, а последние на пункты, не имеющие нумерации: ссылку на параграф под номером 3.2 следует понимать в том смысле, что имеется в виду параграф 2 главы 3. Нумерация определений, теорем и формул сквозная для каждого параграфа: ссылка на определение (теорему, формулу) 3.2.1 означает, что имеется в виду определение (теорема, формула) 1 из параграфа 2 главы 3. Символом «•» обозначается окончание определения; символом «••» обозначается окончание доказательства теоремы, леммы или утверждения; символом «⊗» обозначается окончание примера.

Некоторые параграфы и пункты отмечены символом «*». Это означает, что при первом изучении материал главы, параграфа (пункта) без ущерба для понимания последующего материала можно опустить. Впрочем, таких случаев немного.

Доказательства некоторых теорем набраны мелким шрифтом. Как правило, это сделано в случае, когда автор не смог предоставить более простых доказательств и вынужденно привёл довольно громоздкие «стандартные доказательства», известные из литературы. Эти доказательства предлагаются для ознакомления.

Некоторые разделы линейной алгебры автором намеренно приведены в упрощённом изложении, например, теория определителей изложена в эвристической (делай как сказано) форме, а квадратичные формы рассматриваются только в собственно евклидовом пространстве. Общая теория определителей, теория билинейных и квадратичных форм в абстрактных векторных пространствах и теория вещественных евклидовых пространств вынесены в дополнения. В данной книге не нашлось места тензорной алгебре. По соображениям удобства и полноты решено перенести тензорную алгебру и изложить её как часть тензорного исчисления в более подходящем месте курса.

В начале педагогической деятельности при подготовке лекционных курсов и практических занятий автор пользовался большим числом учебников и задачников, отмеченных в списке литературы. Большинство методических приёмов из упомянутых изданий естественным образом использованы и в данном курсе. Как уже отмечалось выше, в целях приблизить содержание книги к нуждам студентов физических направлений подготовки переработана методика изложения материала. Разумеется, ответственность за форму изложения материала лежит полностью на авторе.

Автор надеется, что представляемая книга будет полезна для изучения высшей математики студентами физических направлений подготовки.

Глава 1. Некоторые сведения из общей алгебры

1.1. Элементы теории множеств

Числовая прямая. Из курса математики средней школы известно, что ***действительные*** (***вещественные***) *числа* можно представить в виде ***бесконечной десятичной дроби***, например 3,1415… или 2,0010…. И наоборот, всякая бесконечная десятичная дробь вида

$$\pm a_{-n}a_{-(n-1)}\ldots a_{-1}a_0,a_1a_2a_3\ldots a_n\ldots,$$

где $a_n\ (n=1,2,3,\ldots)$ — одна из цифр $0,1,2,\ldots,9$, есть некоторое действительное число. Иногда одно и то же действительное число можно задать в виде десятичной дроби по-разному. Например, $0{,}999\ldots=1{,}000\ldots$, и вообще

$$a_{-n}\ldots a_i999\ldots9{,}999\ldots=a_{-n}\ldots(a_i+1)000\ldots0{,}000\ldots,$$

если выполнено условие $a_i\neq 9$. За исключением этого случая представление действительного числа в виде бесконечной десятичной дроби единственно. Если десятичная дробь заканчивается нулями, то нули не пишут, например, $5{,}000\ldots=5$. Так удобнее.

Действительные числа наглядно можно представить в виде точек некоторой прямой линии, например, на рис. 1.1 — это прямая линия L. Выберем на L произвольную точку P и напишем около неё число 0. Отложим на прямой L вправо от точки P отрезок длиной r и у его конца напишем число r. Влево от точки P отложим отрезок той же длины r и у его конца напишем число $-r$. Отложенный отрезок называется ***масштабным отрезком***. Таким образом, каждому действительному числу мы поставили в соответствие некоторую точку прямой L: говорят, что ***число лежит в точке прямой*** L.

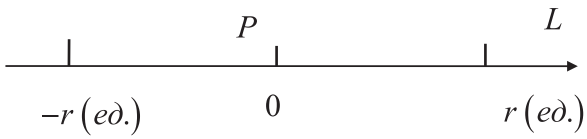


Рис. 1.1

Интуитивно ясно следующее: ***каждому действительному числу соответствует одна и только одна точка прямой L; двум различным числам всегда соответствуют две различные точки на прямой L; каждая точка прямой L соответствует, по крайней мере, одному действительному числу.***

Из этого следует, что каждая точка прямой линии L соответствует *в точности одному действительному числу*. Такое соответствие называется ***взаимно однозначным***, или ***биективным***. Можно сказать, что между точками прямой L и действительными числами описанным способом установлено ***биективное соответствие***. Полагая, что такое соответствие между действительными числами и точками некоторой прямой всегда имеет место, говорят о точке 0, точке 5, ... о точке 11 и т. д. Сама прямая линия, при таком соответствии, называется ***числовой прямой*** и является ***геометрическим образом совокупности всех действительных чисел***.

Понятие множества. Совокупность объектов можно рассматривать как новый объект, называемый в математике ***множество***. Составляющие множество ***объекты*** называются ***элементами множества***.

Пример 1.1.1. Множествами являются: совокупность студентов, присутствующих в данной аудитории; совокупность автомобилей, находящихся на данной автостоянке; совокупность товаров данного типа в ассортименте данного магазина; совокупность всех звёзд на небосклоне. ⊗

Определить понятие множества, по крайней мере, в рамках наивной теории, не представляется возможным, так как оно выражается через другое понятие — «***совокупность***», которое также требует определения. Невозможность определения понятия множества объективно обусловлена его «***изначальностью***»: множество — это самое общее, базовое понятие математики. Поэтому в наивной теории множеств понятие множества принято относить к ***первичным***, или ***неопределяемым*** понятиям.

Каждое множество из примера 1.1.1 состоит из вполне определённого ***конечного*** числа элементов, то есть в каждом из приведённых случаев мы на вопрос «***сколько?***» можем или прямо указать ***целое число*** элементов соответствующего множества, или указать, что такое целое число ***существует***, хотя в данный момент оно может быть нам и не известно. Такие множества называются ***конечными множествами***. В математике часто приходится иметь дело с «***неконечными***» или ***бесконечными*** множествами.

Пример 1.1.2. Бесконечными являются, например, следующие множества: множество всех натуральных чисел; множество всех чётных чисел; множество всех целых чисел, делящихся на 21; множество всех прямых, проходящих через заданную точку плоскости. $\otimes$

Вводится также понятие ***пустого множества***. Пустое множество, по определению, не содержит элементов. Понятие пустого множества весьма удобно, так как при изучении некоторого множества мы можем заранее и не знать, содержит ли оно хотя бы один элемент.

Пример 1.1.3. Множество страусов, находящихся в данный момент за Полярным кругом, вероятно, пусто. Хотя нельзя исключить, что какой-нибудь капитан завёз какого-нибудь страуса за Полярный круг (пример взят из книги П. С. Александрова «Введение в теорию множеств и общую топологию»).$\otimes$

Условимся обозначать произвольные множества большими, а их элементы — малыми буквами латинского алфавита. ***Множество действительных чисел*** принято, например, обозначать буквой R, а ***множество натуральных чисел*** — буквой N. Для пустого множества, в силу его особой роли, применяется специальное небуквенное обозначение $\varnothing$.

Отношения между элементами и множествами. Множество можно ***задать*** разными способами: перечислением всех его элементов, указанием характеристического признака элементов множества и т. д. Но всегда следует иметь в виду следующее соглашение: ***некоторое множество M задано в том и только в том случае, если для каждого элемента установлено, является он элементом множества M или нет.*** Пусть M — некоторое множество. Введём следующие основные ***символьные обозначения***:

$a \in M$, читается «*a* ***есть элемент множества*** *M*», или «*a* ***принадлежит множеству*** *M*», «*a* ***лежит во множестве*** *M*», «*a* ***является элементом из множества*** *M*»;

$a \notin M$, читается «*a* ***не есть элемент множества*** *M*», или «*a* ***не принадлежит множеству*** *M*», «*a* ***не лежит во множестве*** *M*», «*a* ***не является элементом из множества*** *M*»;

$M = \{a, b, \ldots\}$, читается «*M* ***есть множество, состоящее из элементов*** *a*, *b*, ***и так далее***»;

$M = \{a : a \text{ обладает свойством } E\}$, читается: «*M* ***есть множество всех элементов*** *a*, ***обладающих свойством*** *E*, ***и только эти элементы лежат во множестве*** *M*».

Пример 1.1.4. Множество всех натуральных чисел $N = \{1, 2, \ldots, n, \ldots\}$ с помощью символьных обозначений можно задать так:

$$N = \{n : 1 \in N;\, если\; n \in N,\, то\; и\;\; n+1 \in N\}.$$

Множество всех ***целых чисел*** $Z = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ с помощью символьных обозначений можно задать так:

$$Z = \{x : x \in N,\, или\, -x \in N,\, или\; x = 0\}. \otimes$$

Определение 1.1.1. *Два множества M_1 и M_2 называют* ***равными*** *и пишут $M_1 = M_2$ в том и только в том случае, если каждый элемент множества M_1 принадлежит и множеству M_2, а каждый элемент множества M_2 принадлежит и множеству M_1.* •

Таким образом, $M_1 = M_2$ в том и только в том случае, если они содержат тождественно совпадающие элементы.

Определение 1.1.2. *Пусть M_1 и M_2 — два множества. Множество M_1 называется* ***частью***, *или* ***подмножеством множества*** M_2 (***содержится во множестве*** M_2), *когда выполняется следующее условие: если $x \in M_1$, то и $x \in M_2$.* •

Иными словами, множество M_1 состоит из элементов, принадлежащих множеству M_2. В этом случае пишут $M_1 \subset M_2$. Символ $\subset$ называется ***символом включения***, но часто используется и обозначение $M_2 \supset M_1$ ($\supset$ — тоже символ включения).

Пример 1.1.5. 1) Множество студентов-юношей является подмножеством всех студентов, присутствующих в данной аудитории, причём может случиться так, что это множество будет пустым.

2) Множество $N_{2k} = \{n = 2k : k = 1, 2, 3, \ldots\}$ всех чётных натуральных чисел является подмножеством множества N всех натуральных чисел. $\otimes$

Принимается соглашение, что ***пустое множество*** $\varnothing$ ***является частью любого множества*** M, то есть $\varnothing \subset M$, где M — ***множество элементов любой природы***.

Определение 1.1.3. *Если N — часть множества M ($N \subset M$), то множество $C_M(N)$, состоящее из элементов, принадлежащих множеству M и не принадлежащих множеству N, то есть*

$$C_M(N) \overset{def}{=} \{x \in M : x \notin N\},$$

называется ***дополнением*** *множества N до множества M.* •

Пример 1.1.6. 1) Множество студентов-девушек является дополнением множества студентов-юношей до множества всех студентов, присутствующих в данной аудитории, причём может случиться так, что это дополнение будет пустым множеством.

2) Множество

$$N_{2k-1} = \{n = 2k - 1 : k = 1, 2, 3, \ldots\}$$

всех нечётных натуральных чисел является дополнением множества N_{2k} всех чётных натуральных чисел до множества N всех натуральных чисел:

$$N_{2k-1} = C_N(N_{2k}). \bullet$$

Определение 1.1.4. *Пусть M_1 и M_2 — два произвольных множества. Говорят, что между элементами этих множеств установлено* ***взаимно однозначное,*** *или* ***биективное, соответствие****, если:*

1) *каждому элементу $x \in M_1$ соответствует один и только один элемент $y \in M_2$;*

2) *каждым двум различным элементам $x_1, x_2 \in M_1$ соответствуют два различных элемента $y_1, y_2 \in M_2$;*

3) *любой элемент $y \in M_2$ соответствует, по крайней мере, одному элементу $x \in M_1$.* •

Пример 1.1.7. а) Пусть, например,

$$M_1 = \{1, 2, 3, 4, 5, 6, 7, 8\},$$

$$M_2 = \{a, b, c, d, e, f, g, h\},$$

$$M_3 = \{\alpha, \beta, \gamma, \lambda\}.$$

Первые два множества состоят из одного и того же конечного числа элементов и биективное соответствие между их элементами можно установить, например, так:

$$1 \leftrightarrow a,\ 2 \leftrightarrow b,\ 3 \leftrightarrow c,\ 4 \leftrightarrow d,\ 5 \leftrightarrow e,\ 6 \leftrightarrow f,\ 7 \leftrightarrow g,\ 8 \leftrightarrow h,$$

где двойная стрелка обозначает установленное соответствие. В свою очередь, в силу неравенства числа элементов первого и третьего множеств, биективное соответствие можно установить между всеми элементами третьего множества и частью первого множества, например, так:

$$\alpha \leftrightarrow 1,\ \beta \leftrightarrow 2,\ \gamma \leftrightarrow 3,\ \lambda \leftrightarrow 4.$$

б) Множества Z^+ всех целых положительных чисел и Z^- всех целых отрицательных чисел, очевидно, бесконечны. Биективное соответствие между элементами Z^+ и Z^- устанавливается, например, по следующему правилу: каждому элементу $c \in Z^+$ ставится в соответствие элемент $-c \in Z^-$. ⊗

Определение 1.1.5. *Если между элементами множеств M_1 и M_2 установлено биективное соответствие, то эти множества называются* ***количественно эквивалентными****, или просто* ***эквивалентными****.* •

В отношении эквивалентных множеств говорят также, что они ***имеют одинаковую мощность (равномощны)***. Равномощными являются и множества Z^+ и Z^-. Мощность — ***это то общее, что есть у эквивалентных между собой множеств***. Очевидно, что для конечных множеств это «общее» есть количество элементов во множествах: мощность конечного множества — это количество его элементов. Для бесконечных множеств понятие мощности является аналогом понятия количества элементов. Заметим, что из определения 1.5 вытекает следующее утверждение: *если множества M_1 и M_2 эквивалентны третьему множеству M, то они эквивалентны между собой.*

Определение 1.1.6. *Множество $M \neq \varnothing$ называется* ***счётным****, если оно равномощно множеству натуральных чисел N.* •

Очевидно, определение счётного множества можно сформулировать так: непустое множество *M* называется ***счётным***, если все его элементы можно занумеровать и расположить в виде бесконечной последовательности вида $a_1, a_2, a_3, \ldots, a_n, \ldots$ так, чтобы каждый элемент получил единственный номер $n \in N$, а каждое натуральное число *n* было бы присвоено в качестве номера единственному элементу множества *M*.

Определение 1.1.7. *Бесконечное множество $M \neq \varnothing$ (непустое), не являющееся счётным, называется **несчётным** множеством.* •

В математическом анализе доказывается, что множество R^1 действительных чисел несчётно. Точнее, доказывается утверждение: *открытый промежуток $(a, b) \subset R^1$ несчётен.* Отсюда очевидным образом следует такой факт.

Утверждение 1.1. *Бесконечное множество $M \neq \varnothing$, эквивалентное открытому промежутку $(a, b) \subset R^1$, несчётно.*

Определение 1.1.8. *Пусть даны два непустых множества M_1 и M_2. Если по определённому правилу каждому элементу $x \in M_1$ поставлен в соответствие вполне определённый элемент $y \in M_2$, то говорят, что задано **отображение** f **множества** M_1 **во множество** M_2, и пишут $f : M_1 \to M_2$.* •

Об отображении часто говорят как о ***функции (абстрактной)*** f, заданной на множестве M_1 со значениями во множестве M_2, причём M_1 называется ***множеством определения*** функции, а M_2 — ***множеством её значений***. Часто используется ***префиксная*** запись функции: для каждого элемента $x \in M_1$ выполнено условие $y = f(x) \in f(M_1) \equiv M_2$. Здесь введено обозначение $f(M_1)$ для множества значений функции f.

Заметим, что если $M_1 = M_2 = M$, то отображение $f : M \to M$ называется ***преобразованием***.

Отображение $f : M_1 \to M_2$ называется ***инъективным***, если для любых двух элементов $x_1, x_2 \in M_1$, таких, что $x_1 \neq x_2$, неравны значения отображения на этих элементах, то есть $f(x_1) \neq f(x_2)$.

Отображение $f : M_1 \to M_2$ называется ***сюръективным*** отображением, если $f(M_1) = M_2$, то есть для любого элемента $\forall y \in M_2$ найдётся элемент $x \in M_1$, такой, что $f(x) = y$.

Отображение $f : M_1 \to f(M) \equiv M_2$ называется ***биективным***, если оно одновременно инъективно и сюръективно.

Операции над множествами. Закон тождества. Определим теперь операции над множествами.

Определение 1.1.9. ***Пересечением*** *множеств M_1 и M_2 называется множество*

$$M_1 \cap M_2 \overset{def}{=} \{x : x \in M_1 \text{ и } x \in M_2\},$$

то есть множество всех элементов, принадлежащих как множеству M_1, так и множеству M_2. •

Если у M_1 и M_2 нет общих элементов, то полагают, что

$$M_1 \cap M_2 = \varnothing,$$

и говорят, что эти множества ***не пересекаются***.

Операция пересечения множеств обобщается на любое конечное число множеств. Пересечение любого конечного числа множеств $M_1, M_2, \ldots, M_n$ определяется так:

$$M_1 \cap M_2 \cap \ldots \cap M_n \overset{def}{=} \{x : x \in M_1, \text{ и } x \in M_2, \text{ и } \ldots, \text{ и } x \in M_n\}.$$

Определение 1.1.10. ***Объединением*** *множеств M_1 и M_2 называется множество*

$$M_1 \cup M_2 \overset{def}{=} \{x : x \in M_1 \text{ или } x \in M_2\},$$

то есть множество элементов, принадлежащих хотя бы одному из двух данных множеств, включая элементы, принадлежащие им обоим. •

Объединение любого конечного числа множества $M_1, M_2, \ldots, M_n$ определяется так:

$$M_1 \cup M_2 \cup \ldots \cup M_n \overset{def}{=} \{x : x \in M_1, \text{ или } x \in M_2, \text{ или } \ldots, \text{ или } x \in M_n, \ldots\}.$$

Операция объединения множеств обобщается на любое число множеств.

Пример 1.1.8. Пусть даны три множества:

$$M_1 = \{a, b, c, d, e, f, g, h\},$$

$$M_2 = \{a, c, d, e, g\},$$

$$M_3 = \{o, p, q\}.$$

Очевидно, что справедливы следующие отношения:

$$M_1 \cap M_2 = \{a, c, d, e, g\},\ M_1 \cap M_3 = \varnothing,\ M_2 \cap M_3 = \varnothing;$$

$$M_1 \cup M_2 = \{a, b, c, d, e, f, g, h\} = M_1;$$

$$M_1 \cup M_3 = \{a, b, c, d, e, f, g, h, o, p, q\};$$

$$M_2 \cup M_3 = \{a, c, d, e, g, o, p, q\}. \otimes$$

Определение 1.1.11. *Пусть M — некоторое множество.* ***Разностью*** *его частей $M_1 \subset M$ и $M_2 \subset M$, причём возможно, что $M_1 \cap M_2 = \varnothing$, называется множество $M_1 - M_2 = \{x : x \in M_1$ и $x \notin M_2\}$.•*

Для разности множеств используется также обозначение с наклонной чертой:

$$M_1 - M_2 \equiv M_1 \setminus M_2.$$

Пример 1.1.9. Для множеств из предыдущего примера имеем:

$$M_1 - M_2 = \{b, f, h\};$$

$$M_1 - M_3 = \{a, b, c, d, e, f, g, h\} = M_1;$$

$$M_2 - M_3 = \{a, c, d, e, g\} = M_2. \otimes$$

Определение 1.1.12. *Множество $\{x, y\}$, состоящее из двух элементов x и y, называется* ***упорядоченной парой,*** *если указано, какой элемент считается первым, а какой вторым.•*

Определение 1.1.13. ***Произведением*** *$M_1 \times M_2$ двух множеств M_1 и M_2 называется множество всевозможных упорядоченных пар $\{x, y\}$, образованных из элементов $x \in M_1$ и $y \in M_2$, причём пары $\{x, y\}$ и $\{y, x\}$, с не совпадающими тождественно элементами x и y, считаются различными.•*

Вместо термина «***произведение***» часто используются термины «***декартово произведение***» или «***прямое произведение***».

Пример 1.1.10. Каждая точка на декартовой плоскости задаётся её декартовыми координатами — действительными числами x и y, образующими упорядоченную пару. Поэтому вся декартова плоскость является множеством упорядоченных пар $\{x, y\}$ действительных чисел, то есть произведением двух множеств — множества точек оси OX и множества точек оси OY. $\otimes$

Следующая теорема широко применяется для доказательства различных утверждений теории множеств.

Теорема 1.1.1 (закон тождества). *Если M_1 и M_2 — два множества, для которых выполнены включения $M_1 \subset M_2$ и $M_2 \subset M_1$, то $M_1 = M_2$.*

Д о к а з а т е л ь с т в о. Если $x \in M_1$, то и $x \in M_2$, так как $M_1 \subset M_2$. Если, наоборот, $x \in M_2$, то и $x \in M_1$, так как $M_2 \subset M_1$. Итак, множества M_1 и M_2 содержат одни и те же элементы и, следовательно, равны. ••

Высказывания, предикаты и кванторы. Языковой формой суждений об отношениях между понятиями и элементами множеств является ***словесное выражение*** (***повествовательное предложение***), рассматриваемое с точки зрения ***истинности*** или ***ложности*** его значения, которое называется ***высказыванием***.

Не все предложения являются высказываниями. Например, предложение

$$x^2 + y^2 = 4$$

не есть высказывание, так как судить о его истинности или ложности нельзя. При подстановке конкретных значений переменных это предложение превращается в высказывание — истинное или ложное. Например, если

$$x = y = \sqrt{2},$$

то предложение

$$x^2 + y^2 = 4$$

превращается в истинное высказывание, а при $x = 3$, $y = 4$ — в ложное.

Пример 1.1.11. Найти множество истинности предиката $x^2 + 3x + 1 > 0$, $x \in R^1$.

Р е ш е н и е. Так как

$$x^2 + 3x + 1 = \left(x + \tfrac{3}{2}\right)^2 - \tfrac{5}{4},$$

то имеем параболу с вершиной в точке $\left(-\tfrac{3}{2}, -\tfrac{5}{4}\right)$, оси параболы направлены вверх. Точки пересечения с осью абсцисс:

$$x_{1,2} = \frac{-3 \mp \sqrt{5}}{2}.$$

Теперь очевидно, что множество истинности указанного предиката есть

$$M = \left(-\infty, \left(-3 - \sqrt{5}\right)/2\right) \cup \left(\left(-3 + \sqrt{5}\right)/2, +\infty\right). \otimes$$

Приведём значение основных ***логических символов*** (***кванторов***).

Знак всеобщности $\forall$: *читается «для любого», «для каждого», «для всех».*

Знак существования $\exists$: *читается «существует», «найдётся».*

Знак следования (***импликации***) $\Rightarrow$: *запись* $A \Rightarrow B$ *означает, что A влечёт B или B следует из A* (*B — необходимое условие или признак A, A — достаточное условие или признак B*).

Знак равносильности $\Leftrightarrow$: *запись* $A \Leftrightarrow B$ *означает, что B следует из A и A следует из B* (*A равносильно B; A необходимо и достаточно для B; A тогда и только тогда, когда B; A в том и только в том случае, если B*).

Знак дизъюнкции $\vee$: *заменяет союз «или»* (*запись* $A \vee B$ *означает, что имеет место хотя бы одно из высказываний A или B*).

Знак конъюнкции $\wedge$: *заменяет союз «и».*

Знак отрицания $\neg$: *запись* $\neg A$ *означает «не A»* (*отрицание высказывания A*).

Ниже для замены слов «***такой, что ...***» часто будем использовать символ двоеточия, который не является общепринятым.

Пример 1.1.12. Записать при помощи логических символов следующее предложение: некоторые нечётные числа делятся на 7.

Р е ш е н и е. В этом высказывании имеется *квантор существования*, выраженный словом «некоторые», и высказывание «нечётные числа делятся на 7», заданное на множестве X нечётных чисел x. Обозначая высказывание символом $A(x)$, запишем требуемое предложение в виде

$$(\exists x \in X) : A(x). \otimes$$

Бинарная алгебраическая операция. Среди всевозможных множеств можно выделить такие множества, над элементами которых возможно выполнение некоторых действий, в результате которых можно получить элемент того же множества, или элемент некоторого другого множества. Для дальнейшего изложения большое значение имеет понятие ***бинарной алгебраической операции***.

Определение 1.1.14. *Пусть дано некоторое непустое множество M. Говорят, что на множестве M* ***определена бинарная алгебраическая операция****, если указан закон* φ*, по которому каждой упорядоченной паре элементов* $\{x, y\} \in M \times M$ *однозначным образом ставится в соответствие вполне определённый третий элемент* $z \in M$. •

В соответствии с определением, бинарная алгебраическая операция — это отображение $\varphi : M \times M \to M$. Однако бинарная операция может быть определена и для элементов разных множеств, что также следует из определения 1.1.14. Например, бинарная операция $g : P \times M \to M$, где $M \cap P = \varnothing$, имеет весьма большое значение для всего дальнейшего изложения, хотя и не является алгебраической. По этой причине алгебраическую бинарную операцию иногда называют ***внутренней бинарной операцией***, а неалгебраическую — ***внешней бинарной операцией*** (по отношению к множеству *M*).

Далее мы будем изучать множества с определёнными на них исключительно бинарными операциями, поэтому вместо термина «***бинарная алгебраическая операция***» будем часто говорить просто ***алгебраическая операция***. Если операция называется ***сложением***, используется обозначение $z = x + y$; если — ***умножением***, используется обозначение $z = xy$. Для обозначения абстрактной алгебраической операции используются также символы: $*, \circ, \bullet, \times, \perp, \ldots$.

Алгебраические операции по своим свойствам подразделяются на типы.

Определение 1.1.15. *Алгебраическая операция* $(*)$ *называется* ***коммутативной****, если результат её выполнения не зависит от порядка следования элементов, то есть если*

$\left(\forall x, y \in M\right) x * y = y * x.$

В противном случае операция называется ***некоммутативной***. •

Пример 1.1.13. Обычные арифметические операции над действительными числами — ***сложение*** и ***умножение*** — являются алгебраическими, причём они коммутативны. Операции деления и вычитания действительных чисел, являясь алгебраическими, что легко проверить, ***некоммутативны***. ⊗

Определение 1.1.16. *Алгебраическая операция* $(*)$*, заданная на некотором множестве M, называется* ***ассоциативной****, если*

$$\left(\forall x, y, z \in M\right) x * \left(y * z\right) = \left(x * y\right) * z. \bullet$$

В силу определения имеем

$$x * \left(y * z\right) = \left(x * y\right) * z = x * y * z.$$

В случае ассоциативной алгебраической операции значением выражения $x_1 * x_2 * \ldots * x_n$ при любом конечном числе элементов $x_1, x_2, \ldots, x_n$ будет результат расстановки скобок в данном выражении произвольным образом, при условии, что его можно определить путём последовательного применения алгебраической операции к парам элементов. В выражении $x_1 * x_2 * x_3 * x_4 * x_5$ можно расставить скобки, например, так

$$x_1 * \left(\left(x_2 * x_3\right) * \left(x_4 * x_5\right)\right), \left(\left(x_1 * x_2\right) * x_3\right) * \left(x_4 * x_5\right), \ldots.$$

Оформим высказанные соображения в виде теоремы.

Теорема 1.1.2. *Для* ***ассоциативной*** *алгебраической операции результат вычислений не зависит от способа расстановки скобок.*

Доказательство. Применим для доказательства метод математической индукции. Для $n = 3$ это утверждение справедливо по определению ассоциативности. При $n > 3$ считаем, что $(\forall k < n)$ утверждение справедливо. При расстановке скобок в выражении $a_1 * a_2 * \ldots * a_n$ всегда последним шагом будет выполнение операции над двумя элементами

$$a_1 * a_2 * \ldots * a_k,\ a_{k+1} * a_{k+2} * \ldots * a_n$$

для некоторого k, удовлетворяющего условию $3 < k < n, n > 3$. Так как оба выражения содержат меньше, чем n элементов, то они определены однозначно. Остаётся доказать, что для любых k и $l = n - k$ выполняется равенство

$$\left(a_1 * a_2 * \ldots * a_k\right) * \left(a_{k+1} * a_{k+2} * \ldots * a_n\right) = a_1 * a_2 * \ldots * a_n.$$

При $k = n - 1$ — равенство справедливо как следствие из определения ассоциативной операции. Если $k < n - 1$, то, применяя несколько раз ассоциативность, получаем:

$$\begin{gathered}\left(a_1 * a_2 * \ldots * a_k\right) * \left(a_{k+1} * a_{k+2} * \ldots * a_n\right) = \\ = \left(a_1 * a_2 * \ldots * a_k\right) * \left(a_{k+1} * \left(a_{k+2} * \ldots * a_n\right)\right) = \\ = \left(\left(a_1 * a_2 * \ldots * a_k\right) * a_{k+1}\right) * \left(a_{k+2} * \ldots * a_n\right) = \\ = \left(a_1 * a_2 * \ldots * a_k * a_{k+1}\right) * \left(a_{k+2} * \ldots * a_n\right).\end{gathered}$$

Если $k + 1 = n - 1$, то последняя строка равна $a_1 * a_2 * \ldots * a_n$. В противном случае повторяем действие до тех пор, пока в правой скобке останется a_n. ••

Определение 1.1.17. *Алгебраическая операция* $(\circ)$ *называется* ***дистрибутивной*** *относительно алгебраической операции* $(*)$, *если* $(\forall\, x, y, z \in M)$

1) $(x * y) \circ z = (x \circ z) * (y \circ z)$,

2) $z \circ (x * y) = (z \circ x) * (z \circ y)$. •

Если в определении 1.1.17 выполняется только равенство 1, то операция $(\circ)$ называется ***дистрибутивной справа*** относительно операции $(*)$. Если выполняется только равенство 2, то операция $(\circ)$ называется ***дистрибутивной слева*** относительно операции $(*)$. Дистрибутивность справа и слева не различаются, если операция $(\circ)$ коммутативна.

Определение 1.1.18. *Алгебраическая операция* $(*)$, *заданная на некотором множестве X, называется* ***сократимой***, *если для некоторого фиксированного элемента* $a \in X$ *и любых элементов* $x, y \in X$ *из условий*

$$a * x = a * y;\;\; x * a = y * a$$

следует, что $x = y$.•

Сократима операция сложения действительных чисел, так как из равенств

$$a + x = a + y;\;\; x + a = y + a$$

следует, что $x = y$.

Определение 1.1.19. *Пусть* $(*)$ *— сократимая коммутативная операция, заданная на множестве X. Операция* $(\circ)$ *называется* ***обратной*** *для операции* $(*)$, *если* $z = x \circ y$ *в том и только в том случае, когда* $x = z * y$.•

Операция вычитания на множестве действительных чисел является обратной для операции сложения, действительно: $z = x - y \Leftrightarrow x = y + z$.

Понятие бинарного отношения. При решении различных задач часто возникает необходимость иметь способ сравнения элементов множеств M_1 и M_2 (или множества M) между собой относительно некоторого признака. Причём, если сравниваются элементы одного множества, последнее разбивают на ***группы*** (подмножества) элементов, объединяя их по данному признаку. Если при таком разбиении полученные подмножества элементов не пересекаются, а их объединение составляет само множество, то говорят о разбиении множества на ***непересекающиеся группы*** или ***классы***.

Признаки, по которым элементы множества подразделяются на классы, могут быть различными, но не могут быть произвольными. Пусть, например, мы хотим разбить на классы множество действительных чисел R^1, помещая в один класс элементы a и b в том и только в том случае, если $b > a$. Тогда ни одно число a не может попасть в один класс с самим собой, так как a не больше, чем a. Никакого разбиения на классы не получается.

Определение 1.1.20. *Говорят, что между элементами x и y множеств M_1 и M_2 установлено* ***соответствие****, если в прямом произведении $M_1 \times M_2$ задано некоторое подмножество $\Re \subset M_1 \times M_2$. Если соответствие $\Re$ установлено между элементами одного и того же множества M, то есть $\Re \subset M \times M$, то соответствие называется* ***бинарным отношением***. •

Буква готического алфавита $\Re$ читается ***raida*** (райда), или ***reda*** (реда).

Задание соответствия $\Re$ равносильно заданию множества всех упорядоченных пар $\{x, y\}$ элементов $x \in M_1$ и $y \in M_2$, таких, что

$$\{x, y\} \in \Re \subset M_1 \times M_2.$$

Для упорядоченной пары $\{x, y\}$ элементов $x \in M_1$ и $y \in M_2$, удовлетворяющих этому условию, пишут $x \Re y$ и говорят, ***что элемент** y **поставлен в соответствие** $\Re$ **элементу** x*. Отметим, что задание соответствия $\Re \subset M_1 \times M_2$ фактически равносильно заданию некоторого ***отображения*** $F : M_1 \to M_2$, то есть понятие соответствия включает в себя понятие отображения. При этом множество M_1, как и в определении 1.1.20, называется ***множеством определения соответствия (или бинарного отношения)*** $\Re$, а множество M_2 — ***множеством его значений***. В случае бинарного отношения $\Re \subset M \times M \equiv M^2$ реализуется, очевидно, ***преобразование*** $f : M \to M$. Дальше сосредоточим внимание на бинарных отношениях.

Пример 1.1.14. Пусть дано множество $M = \{2, 3, 4, 6, 8, 9\}$. Рассмотрим на этом множестве отношение «*быть кратным*». Для этого объединяем элементы множества в пары по признаку кратности:

$$\{2, 2\}, \{3, 3\}, \{4, 4\}, \{6, 6\}, \{8, 8\}, \{9, 9\};$$

$$\{4, 2\}, \{6, 2\}, \{8, 2\};$$

$$\{6, 3\}, \{9, 3\}, \{8, 4\}.$$

Получили некоторое подмножество множества $M \times M$. Изобразим данное отношение графически. Для этого элементам множества M поставим в соответствие точки на плоскости, а высказыванию «быть кратным» поставим в соответствие стрелку, соединяющую элементы множества (рис. 1.2).

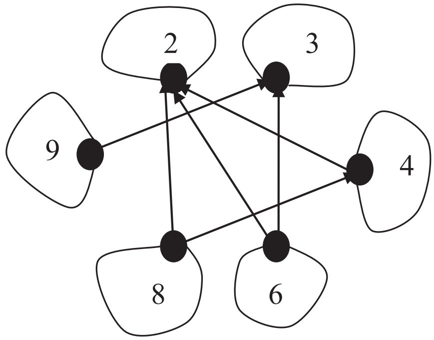


Рис. 1.2

Данное графическое изображение называется ***графом***. Наличие на графе петель, соединяющих элементы с собой, вызвано тем, что каждое число кратно самому себе. ⊗

Отношение эквивалентности. Особое значение имеет ***отношение эквивалентности***, которое является широким обобщением отношения равенства.

Определение 1.1.21. *Бинарное отношение $\Re$, определённое на множестве M, называется* ***отношением эквивалентности****, если оно обладает следующими тремя свойствами:*

a) $(\forall x \in M)\{x, x\} \in \Re$ — ***рефлексивность;***

b) $\{x, y\} \in \Re \Rightarrow \{y, x\} \in \Re$ — ***симметричность;***

c) $\{x, y\} \in \Re$ *и* $\{y, z\} \in \Re \Rightarrow \{x, z\} \in \Re$ — ***транзитивность***.•

Вместо $\{x, y\} \in \Re$ пишут $x \underset{\Re}{\approx} y$ или $x \equiv y(\mathrm{mod}\,\Re)$ (читается «*x* ***конгруэнтно y по модулю*** $\Re$»), или ещё проще $x \approx y$ и $x \equiv y$, если нет необходимости указывать, что речь идёт об одном и том же отношении $\Re$.

Определение 1.1.22. *Пусть для элементов произвольного множества M определено некоторое отношение эквивалентности $\Re$. Часть множества M, образованная из всех эквивалентных между собой элементов, будем называть* ***классом эквивалентности***. •

Пример 1.1.15. Рассмотрим обычную плоскость с известной из школы декартовой прямоугольной системой координат. Пусть на плоскости изображена система концентрических окружностей с центром в начале системы координат и радиусами R_k (рис. 1.3). За бинарное отношение (признак) примем свойство точек плоскости принадлежать окружности с определённым радиусом.

Вполне очевидно, что это отношение является отношением эквивалентности (свойств *a*, *b* и *c* проверяются элементарно). Классами эквивалентности как раз и являются заданные окружности.⊗

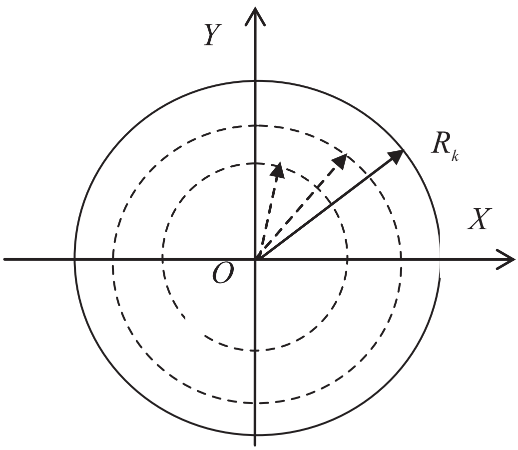


Рис. 1.3

Теорема 1.1.3. *Два класса эквивалентности всегда либо совпадают, либо не пересекаются.*

Доказательство. Пусть K_a и K_b — два класса эквивалентности на множестве M. Если они не пересекаются, доказывать нечего. Поэтому предположим, что K_a и K_b пересекаются и c — некоторый их общий элемент. Пусть $a \in K_a$ и $b \in K_b$ — некоторые элементы из соответствующих классов эквивалентности. Так как $c \in K_a \cap K_b$, то $c \approx a$ и $c \approx b$. В силу симметрии отношения эквивалентности (свойство 2) имеем $a \approx c$, а в силу транзитивности (свойство 3) имеем $a \approx b$ и, конечно, $b \approx a$. Пусть теперь $x \in K_a$, тогда $x \approx a$ и, следовательно, $x \approx b$, то есть $x \in K_b$. Аналогично, если $x \in K_b$, то $x \in K_a$. Таким образом, два класса эквивалентности, имеющие хотя бы один общий элемент, совпадают. ••

Отношение порядка. Часто возникает вопрос, в каком отношении между собой находятся неэквивалентные элементы множества?

Определение 1.1.23. *Бинарное отношение* $\Re$ *на множестве M называется **отношением порядка**, если оно обладает следующими свойствами:* $(\forall x, y, z \in M)$

a) $\{x, x\} \in \Re$ — ***рефлексивность;***

b) $\{x, y\} \in \Re \wedge \{y, z\} \in \Re \Rightarrow \{x, z\} \in \Re$ — ***транзитивность;***

c) $\{x, y\} \in \Re \wedge \{y, x\} \in \Re \Rightarrow x \equiv y$ — ***антисимметричность***. •

Если отношение порядка $\Re$ было определено заранее, то вместо $\{x, y\} \in \Re$ пишут $x \underset{\Re}{\leq} y$, или $x \leq y$. Запись $x \leq y$ читается: «*x **предшествует** y*», или «*y **следует за** x*». При таких обозначениях $y \geq x \Leftrightarrow x \leq y$. Теперь определение отношения порядка записывается так: $(\forall x, y, z \in M)$

a) $x \leq x$ — ***рефлексивность***;

b) $x \leq y$ *и* $y \leq z \Rightarrow x \leq z$ — ***транзитивность;***

c) $x \leq y \wedge y \leq x \Rightarrow x = y$ — ***антисимметричность.***

Пример 1.1.16. Отношение порядка $\Re$ во множествах N натуральных чисел, Z всех целых чисел, Q рациональных чисел и R^1 действительных чисел вводится высказыванием: «$\{x, y\} \in \Re$ ***в том и только в том случае, если*** $x \leq y$ ». ⊗

Пример 1.1.17. Пусть a и b произвольные действительные числа. Тогда или $a < b$, или $b < a$, или $a = b$; имеет место в точности одно из трёх этих отношений.

Пусть $a < b$. Тогда множество

$$I = (a, b) = \{x : a < x < b\},$$

где $a < x < b \Leftrightarrow a < x \wedge x < b$, называется ***открытым промежутком***, или ***интервалом*** *с концами a и b*. Присоединяя к промежутку (a, b) концы a и b, получаем ***замкнутый промежуток***, или ***сегмент***

$$\bar{I} = [a, b] = \{x : a < x < b \vee x = a \vee x = b\}.$$

Рассматриваются также и ***полуоткрытые промежутки*** числовой прямой:

$$[a, b) = \{x : x = a \vee a < x < b\};\ (a, b] = \{x : x = b \vee a < x < b\}. \otimes$$

1.2. Алгебраические системы

Математические объекты, представляющие интерес с точки зрения физики, естествознания, экономики, техники и так далее, являются элементами множеств, которые оснащены внутренними и внешними операциями. Такие множества называются ***алгебраическими системами***. Первоначальными в теории алгебраических систем являются понятия ***группоида*** и ***полугруппы***.

Пара $\{M, *\}$, где M — множество, а $(*)$ — заданная на нём бинарная алгебраическая операция, называется ***группоидом***. Если операция $(*)$ ассоциативная, то группоид называется ***полугруппой***.

Множества с одной алгебраической операцией, понятие группы. Простейшей, но весьма важной для приложений, алгебраической системой является ***группа***. Сформулируем определение группы, используя логическую символику.

Определение 1.2.1. *Множество $G \neq \varnothing$ с заданной на нём бинарной алгебраической (внутренней) операцией $(*)$, называется* ***группой****, если выполнены следующие аксиомы:*

1) $(\forall\, x, y, z \in G)\; x*(y*z)=(x*y)*z$;
2) $(\exists\, e \in G):(\forall\, x \in G)\; x*e=e*x=x$;
3) $(\forall\, x \in G)(\exists\, x^{-1} \in G): x*x^{-1}=x^{-1}*x=e$. •

Бинарная операция, определяющая группу, называется ***групповой операцией***, причём группу обозначают $\{G, *\}$ или $(G, *)$. Отметим, что смысл аксиом состоит в следующем: первая аксиома постулирует ***ассоциативность групповой операции*** $(*)$, вторая — наличие в группе ***единичного элемента***, третья — наличие в группе ***обратного элемента***.

Из аксиом группы можно вывести некоторые простые следствия.

Следствие 1. *Единичный элемент $e \in G$ единственен.*

Доказательство. Действительно, пусть $e_1 \in G$ и $e_2 \in G$ — два единичных элемента. Тогда $e_1 = e_1 * e_2 = e_2$. ••

Следствие 2. *Обратный элемент $x^{-1} \in G$ единственен.*

Доказательство. Действительно, пусть $x' \in G$ и $x'' \in G$ — два обратных элемента. Тогда $(\forall x \in G)$

$$x' = x' * e = x' * (x * x'') = (x' * x) * x'' = e * x'' = x''. \bullet\bullet$$

Следствие 3. *Для любых $a, b \in G$ уравнение $x * a = b$ имеет единственное решение, обозначаемое $x = b * a^{-1}$.*

Доказательство. Пусть c — некоторое решение указанного уравнения. Тогда, используя аксиомы группы, получаем:

$$c * a = b \Rightarrow (c * a) * a^{-1} = b * a^{-1} \Rightarrow c * (a * a^{-1}) = b * a^{-1} \Rightarrow$$

$$\Rightarrow c * e = b * a^{-1} \Rightarrow c = b * a^{-1}. \bullet\bullet$$

Группа называется ***коммутативной***, или ***абелевой***, если групповая операция коммутативна. В этом случае групповая операция обозначается символом $(+)$ и называется ***сложением***. Если групповая операция некоммутативная, то она обычно называется ***умножением*** и обозначается как обычное умножение точкой $(\cdot)$, или, например, просто xy, а сама группа называется ***некоммутативной***. Если групповой операцией является сложение, то группа называется ***аддитивной***, если же групповой операцией является умножение — ***мультипликативной***.

Может показаться, что вторая и третья аксиомы группы выполняются только для коммутативной группы. Покажем, что это не так. Действительно, пусть в группе имеются два элемента — e' и e'', удовлетворяющие условиям $(\forall x \in G)\ e' * x = x$ и $x * e'' = x$ — так называемые ***левый*** и ***правый единичные элементы***. Нетрудно видеть, что эти элементы совпадают, что следует из равенств: $e' * e'' = e''$, $e' * e'' = e'$.

Пусть теперь в группе имеются ***левый*** и ***правый обратные элементы*** x' и x'', удовлетворяющие условиям: $x' * x = e$ и $x * x'' = e$. Совпадение этих двух элементов вытекает из сравнения элемента $x' * x * x''$, вычисленного двумя способами:

$$x' * x * x'' = x' * (x * x'') = x' * e = x';$$

$$x' * x * x'' = (x' * x) * x'' = e * x'' = x''.$$

Сформулируем определения аддитивной и мультипликативной абелевых групп.

Определение 1.2.2. ***Аддитивной абелевой группой*** *называется непустое множество A с заданной на нём операцией* ***сложения****, удовлетворяющей следующим аксиомам:*

1) $(\forall x, y \in A)\ x + y = y + x$ ***(коммутативность);***

2) $(\forall x, y, z \in A)\ (x + y) + z = x + (y + z)$ ***(ассоциативность);***

3) $(\exists 0 \in A) : (\forall x \in A)\ x + 0 = x$ ***(существование нулевого элемента);***

4) $(\forall x \in A)\ (\exists (-x) \in A) :\ x + (-x) = 0$ ***(существование противоположного элемента).*** •

Таким образом, в случае аддитивной абелевой группы единичный элемент называется ***нулевым элементом*** (или просто ***нулём группы***) и обозначается 0, а обратный элемент называется ***противоположным*** и обозначается $-x$. Нетрудно показать, что $(\forall a, b \in A)$ *уравнение* $x + a = b$ *имеет единственное решение, равное* $b + (-a)$ *и обозначаемое* $x = b - a$.

Действительно, пусть c — некоторое решение уравнения $x + a = b$, то есть $c + a = b$. Тогда, используя аксиомы, получаем:

$$c + a = b \Rightarrow (c + a) + (-a) = b + (-a) \Rightarrow$$

$$\Rightarrow c + (a + (-a)) = b + (-a) \Rightarrow$$

$$\Rightarrow c + 0 = b + (-a) \Rightarrow c = b + (-a).$$

Определение 1.2.3. ***Мультипликативной абелевой группой*** *называется непустое множество M с заданной на нём операцией* ***умножения****, удовлетворяющей следующим аксиомам:*

1) $(\forall\ x, y \in M)\ xy = yx$ ***(коммутативность умножения);***

2) $(\forall\ x, y, z \in M)\ x(yz) = (xy)z$ ***(ассоциативность умножения);***

3) $(\exists\, e \in M) : (\forall\ x \in M)\ xe = x$ ***(существование единичного элемента);***

4) $(\forall\ x \in M)\ (\exists\, x^{-1} \in M) : xx^{-1} = e$ ***(существование обратного элемента).*** •

Все следствия из определения группы 1.2.1 автоматически переносятся на случай мультипликативной абелевой группы.

Пример 1.2.1. 1) Выше мы напомнили известное из школьного курса математики понятие множества векторов (плоскости или трёхмерного пространства), снабжённого бинарной операцией — сложением векторов по правилу параллелограмма. Теперь очевидно, что это множество является аддитивной абелевой группой по сложению.

2) Рассмотрим множество столбцов, составленных из трёх действительных чисел, и определим операцию сложения столбцов правилом

$$\begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} + \begin{pmatrix} b^1 \\ b^2 \\ b^3 \end{pmatrix} = \begin{pmatrix} a^1 + b^1 \\ a^2 + b^2 \\ a^3 + b^3 \end{pmatrix}.$$

Если мы, кроме этого, определим нулевой элемент данного множества, как столбец, составленный из одних нулей, то очевидно, что построенное множество станет аддитивной абелевой группой с введённой операцией сложения столбцов. $\otimes$

Множества с двумя алгебраическими операциями, понятие кольца и поля. Если в некотором множестве ввести две алгебраические операции, то придём к новому понятию.

Определение 1.2.4. *Непустое множество K с двумя алгебраическими операциями — **сложением** и **умножением**, называется **кольцом**, если выполнены следующие аксиомы:*

1) *K есть абелева группа по операции сложения (**аддитивная группа кольца**);*

2) *операции сложения и умножения связаны дистрибутивными законами, то есть*

$$\left(\forall x, y, z \in K\right) x(y+z) = xy + xz \text{ и } (y+z)x = yx + zx. \bullet$$

Кольцо называется ***коммутативным***, если операция умножения коммутативна, то есть

$$\left(\forall x, y \in K\right) xy = yx,$$

и ***ассоциативным***, если операция умножения ассоциативна, то есть

$$\left(\forall x, y, z \in K\right)\left(xy\right)z = x\left(yz\right).$$

Элемент $e \in K$ называется ***единицей***, если

$$\left(\forall x \in K\right) ex = xe = x.$$

Единица единственна (доказательство аналогично доказательству для группы). Однако ***в кольце может не быть единицы***.

Получим некоторые следствия из аксиом кольца.

Следствие 1. $x0 = 0x = 0$.

Доказательство. Действительно, пусть $x0 = y$. Тогда:

$$y + y = x0 + x0 = x\left(0+0\right) = x0 = y \Rightarrow y = y - y = 0.$$

Аналогично показывается, что $0x = 0$. $\bullet\bullet$

Следствие 2. $x(-y)=(-x)y=-xy$.

Д о к а з а т е л ь с т в о. Действительно:

$$xy+x(-y)=x\big(y+(-y)\big)=x0=0\Rightarrow x(-y)=-xy.$$

Аналогично показывается, что $(-x)y=-xy$. ••

Следствие 3. $x(y-z)=xy-xz$ и $(x-y)z=xz-yx$.

Д о к а з а т е л ь с т в о. Действительно:

$$x(y-z)+xz=x(y-z+z)=xy.$$

Аналогично показывается, что $(x-y)z+yz=xz$. ••

Пример 1.2.2. 1) Множество действительных чисел R^1 является коммутативным и ассоциативным кольцом относительно обычных операций сложения и умножения. 2) Множество векторов трёхмерного пространства с операциями сложения и векторного умножения является некоммутативным и неассоциативным кольцом. ⊗

Интуитивно кажется верным, что из равенства $xy=0$ следует, что или $x=0$, или $y=0$, но не может быть, чтобы $x\neq 0$ и $y\neq 0$ одновременно. Однако равенство $xy=0$ может иногда выполняться и в случае, когда $x\neq 0$ и $y\neq 0$ одновременно. В этом случае говорят, что кольцо содержит ***делители нуля***. Кольцо, не содержащее делителей нуля, называется ***областью целостности***.

Определение 1.2.5. *Непустое множество P вместе с двумя алгебраическими операциями —* ***сложением*** *и* ***умножением****, называется* ***полем****, если выполняются следующие аксиомы:*

1) *P есть аддитивная абелева группа по сложению;*
2) $P-\{0\}$ *есть мультипликативная абелева группа по умножению;*
3) *операции сложения и умножения связаны дистрибутивными законами, то есть*

$$(\forall x,y,z\in P)\,x(y+z)=xy+xz,\,(y+z)x=yx+zx.\ \bullet$$

Основное отличие поля от кольца состоит в наличии второй аксиомы, превращающей поле в мультипликативную абелеву группу по умножению, в результате чего в поле появляется обратный элемент, то есть

$$(\forall x\in P)\left(\exists x^{-1}\in P\right):x^{-1}x=xx^{-1}=e.$$

По этой причине трудности с возможным существованием делителей нуля исчезают, а именно справедливо следующее утверждение.

Лемма 1.2.1. *Поле не содержит делителей нуля.*

Доказательство. Действительно, пусть $xy = 0$, но $x \neq 0$. Умножая обе части на обратный элемент x^{-1} слева, получим:

$$x^{-1}(xy) = (x^{-1}x)y = ey = y,$$

и, естественно, $x^{-1}0 = 0$. Получаем $y = 0$. ••

По причине отсутствия делителей нуля любое поле является ***областью целостности***.

Пример 1.2.3. Простейшим и в то же время важным для приложений полем является множество действительных чисел R^1 с обычными операциями сложения и умножения. ⊗

Абстрактные векторные пространства и алгебры. Если в аддитивной абелевой группе определить внешнюю бинарную операцию, получим важнейший пример алгебраической системы.

Определение 1.2.6. *Аддитивная абелева группа E называется **абстрактным векторным пространством** (над полем P), если на ней определена внешняя бинарная операция — **умножение элементов группы E на элементы поля P**, то есть* $(\forall x \in E)$ *и* $(\forall \alpha \in P)$ $(\exists \alpha \cdot x \in E)$, *причём выполняются следующие аксиомы:*

1) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x;$

2) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y;$

3) $1 \cdot x = x;$

4) $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$. •

Векторные пространства будут подробно изучаться на протяжении всего курса, начиная со второй главы. Простейшим примером векторного пространства является множество обычных трёхмерных векторов. Здесь отметим лишь, что из определения 1.2.6 следует, что $(\forall x \in E)\ -x = (-1)x$, где 1 — единица поля P, и $0x = 0$. В последнем равенстве 0 в левой части — это нуль поля P; 0 в правой части равенства — это нулевой элемент из группы E. Элементы векторного пространства называются ***абстрактными векторами***.

Система векторов $\{x_1, x_2, \ldots, x_n\} \subset E$ называется ***линейно независимой***, если

$$\left(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n\right) \Leftrightarrow \left(\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0\right),$$

и ***линейно зависимой*** — в противном случае. Если векторное пространство содержит линейно независимую систему из n векторов, а любая система, содержащая $n+1$ векторов, уже линейно зависима, то говорят, что пространство E является ***n-мерным векторным пространством***. Любая линейно независимая система $\{e_1, e_2, \ldots, e_n\} \subset E$ из n векторов образует ***базис*** пространства. Любой вектор $x \in E$ можно представить в виде разложения

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n,$$

где коэффициенты разложения называются ***координатами*** вектора относительно заданного базиса. Подмножество $L \in E$ называется ***линейным многообразием*** в E, если

$$\left(\forall \alpha, \beta \in P\right)\left(x, y \in L\right) \Rightarrow \left(\alpha \cdot x + \beta \cdot y \in L\right).$$

Если на множестве с двумя бинарными алгебраическими операциями — кольце K, определить ещё внешнюю бинарную операцию, получим новый важнейший пример алгебраической системы.

Определение 1.2.7. *Ассоциативное кольцо K называется* ***алгеброй A над полем P****, если кроме двух бинарных алгебраических операций* (*сложения и умножения*) *в нём определена внешняя бинарная операция —* ***умножение элементов кольца K на элементы поля P****, то есть* $\left(\forall x \in \mathrm{A}\right)$ *и* $\left(\forall \alpha \in P\right)$ $\left(\exists \alpha \cdot x \in \mathrm{A}\right)$*, причём выполняются следующие аксиомы:*

1) $\left(\alpha + \beta\right) \cdot x = \alpha \cdot x + \beta \cdot x;$
2) $\alpha \cdot \left(x + y\right) = \alpha \cdot x + \alpha \cdot y;$
3) $1 \cdot x = x$
4) $\left(\alpha \cdot x\right)\left(\beta \cdot y\right) = \left(\alpha\beta\right) \cdot \left(xy\right).$ •

В аксиоме 3 символом 1 обозначена единица поля ***P***. Кольцо K в определении 1.2.7 иногда называют ***образующим кольцом*** алгебры A. Если образующее кольцо алгебры A содержит единицу, то есть

$$\left(\exists e \in A\right) : \left(\forall x \in A\right) ex = xe = x,$$

то алгебра называется ***алгеброй с единицей***. Если кольцо K коммутативно, то A называется ***коммутативной алгеброй***. Последнее означает, что

$$\left(\forall x, y \in A\right) \left[x, y\right] \stackrel{def}{=} xy - yx = 0.$$

Здесь элемент алгебры $[x, y] \overset{def}{=} xy - yx$ называется ***коммутатором*** элементов x и y.

Подмножество $L \in A$ называется ***линейным многообразием*** в алгебре A, если

$$(\forall \alpha, \beta \in P) \text{ и } (\forall x, y \in L)(\alpha \cdot x + \beta \cdot y \in L).$$

Определение 1.2.8. *Пусть A — некоторая алгебра. Линейное многообразие $\Im \subset A$ называется* ***левым (правым) идеалом алгебры*** *A, если $(\forall x \in A \wedge \forall v \in \Im)\ xv \in \Im$ (соответственно $vx \in \Im$).* •

Если алгебра коммутативная, то понятия левого и правого идеалов совпадают.

Пример 1.2.4. Примером алгебры может служить обычное пространство трёхмерных векторов с операциями сложения и векторного умножения векторов (внутренние бинарные операции) и операцией умножения векторов на действительные числа (внешняя бинарная операция). Причём эта алгебра является некоммутативной. ⊗

Алгебраические системы, подсистемы, изоморфизм[*)]**.** Группы, кольца, поля и алгебры являются примерами ***алгебраических систем*** (или ***структур***). Дадим точное определение алгебраической системы.

Определение 1.2.9. ***Алгебраической системой*** *называется множество X элементов $x, y, \ldots$, оснащённое определённым фиксированным набором*

$$\{(x, y) \mapsto x * y, (x, y) \mapsto x \circ y, \ldots\}$$

внутренних бинарных операций (отображений $X \times X \to X$) и, возможно, определённым фиксированным набором

$$\{(\alpha, x) \mapsto \alpha \perp x, (\alpha, x) \mapsto \alpha \cdot x, \ldots\},$$

где α — элементы некоторого числового поля P, внешних бинарных операций (отображений $P \times X \to X$). •

Пусть X и Y — две алгебраические системы с множествами N_X и N_Y внутренних бинарных операций и множествами M_X и M_Y внешних бинарных операций. Если между множествами внутренних и внешних бинарных операций алгебраических систем X и Y установлены взаимно однозначные соответствия $N_X \leftrightarrow N_Y$ и $M_X \leftrightarrow M_Y$, то системы X и Y называются ***однотипными алгебраическими системами***.

Определение 1.2.10. *Пусть* X *и* X' *— две однотипные алгебраические системы. Отображение* $H : X \to X'$ *системы* X *в систему* X' *называется* ***гомоморфизмом****, если*

$$H(x * y) = (Hx) * (Hy),\ H(x \circ y) = (Hx) \circ (Hy), \ldots;$$

$$H(\alpha \perp x) = \alpha \perp (Hx),\ H(\alpha \cdot x) = \alpha \cdot (Hx), \ldots . \bullet$$

Пусть X — алгебраическая система с множеством

$$N_X = \{(x, y) \mapsto x * y, (x, y) \mapsto x \circ y, \ldots\}$$

внутренних и множеством

$$M_X = \{(\alpha, x) \mapsto \alpha \perp x, (\alpha, x) \mapsto \alpha \cdot x, \ldots\}$$

внешних бинарных операций. Если $Y \subset X$ — такое подмножество, что

$$(\forall x, y \in Y)\ X'(\forall \alpha \in P)$$

$$x * y \in Y, x \circ y \in Y, \ldots \wedge \alpha \perp x \in Y, \alpha \cdot x \in Y, \ldots,$$

то говорят, что бинарные операции

$$N_X = \{(x, y) \mapsto x * y, (x, y) \mapsto x \circ y, \ldots\},$$

$$M_X = \{(\alpha, x) \mapsto \alpha \perp x, (\alpha, x) \mapsto \alpha \cdot x, \ldots\}$$

в X ***индуцируют*** в Y бинарные операции

$$N_Y = \{(x, y) \mapsto x *' y, (x, y) \mapsto x \circ' y, \ldots\},$$

$$M_Y = \{(\alpha, x) \mapsto \alpha \perp' x, (\alpha, x) \mapsto \alpha \cdot' x, \ldots\}.$$

Относительно индуцированных (штрихованных) бинарных операций подмножество Y является алгебраической системой, однотипной с системой X. Однако, в системе Y могут не выполняться некоторые аксиомы системы X, например, Y может не содержать нуль системы X. Эта ситуация характеризуется термином ***алгебраическая подсистема***, который определяется ниже.

Определение 1.2.11. *Подмножество* Y *алгебраической системы* X *(группы, кольца, поля, алгебры) называется* ***алгебраической подсистемой*** *системы* X *(****подгруппой****,* ***подкольцом****,* ***подполем****,* ***подалгеброй****), если относительно индуцированных бинарных операций оно является алгебраической* *системой, однотипной с системой X, и при этом в Y сохраняется набор аксиом системы X.* •

Пусть X и X' — две однотипные алгебраические системы и $H : X \to X'$ — гомоморфизм системы X на систему X'. Если гомоморфизм является взаимно однозначным, то он называется ***изоморфизмом***.

Определение 1.2.12. *Алгебраические системы X и X' называются* ***изоморфными****, если существует изоморфизм $H : X \to X'$.* •

Изоморфные алгебраические системы отождествляют, если их рассматривают только с точки зрения алгебраических свойств, то есть свойств, связанных с наборами внутренних и внешних бинарных операций.

Пример 1.2.5. Важным примером изоморфных алгебраических систем являются поля. Используя запись

$$\frac{x}{y} \equiv xy^{-1},$$

нетрудно показать, что во всяком поле сохраняются правила выполнения арифметических операций с дробями — все поля с точки зрения выполнения операций над дробями неотличимы от поля действительных чисел. По этой причине *элементы любого поля часто называют просто числами*. $\otimes$

1.3. Числовые поля

Поле действительных чисел. Аксиомы сложения. В этом параграфе приведены достаточно подробные сведения о двух числовых полях, которые постоянно используются при изучении физики и в физических исследованиях.

Первое поле, обзор теории которого мы приведём в этом параграфе, — поле действительных чисел, знакомо нам из школы и наиболее часто встречается в повседневной жизни. Теорию действительных чисел можно строить по-разному. Можно показать, однако, что все варианты такой теории эквивалентны. По этой причине мы выберем наиболее короткий вариант — аксиоматическое построение теории поля действительных чисел.

Во множестве действительных чисел R^1 определена ***операция сложения***. Приведём аксиомы сложения.

С.0. *Во множестве R^1 задана алгебраическая операция, называемая **сложением**, результат выполнения которой обозначается $a+b$ и называется **суммой** элементов a и b.*

С.1. *Множество R^1 по отношению к операции сложения является аддитивной **абелевой группой**, то есть справедливы следующие свойства операции сложения*:

1) $\left(\forall a, b, c \in R^1\right)(a+b)+c=a+(b+c)$ — ***ассоциативность*** *операции сложения*;

2) $\left(\forall a, b \in R^1\right)\ a+b=b+a$ — ***коммутативность*** *операции сложения*;

3) $\left(\exists 0 \in R^1\right):\left(\forall a \in R^1\right)\ a+0=a$ — *существование единственного* ***числа нуль***;

4) $\left(\forall a \in R^1\right)\left(\exists(-a) \in R^1\right):\ a+(-a)=0$ — *существование числа,* ***противоположного*** *каждому числу $a \in R^1$.*

В аксиоме С.0 постулируется существование суммы двух действительных чисел. Можно определить и сумму любого конечного числа действительных чисел $a_1, a_2, \ldots, a_n$. Для этого нужно придать определённый смысл выражению $a_1+a_2+\ldots+a_n$, что легко сделать на основе аксиомы С. 1. Легко показать, что при любом конечном числе элементов результат вычислений выражения $a_1+a_2+\ldots+a_n$ не зависит от способа расстановки скобок.

Поле действительных чисел. Аксиомы умножения. Во множестве действительных чисел определена ***операция умножения***. Приведём аксиомы умножения.

У.0. *На множестве R^1 задана алгебраическая операция, называемая **умножением**. Результат выполнения операции умножения называется **произведением** и обозначается ab.*

У.1. *Множество $R^*=R^1-\{0\}$ по отношению к операции умножения является мультипликативной **абелевой группой**, то есть справедливы следующие свойства операции умножения*:

1) $\left(\forall a, b, c \in R^1\right)(ab)c=a(bc)$ — ***ассоциативность*** *умножения*;

2) $\left(\forall a, b \in R^*\right)\ ab=ba$ — ***коммутативность*** *операции умножения*;

3) $\left(\exists e \in R^1\right):\left(\forall a \in R^*\right)\ ae=a$ — *существование* ***единицы****, которая обозначается $e \equiv 1$*;

4) $\left(\forall a \in R^*\right)\left(\exists a^{-1} \in R^*\right): aa^{-1} = 1$ — *существование для каждого действительного числа* $a \neq 0$ *действительного числа* a^{-1}, *которое называется* ***обратным*** *числу a.*

Поле действительных чисел. Дистрибутивные законы. Операции сложения (аксиомы С) и умножения (аксиомы У) связаны между собой ***дистрибутивными законами***, которые имеют следующий вид:

Д.1) $\left(\forall a, b, c \in R^1\right)\ a(b+c) = ab + ac$;

Д.2) $\left(\forall a, b, c \in R^1\right)\ (b+c)a = ba + ca$.

Сравнение аксиом сложения и умножения действительных чисел с определением поля показывает, что ***множество действительных чисел*** R^1 ***является полем***. В силу изоморфизма элементы любого поля по своим свойствам неотличимы от действительных чисел. Поэтому элементы произвольных полей также часто называют ***числами***.

Поле действительных чисел. Аксиомы порядка. Выше мы уже использовали свойство упорядоченности множества действительных чисел. Сформулируем аксиомы порядка множества действительных чисел. Для общности рассмотрим так называемую ***замкнутую числовую прямую***, задавая её как объединение множества действительных чисел R^1 и двухэлементного множества $\{-\infty, +\infty\}$, элементами которого являются два абстрактных символа, носящих название ***бесконечно удалённых точек***, или просто точки «***минус бесконечность***» и «***плюс бесконечность***»: $\overline{R^1} = R^1 \cup \{-\infty, +\infty\}$.

П.0. *В* $\overline{R^1}$ *задано отношение* $\leq$, *то есть* $\left(\forall a, b \in \overline{R^1}\right)$ *установлено, выполняется* $a \leq b$ *или нет.*

П.1. $\left(\forall a \in \overline{R^1}\right)\ a \leq a$ *(**рефлексивность отношения** $\leq$).*

П.2. $a \leq b \wedge b \leq a \Leftrightarrow a = b$ *(**антисимметрия отношения** $\leq$).*

П.3. $a \leq b \wedge b \leq c \Rightarrow a \leq c$ *(**транзитивность отношения** $\leq$).*

П.4. $\left(\forall a, b \in \overline{R^1}\right) a \leq b \vee b \leq a \vee a = b$.

П.5. $\left(\forall a \in R^1\right) -\infty \leq a \leq +\infty$.

Определение 1.3.1. *Упорядоченная пара* (M_d, M_u) *множеств* $M_d \subset \overline{R}$ *и* $M_u \subset \overline{R^1}$ *называется* ***дедекиндовым сечением***, *если эти множества удовлетворяют следующим условиям:*

1) $M_d \neq \varnothing$, $M_u \neq \varnothing$;

2) $M_d \cup M_u = \overline{R^1}$;

3) $(\forall a \in M_d) \wedge (\forall b \in M_u)\ a \leq b$.

Последнее условие записывается ещё так: $M_d \leq M_u$.

П.6. Аксиома полноты множества действительных чисел (***аксиома о дедекиндовом сечении***). *Пусть* (M_d, M_u) *— произвольное дедекиндово сечение, тогда*

$$(\exists s \in \overline{R}): (\forall a \in M_d) \wedge (\forall b \in M_u)\, a \leq s \leq b.$$

Высказывание $a \leq s \leq b$ записывается также в виде $M_d \leq s \leq M_u$. Число s называется ***секущим числом*** сечения (M_d, M_u). Наглядный смысл аксиомы П.6 достаточно прост: ***действительные числа распределены непрерывно, без разрывов — на числовой прямой нет «дыр»***.

П.7. *Если* $a, b, c \in R^1$ *и* $a \leq b$, *то* $a + c \leq b + c$.

П.8. $(\forall a, b \in R^1)$ *из* $0 \leq a$ *и* $0 \leq b$ *следует* $0 \leq ab$.

Можно показать, что справедлива следующая теорема.

Теорема 1.3.1 (теорема Архимеда). *Для каждого действительного числа* a *существует натуральное число* $n > a$.

Из приведённых аксиом следуют ***свойства неравенств и правила действий с ними:***

1) *из* $a < b < c$, *следует* $a < c$; *из* $a \leq b < c$ *следует* $a < c$;

2) *из* $a > b \geq c$ *следует* $a > c$; *из* $a \geq b > c$ *следует* $a > c$;

3) *если* $a \leq b$ *и* $c \leq d$, *то* $a + c \leq b + d$; *если* $a \leq b$ *и* $c < d$, *то* $a + c < b + d$;

4) *элемент a является положительным (отрицательным) в том и только в том случае, если элемент –a отрицателен (положителен). В частности, получаем правило знаков для бесконечно удалённых точек:*

$$-(+\infty) = -\infty \text{ и } -(-\infty) = +\infty;$$

5) *произведение положительных чисел положительно, как и произведение отрицательных чисел. Если a положительно, a b отрицательно, то ab отрицательно.*

Правило 5 записывается в виде схемы действий со знаками:

$$(+)(+) = (+);\ (+)(-) = (-);\ (-)(+) = (-);\ (-)(-) = (+);$$

6) *из* $a < b$ *и* $0 < c$ *следует* $ac < bc$;

7) *из* $a < b$ *и* $c < 0$ *следует* $ac > bc$;

8) $0 < 1$;

9) *из* $0 < a$ *и* $a < b$ *следует* $|\alpha| \le |\alpha + \beta| + |\beta| \Rightarrow |\alpha + \beta| \ge |\alpha| - |\beta|$.

Абсолютная величина действительного числа. Важнейшим понятием теории поля действительных чисел является понятие ***абсолютной величины*** действительного числа.

Определение 1.3.2. *Пусть* $a \in R^1$, *тогда число*

$$|a| = \begin{cases} a, \text{если } a > 0, \\ -a, \text{если } a < 0, \\ 0, \text{если } a = 0 \end{cases}$$

называется ***абсолютной величиной*** *действительного числа* a.•

Абсолютная величина действительного числа обладает свойствами, которые устанавливаются в следующей теореме.

Теорема 1.3.2 (свойства абсолютной величины). *Справедливы следующие утверждения:*

1) $\left(\forall a \in R^1\right)|a| \ge 0$, причём $|a| = 0 \Leftrightarrow a = 0$;

2) $\left(\forall a, b \in R^1\right)|ab| = |a||b|$;

3) $\left(\forall \varepsilon \in R^1, \varepsilon > 0\right) \wedge \left(\forall x \in R^1\right)|x| \le \varepsilon \Leftrightarrow -\varepsilon \le x \le \varepsilon$;

4) $\left(\forall a, b \in R^1\right)|a + b| \le |a| + |b|$ (***неравенство треугольника***).

Доказательство. 1) Справедливость этого утверждения следует непосредственно из определения 1.3.2.

2) Если $a \ge 0$ и $b \ge 0$, то $|ab| = \{+ \cdot + = +\} = ab = |a| \cdot |b|$.

Если $a \le 0$ и $b \le 0$, то $|ab| = \{- \cdot - = +\} = ab = (-a)(-b) = |a| \cdot |b|$.

Если $a \ge 0$, а $b \le 0$, то $|ab| = \{+ \cdot - = -\} = -ab = a \cdot (-b) = |a| \cdot |b|$.

3) Покажем, что $|x| \le \varepsilon \Rightarrow -\varepsilon \le x \le \varepsilon$. Для этого рассмотрим два случая.

а) Пусть $x \ge 0$. Тогда, так как $\varepsilon > 0$, то $x \ge -\varepsilon$, а так как $|x| \le \varepsilon$, то $x = |x| \le \varepsilon$. Получаем, что $-\varepsilon \le x \le \varepsilon$.

б) Пусть $x \le 0$. Тогда, так как $\varepsilon > 0$, то $\varepsilon \ge x$, а так как $|x| \le \varepsilon$, то $-x = |x| \le \varepsilon$, или $x \ge -\varepsilon$. Снова получаем, что $-\varepsilon \le x \le \varepsilon$.

4) Очевидно, что

$$(\forall a, b \in R) -|a| \le a \le |a| \wedge -|b| \le b \le |b|$$

Сложив эти неравенства, получаем:

$$-\left(|a|+|b|\right) \le a+b \le |a|+|b|.$$

Из утверждения 3 теперь следует, что $|a+b| \le |a|+|b|$. ••

Следствие из теоремы 1.3.2. *Если* $x_0 \in R^1$, *то множество*

$$U_\varepsilon(x_0) = \left\{x \in R^1 : |x - x_0| < \varepsilon\right\}$$

совпадает с множеством $(x_0 - \varepsilon; x_0 + \varepsilon)$.

Интерпретация этого факта приведена на рис. 1.4.

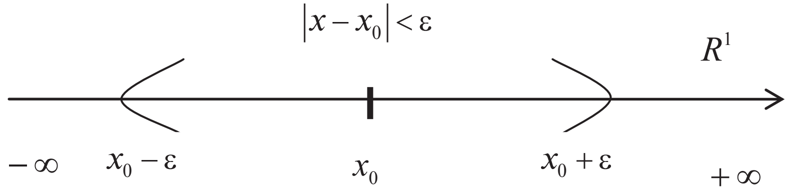


Рис. 1.4

Итак, множество R^1 является упорядоченным полем по отношению к операциям сложения и умножения его элементов — действительных чисел. Существуют, однако, и другие числовые поля.

Поле комплексных чисел. Аксиоматическое построение и теорема существования. Ещё одним полем, весьма важным для практики, является поле, называемое ***полем комплексных чисел***. Необходимость построения поля комплексных чисел диктуется, например, задачей о решении квадратных уравнений, простейшее из которых $x^2 + 1 = 0$ уже не имеет решения в поле действительных чисел. Формальная запись решения этого уравнения $x = \pm\sqrt{-1}$ приводит нас к новому символу, обозначаемому $i \overset{def}{=} \sqrt{-1}$. Из дальнейшего видно, что построение поля комплексных чисел производится путём выхода из одномерного множества (числовой прямой) в двумерное множество — декартову плоскость.

При построении теории комплексных чисел мы применим ***аксиоматический метод***, то есть введём формально некоторые объекты, подчинив их аксиомам, а затем покажем, что введённое множество является полем.

Определение 1.3.3. ***Комплексными числами*** *называются упорядоченные пары действительных чисел вида* (a,b)*, для которых операции сложения и умножения вводятся посредством определения результата их выполнения в соответствии со следующими аксиомами.*

1. *Два комплексных числа* (a,b) *и* (c,d) *считаются* ***равными*** *в том и только в том случае, если* $a=c$ *и* $b=d$*, что при помощи логической символики записывается так:*

$$(a,b)=(c,d)\Leftrightarrow a=c\wedge b=d. \tag{1.3.1}$$

2. ***Сумма*** *двух комплексных чисел* (a,b) *и* (c,d) *является комплексным числом, которое находится по правилу:*

$$(a,b)+(c,d)\overset{def}{=}(a+c,b+d). \tag{1.3.2}$$

3. ***Произведение*** *двух комплексных чисел* (a,b) *и* (c,d) *является комплексным числом, которое находится по правилу:*

$$(a,b)(c,d)\overset{def}{=}(ac-bd,ad+bc). \tag{1.3.3}$$

4. *Комплексное число* $(a,0)$ ***отождествляется с действительным (вещественным) числом*** a*:* $(a,0)\equiv a$*. В частности,* $(0,0)\equiv 0$*.*

Числа вида $(0,b)$ *называются* ***мнимыми числами***.•

Из определения 1.3.3 следует, что операции сложения и умножения комплексных чисел являются алгебраическими. Комплексные числа принято обозначать малыми буквами латинского или греческого алфавита, например, $z=(a,b)$. Множество комплексных чисел обозначается буквой C.

По определению ***степень комплексного числа*** (a,b) вводится как n-кратное произведение этого числа на себя:

$$(a,b)^n\overset{def}{=}\underbrace{(a,b)(a,b)\ldots(a,b)}_{n\text{ раз}}. \tag{1.3.4}$$

Из аксиомы 4 следует отождествление комплексного числа $(1,0)$ с единицей поля действительных чисел: $(1,0)\equiv 1$. По аналогии мнимое число $i=(0,1)$ называется ***мнимой единицей***. Рассмотрим возведение мнимой единицы i во вторую степень. Используя аксиому 3, получаем:

$$i^2 = (0,1)(0,1) = (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 1\cdot 0) = (-1,0) \equiv -1.$$

Имеем важнейшее свойство символа i: $i^2 = -1$.

Покажем, что операции сложения и умножения комплексных чисел ***коммутативны***, ***ассоциативны*** и связаны ***дистрибутивными законами***. Обозначим

$$z_1 = (a,b),\ z_2 = (c,d),\ z_3 = (u,v).$$

Коммутативность операций сложения и умножения комплексных чисел непосредственно следует из сравнения правых частей приведённых ниже равенств с учётом коммутативности операций сложения и умножения действительных чисел и аксиом 2 и 3:

$$1)\begin{cases} z_1 + z_2 = (a,b) + (c,d) = (a+c, b+d), \\ z_2 + z_1 = (c,d) + (a,b) = (c+a, d+b); \end{cases}$$

$$2)\begin{cases} z_1 z_2 = (a,b)(c,d) = (ac-bd, ad+bc), \\ z_2 z_1 = (c,d)(a,b) = (ca-db, cb+da). \end{cases}$$

Ассоциативность операций сложения и умножения комплексных чисел также следует из свойств соответствующих операций над действительными числами, аксиом 2 и 3 и сравнения правых частей следующих равенств:

$$1)\begin{cases} (z_1+z_2)+z_3 = \left[(a,b)+(c,d)\right]+(u,v) = (a+c,b+d)+(u,v) = \\ = \left((a+c)+u,(b+d)+v\right) = (a+c+u, b+d+v), \\ z_1+(z_2+z_3) = (a,b)+\left[(c,d)+(u,v)\right] = (a,b)+(c+u,d+v) = \\ \left(a+(c+u), b+(d+v)\right) = (a+c+u, b+d+v); \end{cases}$$

$$2)\begin{cases} (z_1z_2)z_3 = \left[(a,b)(c,d)\right](u,v) = (ac-bd, ad+bc)(u,v) = \\ = (acu-bdu-adv-bcv, acv-bdv+adu+bcu), \\ z_1(z_2z_3) = (a,b)\left[(c,d)(u,v)\right] = (a,b)(cu-dv, cv+du) = \\ = (acu-adv-bcv-bdu, acv+adu+bcu-bdv). \end{cases}$$

Операции умножения и сложения комплексных чисел связаны дистрибутивными законами, справедливость которых вытекает из аксиом 1–3 и сравнения правых частей следующих равенств:

$$(z_1+z_2)z_3=\left[(a,b)+(c,d)\right](u,v)=(a+c,b+d)(u,v)=$$
$$=(au+cu-bv-dv,av+cv+bu+du);$$
$$z_1z_3+z_2z_3=(a,b)(u,v)+(c,d)(u,v)=(au-bv,av+bu)+$$
$$+(cu-dv,cv+du)=(au+cu-bv-dv,av+cv+bu+du).$$

Исследуем вопрос о существовании обратных операций.

Определение 1.3.4. ***Разностью*** *двух комплексных чисел*

$$z_1=(a,b),\ z_2=(c,d)$$

называется такое комплексное число $z=(x,y)$*, что* $z_2+z=z_1$.•

Итак, по определению для разности имеем $z_2+z=z_1$.

По аксиомам 1 и 2

$$(c,d)+(x,y)=(c+x,d+y)\Rightarrow(c+x,d+y)=(a,b),$$

откуда, применяя аксиому 1, получаем два уравнения

$$c+x=a;\ d+y=b,$$

решения которых дают $z=(a-c,b-d)$.

Итак, во множестве комплексных чисел для операции сложения существует обратная операция, называемая ***вычитанием***, причём её результат — ***разность*** $z=z_1-z_2$ комплексных чисел z_1 и z_2, находится по правилу

$$z=z_1-z_2=(a,b)-(c,d)=(a-c,b-d). \qquad (1.3.5)$$

Очевидно, что комплексное число $(-a,-b)$ — ***противоположно*** комплексному числу (a,b).

Определение 1.3.5. ***Частным комплексных чисел*** $z_1=(a,b)$ *и* $z_2=(c,d)$ $(z_2\neq(0,0))$ *называется такое комплексное число* $z=(x,y)$*, что* $z_1=zz_2$. •

По определению

$$z_1=zz_2\Rightarrow(a,b)=(x,y)(c,d)\Rightarrow(a,b)=(cx-dy,dx+cy),$$

откуда, применяя аксиому 1, получаем систему уравнений

$$\begin{cases}cx-dy=a,\\ dx+cy=b.\end{cases}$$

Решение этой системы уравнений имеет вид

$$x=\frac{ac+bd}{c^2+d^2},\ y=\frac{bc-ad}{c^2+d^2}.$$

Итак, во множестве комплексных чисел для операции умножения существует обратная операция, называемая ***делением***, причём её результат — ***частное*** z комплексных чисел z_1 и z_2 обозначается $z=\frac{z_1}{z_2}$ и находится по правилу

$$z=\frac{z_1}{z_2}=\frac{(a,b)}{(c,d)}=\left(\frac{ac+bd}{c^2+d^2},\frac{bc-ad}{c^2+d^2}\right). \qquad (1.3.6)$$

Очевидно, что если $z_1=z_2$, то из формулы (1.3.6) следует $\frac{z_1}{z_2}=(1,0)$. Следовательно, единицей для операции умножения служит комплексное число $(1,0)$, которое отождествляется с действительным числом 1. Полагая далее в формуле (1.3.6) $z_1=(1,0)$, получаем, что при $z_2\neq(0,0)=0$ комплексное число, обратное комплексному числу z_2, определяется по правилу

$$(z_2)^{-1}=\frac{(1,0)}{(c,d)}=\left(\frac{c}{c^2+d^2},\frac{-d}{c^2+d^2}\right). \qquad (1.3.7)$$

Рассмотрим теперь вопрос о единственности представления комплексного числа z.

Допустим, что комплексное число z может быть представлено в виде двух упорядоченных пар: $z=(a,b)=(c,d)$. Тогда по аксиоме 1 получаем $a=c$ и $b=d$, что и доказывает единственность представления комплексного числа.

Суммируя всё сказанное выше, сформулируем теорему существования и единственности множества комплексных чисел.

Теорема 1.3.3. *Множество комплексных чисел, введённых определением 1.3.1, существует и является полем, включающим в себя поле вещественных чисел в качестве* ***подполя****, причём представление комплексных чисел в виде упорядоченных пар вещественных чисел единственно с точностью до изоморфизма, переводящего все вещественные числа в себя.*

В отличие от поля действительных чисел R, ***поле комплексных чисел C неупорядочено*** — отношение порядка в поле комплексных чисел не определено.

Отметим также следующий результат, вытекающий из всего предыдущего рассмотрения теории комплексных чисел и определения алгебры: *множество C комплексных чисел, являясь полем, одновременно является и* ***коммутативной алгеброй над полем R вещественных чисел***.

Алгебраическая форма комплексного числа. Рассмотрим следующую выкладку

$$(a,b)=(a,0)+(0,b)=(a,0)+(0,1)(b,0),$$

из которой в соответствии с аксиомой 4, постулирующей отождествление комплексных чисел вида $(a,0)$ с вещественными числами, и обозначением для мнимой единицы вытекает следующая форма записи комплексных чисел

$$(a,b)=a+ib\equiv a+bi, \tag{1.3.8}$$

которая называется ***алгебраической формой*** комплексных чисел. Вещественное число a называется ***вещественной частью*** комплексного числа z и обозначается $a\equiv\operatorname{Re}z$, а вещественное число b называется ***мнимой частью*** комплексного числа z и обозначается $b\equiv\operatorname{Im}z$.

Путём прямого сравнения с соответствующими определениями нетрудно показать, что все операции над комплексными числами в алгебраической форме с учётом свойства мнимой единицы $i^2=-1$ сводятся к простому раскрытию скобок.

Пример 1.3.1. Даны комплексные числа $z_1=(2,7)$ и $z_2=(1,4)$. Найти z_1+z_2 и z_1z_2.

Р е ш е н и е. Используя аксиомы поля комплексных чисел, для z_1+z_2 и z_1z_2 получаем:

$$z_1+z_2=(2,7)+(1,4)=(3,11),$$

$$z_1z_2=(2,7)(1,4)=(-26,15).$$

В алгебраической форме имеем:

$$z_1+z_2=(2+7i)+(1+4i)=3+11i,$$

$$z_1z_2=(2+7i)(1+4i)=2+8i+7i+28i^2=-26+15i.\otimes$$

Определение 1.3.6. *Число вида $a-ib$ называется* ***комплексно-сопряжённым*** *к числу $z=a+ib$ и обозначается*

$$a-ib\equiv\bar{z}\equiv z^*\equiv\overline{a+ib}\,.\bullet$$

Сопоставление комплексному числу $z = a + ib$ комплексно-сопряжённого к нему числа $\bar{z} = a - ib$ можно рассматривать как ***операцию комплексного сопряжения***, действующую на множестве всех комплексных чисел. Если операцию комплексного сопряжения обозначить $(*)$, то будем иметь $(*): C \to C$.

Теорема 1.3.4. *Для комплексно-сопряжённых чисел справедливы следующие утверждения:*

1) $\overline{(a+ib)} + \overline{(c+id)} = \overline{(a+c)+i(b+d)}$;

2) $\overline{(a+ib)} \cdot \overline{(c+id)} = \overline{(a+ib)\cdot(c+id)}$;

3) $(a+ib)(a-ib) = a^2 + b^2$.

Д о к а з а т е л ь с т в о. Например, для третьего свойства имеем:

$$(a+ib)(a-ib) = a^2 - iab + iab - (ib)^2 = a^2 - (-1)\cdot b^2 = a^2 + b^2. \bullet\bullet$$

Нахождение частного двух комплексных чисел связано с необходимостью запоминания формулы (1.3.6). Найти частное можно проще, используя понятие комплексно-сопряжённого числа и утверждение 3 теоремы 1.3.4; действительно,

$$\frac{z_1}{z_2} = \frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i,$$

что как раз совпадает с формулой (1.3.6).

Пример 1.3.2. Даны комплексные числа $z_1 = (2, 7)$ и $z_2 = (1, 4)$. Найти частное $\frac{z_1}{z_2}$.

Р е ш е н и е. Представляя комплексные числа в алгебраической форме, получаем:

$$\frac{z_1}{z_2} = \frac{2+7i}{1+4i} = \frac{(2+7i)(1-4i)}{(1+4i)(1-4i)} = \frac{30}{17} - \frac{1}{17}i. \otimes$$

Геометрическая интерпретация и тригонометрическая форма комплексного числа. Проводим на плоскости две взаимно перпендикулярные прямые линии и принимаем точку их пересечения за ***начало системы координат** O*. Выбираем на этих прямых положительное направление, откладывая от начала O один и тот же масштабный отрезок единичной длины. Горизонтальную прямую называют ***осью***

абсцисс и обозначают OX. Вертикальную прямую называют ***осью ординат*** и обозначают OY. При таком построении каждой точке плоскости взаимно однозначным образом ставятся в соответствие два действительных (вещественных) числа x и y, которые называются её ***декартовыми координатами — абсциссой*** и ***ординатой*** соответственно. Приходим к декартовой системе координат.

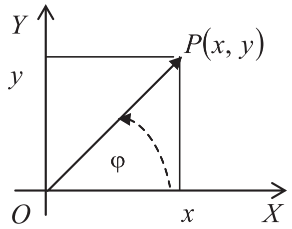


Рис. 1.5

На рис. 1.5 точка P, например, имеет координаты x и y, что обозначается $P(x; y)$. Точке P сопоставляется ***радиус-вектор*** $\vec{r} \equiv \vec{OP}$, координаты которого также равны x и y. Угол, образованный радиус-вектором точки P и положительным направлением оси OX, называется ***полярным углом***, за положительное направление отсчёта полярного угла принимается направление поворота радиус-вектора против часовой стрелки. Итак, каждой точке декартовой плоскости ставятся в соответствие ***полярные координаты*** — длина радиус-вектора r и полярный угол φ. Получаем ***полярную систему координат***.

Принимается соглашение, в соответствии с которым комплексные числа $z = (a, b)$ (или $z = a + ib$) изображают ***точками на декартовой плоскости***, причём число a является абсциссой, а число b — ординатой точки z. Вещественные числа изображаются точками на оси абсцисс, а мнимые числа — точками на оси ординат. В начало координат помещают число $0 = (0, 0)$ (рис. 1.5). При таком соглашении устанавливается взаимно однозначное (биективное) соответствие между точками плоскости и множеством комплексных чисел.

Определение 1.3.7. *Декартова плоскость, на которой изображаются комплексные числа, называется* ***комплексной плоскостью****. Ось абсцисс на комплексной плоскости называется* ***вещественной осью****, а ось ординат называется* ***мнимой осью****.* •

Определение 1.3.8. ***Модулем*** *комплексного числа z называется длина радиус-вектора точки z на комплексной плоскости.* ***Аргументом*** *комплексного числа z называется полярный угол φ точки, изображающей это число.* •

Модуль комплексного числа z обозначается $|z|$, аргумент комплексного числа z обозначается $\arg z$. Аргумент комплексного числа $0=(0,0)$ не имеет смысла. Положительным направлением отсчёта аргумента комплексного числа считается направление от положительной полуоси вещественной оси к положительной полуоси мнимой оси, то есть против часовой стрелки. Очевидно, что все возможные значения аргумента φ комплексного числа z даются формулой

$$\varphi=\varphi_0+2\pi k,$$

где $k=0,\pm 1,\pm 2,\ldots$ — аргумент комплексного числа ***многозначен***. Справедливы формулы:

$$a=r\cos\varphi;\ b=r\sin\varphi;\ r=\sqrt{a^2+b^2}\ ;\ \operatorname{tg}\varphi=\frac{b}{a}.$$

Таким образом, для комплексного числа $z=a+ib$ получаем ***тригонометрическую форму записи***:

$$z=r\cdot(\cos\varphi+i\sin\varphi). \qquad (1.3.9)$$

Для устранения многозначности аргумента комплексного числа полагают, что его ***главное значение*** φ_0 удовлетворяет условию $-\pi\le\varphi_0\le\pi$. Знак главного значения φ_0 выбирается по знаку $\cos\varphi=\dfrac{a}{\sqrt{a^2+b^2}}$ и $\sin\varphi=\dfrac{b}{\sqrt{a^2+b^2}}$.

Пример 1.3.3. Записать в тригонометрической форме комплексное число $z=-1-i$.

Р е ш е н и е. Вычисляем модуль числа z по формуле

$$r=\sqrt{a^2+b^2}:|z|=\sqrt{2}.$$

Так как $\operatorname{tg}\varphi=\dfrac{b}{a}=1$, причём $\operatorname{Re}z=-1<0$ и $\operatorname{Im}z=-1<0$ (точка лежит в третьей четверти), получаем $\varphi_0=-\pi+\dfrac{\pi}{4}=-\dfrac{3\pi}{4}$. Записываем тригонометрическую форму числа z:

$$z = \sqrt{2}\left(\cos\left(-\frac{3\pi}{4} + 2k\pi\right) + i\sin\left(-\frac{3\pi}{4} + 2k\pi\right)\right). \otimes$$

Теорема 1.3.5. *Для модуля суммы и разности двух комплексных чисел* α *и* β *имеют место неравенства:*

1) $|\alpha + \beta| \le |\alpha| + |\beta|$,

2) $|\alpha - \beta| \le |\alpha| + |\beta|$,

3) $|\alpha + \beta| \ge |\alpha| - |\beta|$,

4) $|\alpha - \beta| \ge |\alpha| - |\beta|$.

Доказательство. Докажем, например, неравенство 1. Пусть

$$\alpha = r_1(\cos\varphi_1 + i\sin\varphi_1),\ \beta = r_2(\cos\varphi_2 + i\sin\varphi_2),$$

где $r_1 = |\alpha|$, $r_2 = |\beta|$. Для

$$\alpha + \beta = (r_1\cos\varphi_1 + r_2\cos\varphi_2) + i(r_1\sin\varphi_1 + r_2\sin\varphi_2)$$

получаем:

$$|\alpha + \beta|^2 = (r_1\cos\varphi_1 + r_2\cos\varphi_2)^2 + (r_1\sin\varphi_1 + r_2\sin\varphi_2)^2 =$$

$$= r_1^2\cos^2\varphi_1 + 2r_1r_2\cos\varphi_1\cos\varphi_2 +$$

$$+ r_2^2\cos^2\varphi_2 + r_1^2\sin^2\varphi_1 + 2r_1r_2\sin\varphi_1\sin\varphi_2 + r_2^2\sin^2\varphi_2 =$$

$$= r_1^2\left(\cos^2\varphi_1 + \sin^2\varphi_1\right) +$$

$$+ 2r_1r_2(\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2) + r_2^2\left(\cos^2\varphi_2 + \sin^2\varphi_2\right) =$$

$$= r_1^2 + 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2.$$

Но $\cos(\varphi_1 - \varphi_2) \le 1$. Поэтому

$$|\alpha + \beta|^2 \le r_1^2 + 2r_1r_2 + r_2^2 = (r_1 + r_2)^2.$$

Так как $|\alpha + \beta| > 0$ и $r_1 + r_2 > 0$, то $|\alpha + \beta| \le r_1 + r_2 = |\alpha| + |\beta|$.

Теперь доказательства остальных неравенств получаются почти автоматически.

Для доказательства неравенства 2 заметим, что $|\beta| = |-\beta|$. Действительно, компоненты чисел β и $-\beta$ отличаются только знаками и, следовательно, их квадраты одинаковы. Поэтому имеем:

$$|\alpha - \beta| = |\alpha + (-\beta)| \le |\alpha| + |-\beta| = |\alpha| + |\beta|.$$

Для доказательства неравенства 3 применим неравенство 2 к тождеству

$$\alpha = (\alpha + \beta) - \beta,$$

получим:

$$|\alpha| \le |\alpha + \beta| + |\beta| \Rightarrow |\alpha + \beta| \ge |\alpha| - |\beta|.$$

Для неравенства 4 имеем:

$$|\alpha - \beta| = |\alpha + (-\beta)| \ge |\alpha| - |\beta|. \bullet\bullet$$

Если записать комплексные числа в тригонометрической форме

$$\alpha_1 = r_1(\cos\varphi_1 + i\sin\varphi_1),\ \alpha_2 = r_2(\cos\varphi_2 + i\sin\varphi_2),$$

то для их произведения легко получается формула:

$$\alpha_1\alpha_2 = r_1r_2\big((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) +$$

$$+i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2)\big) =$$

$$= r_1r_2\big(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\big). \qquad (1.3.10)$$

Из формулы (1.3.10) следует, что
1) $|\alpha_1\alpha_2| = |\alpha_1|\cdot|\alpha_2|$, 2) $\arg(\alpha_1\alpha_2) = \arg\alpha_1 + \arg\alpha_2$.

Эти формулы распространяются на произведение любого конечного числа сомножителей.

Формула для ***частного двух комплексных чисел***, записанных в тригонометрической форме, получается аналогично:

$$\frac{\alpha_1}{\alpha_2} = \frac{r_1}{r_2}\big(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)\big). \qquad (1.3.11)$$

Рассмотрим ***возведение комплексного числа в степень с натуральным показателем***. Для этого в формуле

$$r_1(\cos\varphi_1 + i\sin\varphi_1)\cdot\ldots\cdot r_n(\cos\varphi_n + i\sin\varphi_n) =$$

$$= r_1r_2\ldots r_n\big(\cos(\varphi_1 + \ldots + \varphi_n) + i\sin(\varphi_1 + \ldots + \varphi_n)\big)$$

положим $r_1 = r_2 = \ldots = r_n = r$, $\varphi_1 = \varphi_2 = \ldots = \varphi_n = \varphi$. Получаем

$$\big(r(\cos\varphi + i\sin\varphi)\big)^n = r^n(\cos n\varphi + i\sin n\varphi). \qquad (1.3.12)$$

При $r=1$ из формулы (1.3.12) получаем известную ***формулу Муавра***:

$$(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi. \qquad (1.3.13)$$

Определение 1.3.9. ***Корнем n-й степени из комплексного числа*** α *называется такое комплексное число* $\beta = \sqrt[n]{\alpha}$*, что справедливо равенство*

$$\beta^n \overset{def}{=} \alpha.\bullet$$

Формула извлечения корня n-й степени из комплексного числа в тригонометрической форме записи является результатом доказательства следующей теоремы.

Теорема 1.3.6. *Корни n-й степени из комплексного числа* α *существуют и все они даются формулой*

$$\beta_k = r^{\frac{1}{n}}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right) \qquad (1.3.14)$$

при любом целом числе k.

Д о к а з а т е л ь с т в о. Для $\alpha = 0$ единственное значение $\sqrt[n]{\alpha} = 0$. Если $\alpha \neq 0$, то для числа

$$\alpha = r(\cos\varphi + i\sin\varphi)$$

и корня n-й степени

$$\beta = R(\cos\theta + i\sin\theta),$$

учитывая определение $(\beta^n = \alpha)$, имеем

$$R^n(\cos n\theta + i\sin n\theta) = r(\cos\varphi + i\sin\varphi).$$

Откуда $R^n = r$, $n\theta = \varphi + 2\pi k$, где k — целое число. Так как $r > 0$ и $R > 0$, то $R = r^{\frac{1}{n}}$. Аргумент частного получается простым делением:

$$\theta = \frac{\varphi + 2\pi k}{n}.\bullet\bullet$$

Следствие из теоремы 1.3.6. *Существует ровно n различных значений корня n-й степени из комплексного числа* $\alpha \neq 0$*, которые вычисляются по формуле (1.3.14) при* $k = 0, 1, \ldots, n-1$.

Д о к а з а т е л ь с т в о. Действительно, числа $\beta_0, \beta_1, \ldots, \beta_{n-1}$ различны, так как их аргументы

$$\theta_0 = \frac{\varphi}{n};\ \theta_1 = \frac{\varphi}{n} + \frac{2\pi}{n};\ \ldots;\ \theta_{n-1} = \frac{\varphi}{n} + \frac{2\pi(n-1)}{n}$$

различны и отличаются друг от друга менее чем на 2π. Далее, $\beta_n = \beta_0$, так как $|\beta_n| = |\beta_0|$ и $\theta = \varphi + 2\pi k$.

Аналогично

$$|\beta_{n+1}| = |\beta_1|;\ |\beta_{-1}| = |\beta_{n-1}|$$

и так далее. ••

Глава 2. Векторные пространства

2.1. Трёхмерное евклидово пространство

Понятие вектора. Считаем, что имеется интуитивное понимание размерности обычного трёхмерного пространства. Примером двумерного пространства — ***плоскости***, является учебная доска.

Под ***упорядоченной парой*** $\{A, B\}$ ***точек*** *обычного трёхмерного пространства (или плоскости) будем понимать такую пару точек, для которой указано, какая точка является первой, а какая второй.*

При таком ***соглашении*** по определению полагают, что ***вектор — это направленный отрезок*** AB, соединяющий между собой элементы ***упорядоченной пары*** $\{A, B\}$ точек пространства, отмеченных на некоторой прямой линии с выбранным направлением (числовой оси). Обозначаются векторы либо двумя большими латинскими буквами со стрелкой вверху, например $\vec{AB}$, либо одной малой латинской буквой со стрелкой вверху, например $\vec{a}$ (рис. 2.1, *а*). Точка A называется ***началом вектора*** $\vec{AB}$, а точка B — ***концом вектора*** $\vec{AB}$.

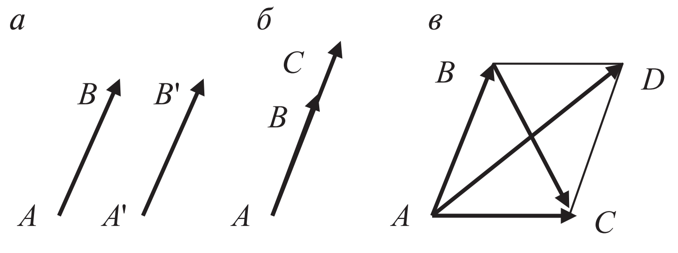


Рис. 2.1

Если соединить начальные точки векторов отрезком прямой линии и передвигать один из них вдоль этого отрезка с сохранением его направления до совпадения начальных точек обоих векторов, будем говорить о ***параллельном переносе этого вектора***. Два вектора считаются ***равными***, если они могут быть полностью совмещены как направленные отрезки путём ***параллельного переноса***. Так, на рис. 2.1, *а* векторы $\vec{AB}$ и $\vec{A'B'}$ очевидно равны.

Длиной $\left\|\vec{AB}\right\|$ ***вектора*** $\vec{AB}$ называется длина определяющего его направленного отрезка AB. ***Величина*** $\left\{\vec{x}\right\}$ ***вектора*** $\vec{x}$ по определению равна его длине, взятой со знаком «***плюс***», если его направление совпадает с выбранным направлением оси, на которую он помещён, и со знаком «***минус***» — в противном случае.

Для векторов определяются две операции — ***умножение вектора на число*** (обозначается $\alpha \cdot \vec{a}$ или $\alpha \cdot \vec{AB}$, где $\alpha \in R^1$) и ***сложение векторов*** (обозначается $\vec{a} + \vec{b}$ или $\vec{AB} + \vec{CD}$).

Сложение двух векторов осуществляется по известному из школьного курса ***правилу параллелограмма***. Например, вектор $\vec{AD} = \vec{c}$ на рис. 2.1, *в*, является суммой векторов $\vec{AB} = \vec{a}$ и $\vec{AC} = \vec{b}$. Кроме этого, легко видеть, что так как $\vec{AC} = \vec{BD}$, сложение векторов может производиться по ***правилу треугольника***: *начало вектора* $\vec{b}$ *путём параллельного переноса совмещаем с концом вектора* $\vec{a}$*, получаем вектор* $\vec{c} \equiv \vec{AD}$*, который и является суммой векторов* $\vec{a}$ *и* $\vec{b}$. Правило треугольника обобщается на правило ***многоугольника***.

Отметим, что ***операция сложения векторов является алгебраической***, так как она каждым двум векторам трёхмерного пространства (плоскости) ставит в соответствие третий вектор того же пространства (плоскости). Используя правило треугольника, легко показать, что операция сложения векторов ***коммутативна***, то есть

$$\vec{AD} = \vec{AB} + \vec{AC} = \vec{AC} + \vec{AB},$$

и ***ассоциативна***, то есть

$$\left(\vec{AB} + \vec{BD}\right) + \vec{DC} = \vec{AB} + \left(\vec{BD} + \vec{DC}\right)$$

(см. рис. 2.1, *в*). Кроме этого, операция сложения векторов имеет обратную операцию, которая называется ***вычитанием*** векторов: вектор $\vec{BC}$ — является ***разностью*** векторов $\vec{AC}$ и $\vec{AB}$. Обозначается разность векторов обычным образом:

$$\vec{BC} = \vec{AC} - \vec{AB}.$$

Если векторы равны, например,

$$\vec{AB} = \vec{AC},$$

то принимается соглашение, что их разность — это вектор, не имеющий направления (направление не определено) с длиной, равной нулю. Такой вектор называется ***вектор нуль***, или ***нуль-вектор***. Нуль-вектор можно представить как вектор, «сосредоточенный в точке».

Результат умножения вектора на число снова есть вектор, причём при умножении вектора на некоторое число $\alpha \in R^1$ его длина умножается на абсолютную величину $|\alpha|$ этого числа, направление при $\alpha > 0$ не изменяется, а при $\alpha < 0$ изменяется на противоположное (вектор ***переопрокидывается***). На рис. 2.1, *б* вектор $\vec{AC}$ есть результат умножения вектора $\vec{AB}$ на некоторое число $\alpha > 0$, что записывается так:

$$\vec{AC} = \alpha \cdot \vec{AB} \vee \vec{y} = \alpha \cdot \vec{x}.$$

Заметим, что связанные подобными соотношениями векторы называются ***коллинеарными***. Отметим также, точка в произведении вектора на число часто не ставится.

Помещая некоторый вектор на соответствующую ось и используя определение операции сложения векторов, нетрудно показать, что операция умножения вектора на число обладает следующими свойствами:

$$1 \cdot \vec{x} = \vec{x};$$

$$(\alpha + \beta)\vec{x} = \alpha\vec{x} + \beta\vec{x};$$

$$\alpha\left(\vec{x} + \vec{y}\right) = \alpha\vec{x} + \beta\vec{y};$$

$$(\alpha\beta)\vec{x} = \alpha\left(\beta\vec{x}\right).$$

В отличие от операции сложения векторов операция умножения вектора на число является ***внешней бинарной операцией*** (неалгебраической), так как определена для пары, один элемент которой принадлежит множеству векторов пространства (плоскости), а второй множеству действительных чисел.

Декартова система координат. Координаты точек и векторов. ***Ортогональной декартовой системой координат*** в пространстве (на плоскости) называется совокупность трёх (двух) пересекающихся, взаимно перпендикулярных осей и точки их пересечения — ***начала координат***. Напомним известное из средней школы построение декартовой системы координат для случая плоскости.

Проводим на плоскости две взаимно перпендикулярные прямые и обозначаем точку их пересечения буквой O (рис. 2.2). На горизонтальной и вертикальной прямых откладываем отрезки OE_1 и OE_2 одной и той же длины. Упорядочим тройку точек $\{O; E_1; E_2\}$ так, чтобы кратчайший поворот отрезка OE_1 до совмещения с отрезком OE_2 осуществлялся против часовой стрелки — это направление принимаем за ***положительное***.

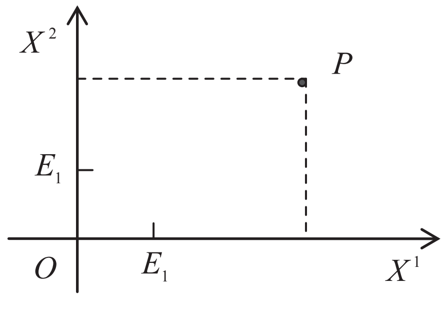


Рис. 2.2

Прямую, проходящую через точки O и E_1, назовём осью OX^1, а прямую, проходящую через точки O и E_2, назовём осью OX^2. На каждой оси поместим действительные числа, отмечая их положение пропорционально величине соответствующего числа: число 0 поместим в точку O, число 1 в точки E_1 и E_2, и так далее. Теперь каждой упорядоченной паре чисел $\{x^1; x^2\}$ ставится во взаимно однозначное соответствие точка плоскости, в проекции которой на оси OX^k находится число x^k $(k = 1, 2)$. Индекс вверху — это номер. Смысл выяснится ниже.

Таким образом, между точками плоскости и парами действительных чисел — их координатами x^1 на оси OX^1 и x^2 на оси OX^2, установлено биективное соответствие, то есть плоскость является прямым произведением $R^1 \times R^1 = R^2$. По этой причине плоскость называется пространством R^2 (читается «пространство эр два»). Поясним, что речь идёт о ***модели плоскости***, но слово модель обычно опускают.

Если ортогональная декартова система координат введена в трёхмерном пространстве, то точке P ставятся в соответствие три действительных числа — ***декартовы координаты***, которые по определению равны ***величинам координатных проекций*** $\vec{OA}$, $\vec{OB}$ ***и*** $\vec{OC}$ ***радиус-вектора*** $\vec{OP}$ ***точки*** P, началом которого является начало системы координат, а концом — точка P. На рис. 2.3 величины x^1, x^2, x^3 — это координаты точки P и её радиус-вектора $\vec{OP}$. ***Началу*** системы координат O ставится в соответствие ***нуль-вектор*** — направленный отрезок с совпадающими начальной и конечной точками и, следовательно, имеющий ***нулевую длину***. Для нуль-вектора ***направление не определено***.

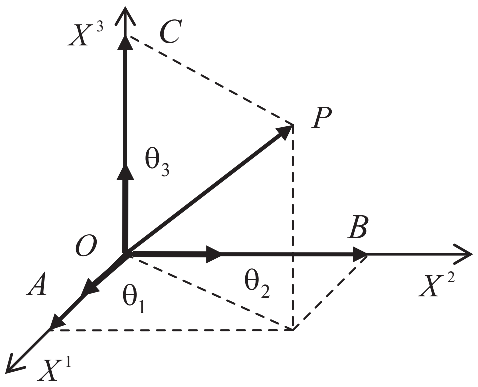


Рис. 2.3

Применяя теорему Пифагора, нетрудно показать, что длина или ***норма*** радиус-вектора $\vec{x}$ в трёхмерном пространстве вычисляется по формуле

$$\left\|\vec{x}\right\| \stackrel{def}{=} \sqrt{\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2}. \tag{2.1.1}$$

В случае плоскости в формуле отсутствует третья координата.

Направление вектора $\vec{x}$ определяется тремя углами $\theta_1, \theta_2, \theta_3$, которые этот вектор образует с положительными полуосями координат OX^1, OX^2, OX^3, причём $0 \le \theta_i \le \pi \ (i = 1, 2, 3)$, или ***направляющими косинусами***:

$$\cos\theta_i = \frac{x^i}{\left\| \vec{x} \right\|} \ (i = 1, 2, 3).$$

Очевидно, что обычное трёхмерное пространство является произведением

$$R^1 \times R^1 \times R^1 = R^3.$$

Представление радиус-вектора в виде разложения по базисным векторам. Рассмотрим случай радиус-векторов, причём для простоты и наглядности ограничимся случаем векторов на плоскости. Сначала рассмотрим представление радиус-вектора через его координаты.

На рис. 2.4 представлен радиус-вектор $\vec{x}$ в декартовой системе координат X^1OX^2. Векторы $\vec{OA}$ и $\vec{OB}$ — это ***составляющие*** вектора $\vec{x}$ по осям координат. Очевидно, что в соответствии с правилом параллелограмма

$$\vec{x} = \vec{OA} + \vec{OB}.$$

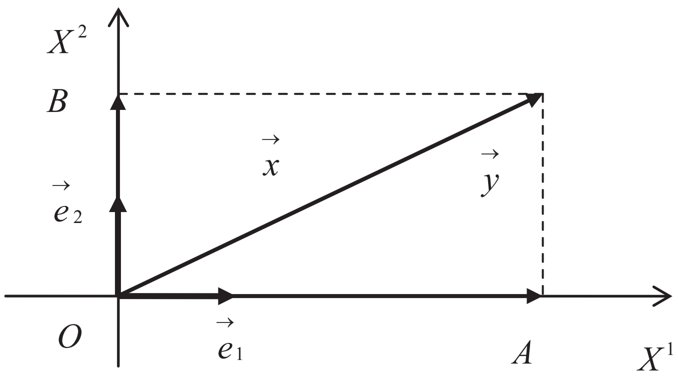


Рис. 2.4

Векторы $\vec{e}_1$ и $\vec{e}_2$, расположенные на осях OX^1 и OX^2 соответственно, направленные в положительных направлениях осей и имеющие единичную длину, образуют ***канонический базис*** декартовой системы координат (глубокий смысл этого важнейшего понятия будет выяснен в ближайшее время).

Обозначим величины составляющих вектора $\vec{x}$ по соответствующей оси фигурными скобками $x^1 = \left\{\vec{OA}\right\}$, $x^2 = \left\{\vec{OB}\right\}$. Тогда очевидно, что

$$\vec{OA} = \left\{\vec{OA}\right\}\vec{e}_1 \equiv x^1\vec{e}_1, \vec{OB} = \left\{\vec{OB}\right\}\vec{e}_2 \equiv x^2\vec{e}_2.$$

Теперь вектор $\vec{x}$ с учётом правила параллелограмма представляется в виде следующей суммы:

$$\vec{x} = \vec{OA} + \vec{OB} = \left\{\vec{OA}\right\}\vec{e}_1 + \left\{\vec{OB}\right\}\vec{e}_2 \equiv x^1\vec{e}_1 + x^2\vec{e}_2. \qquad (2.1.2)$$

Формула (2.1.2) является важнейшей для дальнейшего изложения и называется ***разложением вектора*** $\vec{x}$ по базисным векторам $\left\{\vec{e}_1, \vec{e}_2\right\}$.

Переходя к случаю трёхмерного пространства, можем по аналогии с разложением (2.1.2) написать следующее равенство:

$$\vec{x} = \vec{OA} + \vec{OB} + \vec{OC} =$$

$$= \left\{\vec{OA}\right\}\vec{e}_1 + \left\{\vec{OB}\right\}\vec{e}_2 + \left\{\vec{OC}\right\}\vec{e}_3 \equiv x^1\vec{e}_1 + x^2\vec{e}_2 + x^3\vec{e}_3. \qquad (2.1.3)$$

В качестве иллюстрации этого равенства можно обратиться к рис. 2.3, на котором векторы ***базиса*** $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ указаны, но не подписаны.

Величины составляющих радиус-вектора $\vec{x}$ по осям системы координат $\vec{OA}$ и $\vec{OB}$ в равенстве (2.1.2) и $\vec{OA}$, $\vec{OB}$ и $\vec{OC}$ в равенстве (2.1.3) совпадают, очевидно, с координатами конечной точки вектора и называются также ***координатами*** радиус-вектора $\vec{x}$.

В силу того, что векторы $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ являются обычными векторами пространства R^3, из равенства (2.1.3) вытекают следующие разложения:

$$\vec{e}_1 = q_1^1 \vec{e}_1 + q_1^2 \vec{e}_2 + q_1^3 \vec{e}_3,$$
$$\vec{e}_2 = q_2^1 \vec{e}_1 + q_2^2 \vec{e}_2 + q_2^3 \vec{e}_3, \tag{2.1.4}$$
$$\vec{e}_3 = q_3^1 \vec{e}_1 + q_3^2 \vec{e}_2 + q_3^3 \vec{e}_3.$$

Очевидно, что координаты

$$q_1^1 = 1,\ q_2^2 = 1,\ q_3^3 = 1, \tag{2.1.5}$$

а остальные координаты равны нулю. Аналогичные равенства справедливы, естественно, и для пространства R^2. Кроме этого, принимается соглашение, что нуль-вектор имеет ***нулевые координаты***:

$$\vec{0} = 0\vec{e}_1 + 0\vec{e}_2 + 0\vec{e}_3. \tag{2.1.6}$$

Выражение операций над векторами через их координаты. Сначала выясним, как связаны координаты равных векторов. Для этого снова обратимся к рис. 2.4. На нём изображены равные векторы $\vec{x}$ и $\vec{y}$, которые в соответствии с данным выше определением равенства векторов совмещены как направленные отрезки. Их составляющие по осям OX^1 и OX^2 очевидно также совпадают как направленные отрезки. Следовательно, если $\vec{x} = \vec{y}$, то

$$x^1 = \left\{\vec{OA}\right\} = y^1,\ x^2 = \left\{\vec{OB}\right\} = y^2. \tag{2.1.7}$$

Аналогичные выражению (2.1.7) равенства справедливы, конечно, и в случае трёхмерного пространства, а именно, если $\vec{x} = \vec{y}$, то

$$x^1 = y^1,\ x^2 = y^2,\ x^3 = y^3. \tag{2.1.7'}$$

Итак, ***если векторы равны, то и их координаты относительно декартовой системы координат также равны***.

Справедливо, очевидно, и обратное, а именно: ***если координаты двух векторов относительно некоторой фиксированной системы координат равны, то они равны*** в том смысле, что их можно полностью совместить как направленные отрезки путём параллельного переноса.

Рассмотрим теперь операцию сложения радиус-векторов по правилу параллелограмма для случая плоскости, который представлен на рис. 2.5. Записывая для векторов $\vec{x}$, $\vec{y}$ и $\vec{z}$ разложения вида (2.1.2) и учитывая, что вектор $\vec{z}$ есть результат сложения векторов $\vec{x}$ и $\vec{y}$ по правилу параллелограмма, имеем:

$$\vec{z} = \vec{x} + \vec{y} = \left(x^1 \vec{e}_1 + x^2 \vec{e}_2 \right) + \left(y^1 \vec{e}_1 + y^2 \vec{e}_2 \right) =$$
$$= \left(x^1 + y^1 \right) \vec{e}_1 + \left(x^2 + y^2 \right) \vec{e}_2 = z^1 \vec{e}_1 + z^2 \vec{e}_2.$$

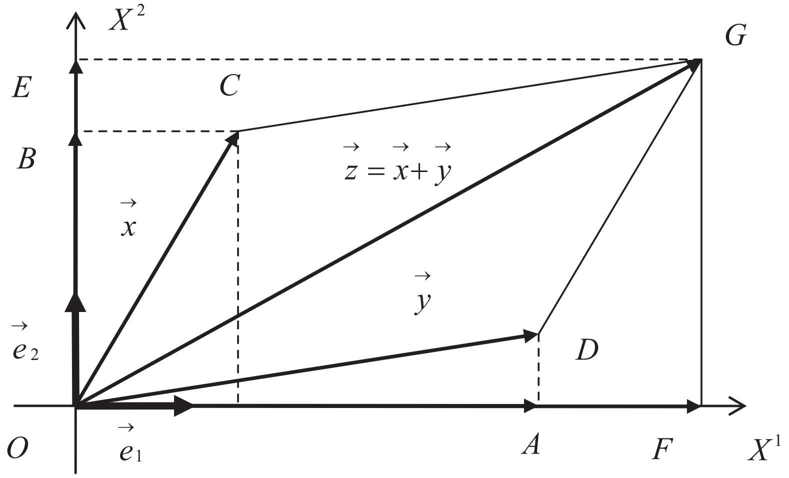


Рис. 2.5

Учитывая равенства (2.1.7), получаем

$$z^1 = x^1 + y^1,\ z^2 = x^2 + y^2. \tag{2.1.8}$$

Таким образом, ***при сложении радиус-векторов их координаты относительно канонического базиса складываются***.

Конечно, аналогичный результат для суммы векторов справедлив с учётом выражений (2.1.3), (2.1.7’) и в пространстве R^3:

$$\vec{z} = \vec{x} + \vec{y} = \left(x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3 \right) + \left(y^1 \vec{e}_1 + y^2 \vec{e}_2 + y^3 \vec{e}_3 \right) =$$
$$= \left(x^1 + y^1 \right) \vec{e}_1 + \left(x^2 + y^2 \right) \vec{e}_2 + \left(x^3 + y^3 \right) \vec{e}_3 = z^1 \vec{e}_1 + z^2 \vec{e}_2 + z^3 \vec{e}_3, \tag{2.1.9}$$

$$z^1 = x^1 + y^1,\ z^2 = x^2 + y^2,\ z^3 = x^3 + y^3. \tag{2.1.10}$$

Таким образом, ***при сложении векторов их соответствующие координаты относительно канонического базиса декартовой системы координат складываются***.

Рассмотрим умножение вектора на число (рис. 2.6), начиная снова для простоты со случая пространства R^2. Рассмотрим только координату вектора по оси OX^1. Из подобия треугольников OCD и OAB, а также из того, что длины векторов $\vec{x}$ и $\vec{z}$ пропорциональны, следует, что $z^1 = \alpha x^1$, $z^2 = \alpha x^2$. Поэтому, получаем, что

$$\vec{z} = z^1 \vec{e}_1 + z^2 \vec{e}_2 = \alpha \vec{x} = \alpha\left(x^1 \vec{e}_1 + x^2 \vec{e}_2 \right) = \alpha x^1 \vec{e}_1 + \alpha x^2 \vec{e}_2. \quad (2.1.11)$$

Аналогичное равенство имеет место и в случае пространства R^3:

$$\vec{z} = z^1 \vec{e}_1 + z^2 \vec{e}_2 + z^3 \vec{e}_3 = \alpha \vec{x} =$$

$$= \alpha\left(x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3 \right) = \alpha x^1 \vec{e}_1 + \alpha x^2 \vec{e}_2 + \alpha x^3 \vec{e}_3. \quad (2.1.12)$$

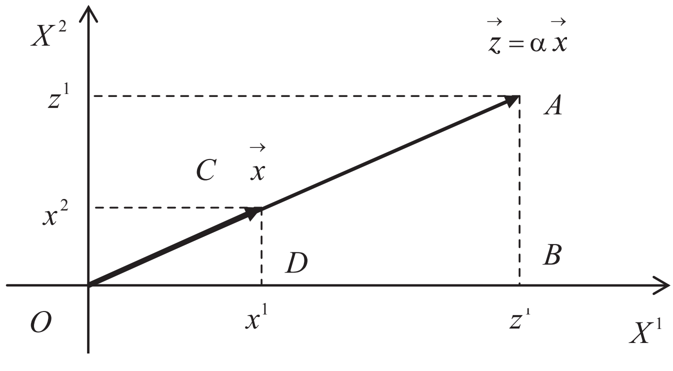


Рис. 2.6

Из равенств (2.1.11) и (2.1.12) следует, что

$$z^1 = \alpha \cdot x^1,\ z^2 = \alpha \cdot x^2,\ z^3 = \alpha \cdot x^3. \quad (2.1.13)$$

Таким образом, ***при умножении вектора на число все его координаты относительно канонического базиса декартовой системы координат умножаются на это число***.

Установим важное свойство системы векторов $\left\{ \vec{e}_1, \vec{e}_2, \vec{e}_3 \right\}$ (рассматриваем сразу пространство R^3). Умножая каждый из векторов на соответствующий множитель $\{\alpha_1, \alpha_2, \alpha_3\}$, складывая получившиеся векторы и приравнивая результат нулевому вектору, получаем:

$$\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \alpha_3 \vec{e}_3 = \vec{0}. \tag{2.1.14}$$

Учитывая выражения (2.1.6) и (2.1.7'), видим, что

$$\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0. \tag{2.1.15}$$

Векторная величина в левой части выражения (2.1.14) называется ***линейной комбинацией векторов системы*** $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ ***с коэффициентами*** $\{\alpha_1, \alpha_2, \alpha_3\}$. Равенства (2.1.15) означают, что линейная комбинация векторов канонического базиса имеет своим значением нулевой вектор только при их выполнении. Это свойство относится ко всей системе векторов канонического базиса и называется ***линейной независимостью*** системы. Свойство системы векторов быть линейно независимой выполняется не только для векторов канонического базиса, но и для бесконечного числа систем трёх некомпланарных (в случае плоскости — для бесконечного числа систем неколлинеарных) векторов.

2.2. Скалярное произведение. Векторное и смешанное произведения

Скалярное произведение векторов и его свойства, ортогональность. Пусть в трёхмерном пространстве R^3 задана упорядоченная пара $\left\{\vec{x}, \vec{y}\right\}$, векторы которой представлены своими разложениями по векторам канонического базиса $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$:

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3, \; \vec{y} = y^1 \vec{e}_1 + y^2 \vec{e}_2 + y^3 \vec{e}_3.$$

Аналогичные разложения, содержащие два слагаемых, справедливы, конечно, и в двумерном пространстве R^2.

Отображение $\varphi\left(\vec{x}, \vec{y}\right): R^3 \times R^3 \to R^1$*, ставящее в соответствие каждой упорядоченной паре трёхмерных (двумерных) векторов однозначно* *определённое действительное число, называется* ***действительной функцией двух векторных аргументов (переменных)*** $\vec{x}$ ***и*** $\vec{y}$.

Действительная функция двух векторных аргументов, значение которой находится в каноническом базисе пространства R^3 по правилу

$$\left(\vec{x}, \vec{y}\right) \overset{def}{=} x^1 y^1 + x^2 y^2 + x^3 y^3 = \sum_{r=1}^{3} x^k y^k, \tag{2.2.1}$$

называется ***скалярным произведением*** векторов $\vec{x}$ и $\vec{y}$. В формуле (2.2.1), как это принято исторически, символ функции опущен.

Подчеркнём, что формула (2.2.1) определяет ***скалярную функцию двух векторных аргументов***, а «скалярное произведение» — ***это сложившееся исторически название данной функции***. Поэтому термин «скалярное произведение» не следует понимать, как обозначение операции, это неверно.

Если $\vec{x} = \vec{y}$, то величина

$$\left(\vec{x}, \vec{x}\right) = x^1 x^1 + x^2 x^2 + x^3 x^3 = \sum_{r=1}^{3} \left(x^k\right)^2 \tag{2.2.1'}$$

называется скалярным квадратом радиус-вектора $\vec{x}$. Очевидно, что теперь можно записать ***норму*** (длину) радиус-вектора $\vec{x}$ в виде:

$$\left\|\vec{x}\right\| \overset{def}{=} \sqrt{\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2} = \sqrt{\left(\vec{x}, \vec{x}\right)}. \tag{2.2.2}$$

Из определения (2.2.1) нетрудно получить следующие ***свойства скалярного произведения:***

$$\begin{aligned}
&1)\ \left(\vec{x}, \vec{y}\right) = \left(\vec{y}, \vec{x}\right);\\
&2)\ \left(\alpha \cdot \vec{x}, \vec{y}\right) = \alpha \cdot \left(\vec{x}, \vec{y}\right), \left(\vec{x}, \alpha \cdot \vec{y}\right) = \alpha \cdot \left(\vec{x}, \vec{y}\right);\\
&3)\ \left(\vec{x} + \vec{y}, \vec{z}\right) = \left(\vec{x}, \vec{z}\right) + \left(\vec{y}, \vec{z}\right), \left(\vec{x}, \vec{y} + \vec{z}\right) = \left(\vec{x}, \vec{y}\right) + \left(\vec{x}, \vec{z}\right);\\
&4)\ \left(\forall \vec{x} \neq \vec{0}\right)\left(\vec{x}, \vec{x}\right) > 0.
\end{aligned} \tag{2.2.3}$$

Здесь $\alpha \in R^1$, то есть действительное число. Рекомендуется провести вывод этих свойств скалярного произведения самостоятельно.

Первое свойство выражает тот факт, что скалярное произведение ***симметрично*** относительно перестановки аргументов. Второе свойство означает, что скалярное произведение является ***однородным*** относительно каждого из своих аргументов. Третье свойство выражает тот факт, что скалярное произведение ***аддитивно*** относительно каждого из своих аргументов. И, наконец, четвёртое свойство означает, что скалярное произведение ***невырождено***.

Вектор, ***длина которого равна единице***, называется ***нормированным***. Каждый ненулевой вектор можно ***нормировать***, умножив его на ***нормирующий множитель***

$$\lambda = \left\| \vec{x} \right\|^{-1}.$$

Действительно, пусть $\lambda \cdot \vec{x}$, где $\lambda \in R^1$ — вектор с длиной, равной 1, следовательно, его скалярный квадрат в соответствии с формулой (2.2.2) также равен единице. Учитывая свойства (2.2.17), запишем скалярный квадрат вектора $\lambda \cdot \vec{x}$:

$$\left(\lambda \cdot \vec{x}, \lambda \cdot \vec{x} \right) = \lambda^2 \cdot \left(\vec{x}, \vec{x} \right) = 1.$$

Откуда имеем (длина всегда положительна):

$$\lambda^2 = \left(\vec{x}, \vec{x} \right)^{-1} \Rightarrow \lambda = \left(\vec{x}, \vec{x} \right)^{-1/2} = \left\| \vec{x} \right\|^{-1}.$$

Очевидно, что вектор

$$\vec{e}_x \overset{def}{=} \left\| \vec{x} \right\|^{-1} \cdot \vec{x} \tag{2.2.4}$$

имеет норму, равную 1. Этот вектор называется ***ортом вектора*** $\vec{x}$.

Два вектора

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3, \; \vec{y} = y^1 \vec{e}_1 + y^2 \vec{e}_2 + y^3 \vec{e}_3,$$

называются ***ортогональными***, если их скалярное произведение равно нулю, то есть выполняется следующее условие:

$$\left(\vec{x}, \vec{y} \right) = x^1 y^1 + x^2 y^2 + x^3 y^3 = 0. \tag{2.2.5}$$

Теорема 2.2.1. *Скалярное произведение векторов в трёхмерном (двумерном) евклидовом пространстве равно произведению их норм (длин) на косинус угла между векторами*

$$\left(\vec{x}, \vec{y}\right) = \left\|\vec{x}\right\| \cdot \left\|\vec{y}\right\| \cdot \cos\varphi. \tag{2.2.6}$$

Д о к а з а т е л ь с т в о. Пусть $\vec{x}, \vec{y}$ — данные ненулевые векторы, а φ — угол между ними. Тогда по свойствам (2.2.3) имеем:

$$\left(\vec{x}+\vec{y}, \vec{x}+\vec{y}\right) = \left(\vec{x}, \vec{x}\right) + 2\left(\vec{x}, \vec{y}\right) + \left(\vec{y}, \vec{y}\right).$$

Учитывая выражение (2.2.4), получаем:

$$\left\|\vec{x}+\vec{y}\right\|^2 = \left\|\vec{x}\right\|^2 + 2\left(\vec{x}, \vec{y}\right) + \left\|\vec{y}\right\|^2.$$

Видим, что скалярное произведение векторов $\left(\vec{x}, \vec{y}\right)$ выражается через их нормы (длины) и, следовательно, не зависит от выбора системы координат.

Выберем декартову ортогональную систему координат, как показано на рис. 2.7 (ось OX^3 направлена перпендикулярно плоскости страницы). Тогда для координат векторов имеем:

$$x^1 = \left\|\vec{x}\right\|;\ x^2 = 0;\ x^3 = 0;\ y^1 = \left\|\vec{y}\right\|\cos\varphi;\ y^2 = \left\|\vec{y}\right\|\sin\varphi;\ y^3 = 0.$$

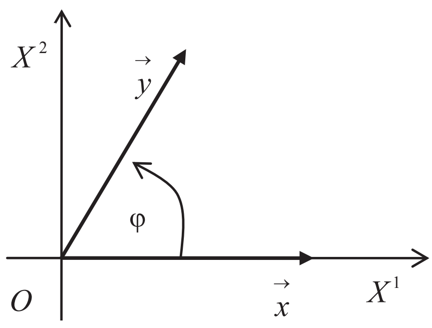


Рис. 2.7

Подставляя выражения координат в выражение (2.2.1), приходим к формуле (2.2.6):

$$\left(\vec{x}, \vec{y}\right) = \left\|\vec{x}\right\| \cdot \left\|\vec{y}\right\| \cdot \cos\varphi + 0 \cdot \left\|\vec{y}\right\| \cdot \sin\varphi + 0 \cdot 0 = \left\|\vec{x}\right\| \cdot \left\|\vec{y}\right\| \cdot \cos\varphi. \bullet\bullet$$

Из теоремы виден смысл термина «***ортогональность***»: *если векторы перпендикулярны* (с точки зрения элементарной геометрии)*, то их скалярное произведение равно нулю, так как* $\cos\frac{\pi}{2} = 0$*, то есть векторы ортогональны*. Справедливо и обратное: *если скалярное произведение векторов равно нулю (векторы ортогональны), то они перпендикулярны*. Таким образом, ортогональность и перпендикулярность — ***синонимы*** одного и того же свойства векторов.

Нетрудно видеть, что ***векторы канонического базиса попарно ортогональны***, что проверяется непосредственным вычислением. Кроме этого, векторы $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ ***нормированы***. Если векторы системы $\left\{\vec{x}_1, \vec{x}_2, \vec{x}_3\right\}$ попарно ортогональны (ортонормированны)*,* то эта система называется ***ортогональной*** (***ортонормированной***).

Любая ортогональная (ортонормированная) система векторов линейно независима. Действительно, если система $\left\{\vec{x}_1, \vec{x}_2, \vec{x}_3\right\}$ ортогональна, то, находя значение скалярного произведения обеих частей векторного равенства

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \alpha_3 \vec{x}_3 = \vec{0}$$

и векторов $\vec{x}_1, \vec{x}_2, \vec{x}_3$ последовательно, получаем

$$\alpha_1\left(\vec{x}_1, \vec{x}_1\right) = 0, \alpha_2\left(\vec{x}_2, \vec{x}_2\right) = 0, \alpha_3\left(\vec{x}_3, \vec{x}_3\right) = 0.$$

Так как векторы ненулевые, то есть их скалярные квадраты нулю не равны, то

$$\alpha_1 = \alpha_2 = \alpha_3 = 0,$$

что и доказывает линейную независимость системы векторов.

Отметим, что обычное трёхмерное пространство с введённой ортогональной декартовой системой координат является ***третьей степенью множества действительных чисел***

$$R^1 \times R^1 \times R^1 = R^3,$$

поэтому обозначение R^3 (читается ***эр три***) вполне оправдано.

Определение 2.2.1. *Множество трёхмерных векторов с определёнными выше внутренней бинарной (алгебраической) операцией —* ***сложением векторов****, внешней бинарной операцией —* ***умножением вектора на число*** *и заданной на нём функцией —* ***скалярным произведением векторов****, называется* ***трёхмерным евклидовым пространством*** R^3**.•**

Измерения в пространстве R^3. По теореме 2.2.1 (формула (2.2.6))

$$\cos\varphi = \frac{1}{\left\|\vec{x}\right\| \cdot \left\|\vec{y}\right\|}\left(\vec{x}, \vec{y}\right),$$

откуда с учётом определения (2.2.5) и формулы (2.2.2) получаем формулу для вычисления ***косинуса угла между векторами в пространстве*** R^3:

$$\cos\varphi = \frac{x^1 y^1 + x^2 y^2 + x^3 y^3}{\sqrt{\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2}\sqrt{\left(y^1\right)^2 + \left(y^2\right)^2 + \left(y^3\right)^2}}. \qquad (2.2.7)$$

Величина, определённая формулой

$$\Pr_{\vec{y}}\left\{\vec{x}\right\} \overset{def}{=} \left\|\vec{x}\right\| \cdot \cos\varphi = \frac{\left(\vec{x}, \vec{y}\right)}{\left\|\vec{y}\right\|} = \frac{x^1 y^1 + x^2 y^2 + x^3 y^3}{\sqrt{\left(y^1\right)^2 + \left(y^2\right)^2 + \left(y^3\right)^2}} \qquad (2.2.8)$$

называется ***проекцией вектора*** $\vec{x}$ ***на направление вектора*** $\vec{y}$.

До сих пор проводили все рассмотрения для радиус-векторов. Для определения расстояния между точками пространства R^3 нам потребуется понятие длины (нормы) произвольного вектора.

Рассмотрим, для простоты, двумерный случай (см. рис. 2.8). По определению, координатами вектора $\vec{z} = \vec{x} - \vec{y}$, начальной точкой которого является конечная точка вектора $\vec{y}$, а конечной точкой — конечная точка вектора $\vec{x}$, называются ***величины координатных проекций*** $\overrightarrow{M'N'}$ ***и*** $\overrightarrow{M''N''}$ ***вектора*** $\vec{x} - \vec{y}$, *то есть*

$$z^1 = \left\{\overrightarrow{M'N'}\right\}, \; z^2 = \left\{\overrightarrow{M''N''}\right\}.$$

Из рис. 2.8 легко видеть, что координаты z^1 и z^2 вектора $\vec{x} - \vec{y}$ равны

$$z^1 = x^1 - y^1, \; z^2 = x^2 - y^2.$$

Поэтому расстояние $\rho(M, N)$ между начальной точкой M вектора $\vec{x}-\vec{y}$ и его конечной точкой N равно длине вектора $\vec{MN} = \vec{x}-\vec{y}$ и вычисляется так

$$\rho(M, N) = \left\|\vec{MN}\right\| \equiv \left\|\vec{x}-\vec{y}\right\| = \sqrt{\left(x^1 - y^1\right)^2 + \left(x^2 - y^2\right)^2}.$$

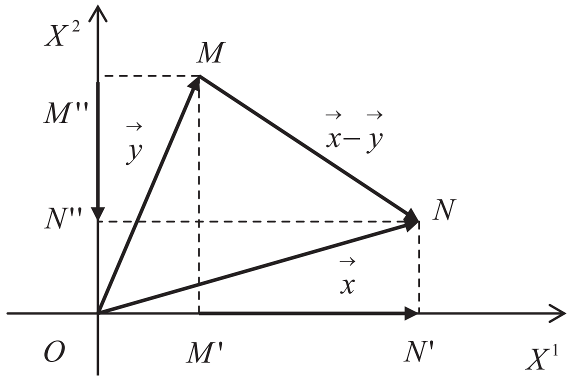


Рис. 2.8

Очевидно, что в трёхмерном случае расстояние между двумя точками M и N пространства R^3 вычисляется по формуле

$$\rho(M, N) = \sqrt{\left(x^1 - y^1\right)^2 + \left(x^2 - y^2\right)^2 + \left(x^3 - y^3\right)^2}, \tag{2.2.9}$$

где

$$\vec{x} = x^1\vec{e}_1 + x^2\vec{e}_2 + x^3\vec{e}_3,\ \vec{y} = y^1\vec{e}_1 + y^2\vec{e}_2 + y^3\vec{e}_3$$

— радиус-векторы точек N и M соответственно.

Векторное и смешанное произведения векторов. В трёхмерном евклидовом пространстве кроме операции сложения векторов существует ещё одна алгебраическая операция — ***векторное умножение векторов***, наличие которой превращает трёхмерное пространство в ***алгебру***.

Определение 2.2.2. *Тройка некомпланарных векторов* $\left(\vec{x}, \vec{y}, \vec{z}\right)$ *называется* ***правой (левой)****, если эти векторы располагаются так, как могут быть расположены, соответственно, большой, не согнутые указательный и средний пальцы правой (левой) руки.* •

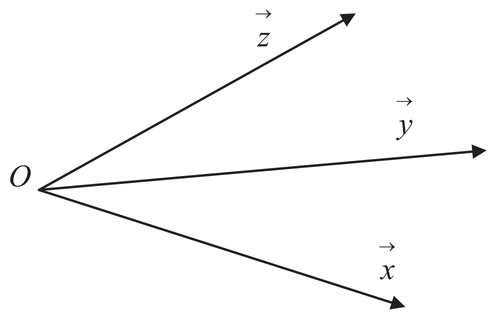


Рис. 2.9

Например, тройка векторов $\left(\vec{x}, \vec{y}, \vec{z}\right)$ на рис. 2.9 — правая. Из любых трёх некомпланарных векторов $\vec{x}, \vec{y}, \vec{z}$ можно составить шесть троек:

$$\left\{\left(\vec{x}, \vec{y}, \vec{z}\right), \left(\vec{y}, \vec{z}, \vec{x}\right), \left(\vec{z}, \vec{x}, \vec{y}\right)\right\},$$

$$\left\{\left(\vec{y}, \vec{x}, \vec{z}\right), \left(\vec{x}, \vec{z}, \vec{y}\right), \left(\vec{z}, \vec{y}, \vec{x}\right)\right\}.$$

Первые три тройки векторов того же наименования, что и тройка $\left(\vec{x}, \vec{y}, \vec{z}\right)$, то есть ***правые***; следующие три тройки — ***левые***.

Определение 2.2.3. *Декартова система координат называется* ***правой (левой)*** *системой координат, если тройка базисных векторов правая (левая).* •

Дальше рассматриваем только правые системы координат.

Определение 2.2.4. *Пусть дана упорядоченная пара векторов* $\left\{\vec{x}, \vec{y}\right\}$. *Поставим этим векторам в соответствие третий вектор* $\vec{z}$, *потребовав выполнения условий:*

1) *вектор* $\vec{z}$ *ортогонален векторам* $\vec{x}, \vec{y}$;

2) *тройка векторов* $\left(\vec{x}, \vec{y}, \vec{z}\right)$ *правая*;

3) *норма (длина)* $\left\|\vec{z}\right\|$ *вектора* $\vec{z}$ *численно равна площади S параллелограмма, построенного на приведённых к общему началу векторах* $\vec{x}, \vec{y}$ *(рис. 2.10)*;

4) *Если векторы* $\vec{x}, \vec{y}$ *коллинеарны, то полагаем* $\left\|\vec{z}\right\| = 0$.

Построенное соответствие является ***алгебраической операцией*** *над векторами в пространстве* R^3 *и называется операцией* ***векторного умножения вектора*** $\vec{x}$ ***на вектор*** $\vec{y}$**.** •

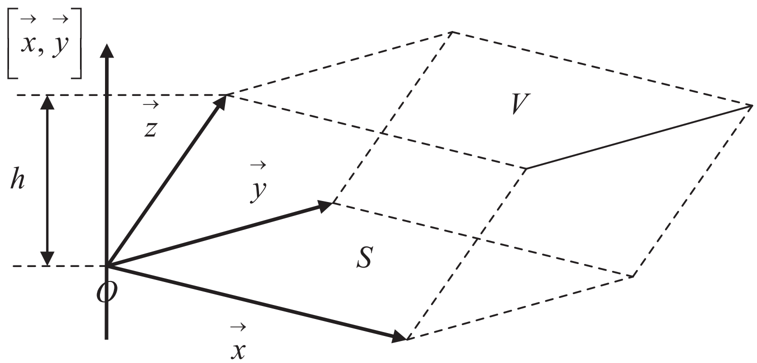


Рис. 2.10

Вектор $\vec{z}$ из определения 2.2.4 называется ***векторным произведением вектора*** $\vec{x}$ ***на вектор*** $\vec{y}$. Для операции векторного умножения используется одно из следующих обозначений:

$$\vec{z} = \left[\vec{x}, \vec{y}\right] = \left[\vec{x} \times \vec{y}\right] = \left(\vec{x} \times \vec{y}\right) = \vec{x} \times \vec{y}.$$

Пусть $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ — базис декартовой системы координат, тогда по определению 2.2.4 имеем:

$$\begin{aligned}
&\left[\vec{e}_1, \vec{e}_1\right] = \vec{0}, \left[\vec{e}_1, \vec{e}_2\right] = \vec{e}_3, \left[\vec{e}_1, \vec{e}_3\right] = -\vec{e}_2;\\
&\left[\vec{e}_2, \vec{e}_1\right] = -\vec{e}_3, \left[\vec{e}_2, \vec{e}_2\right] = \vec{0}, \left[\vec{e}_2, \vec{e}_3\right] = \vec{e}_1;\\
&\left[\vec{e}_3, \vec{e}_1\right] = \vec{e}_2, \left[\vec{e}_3, \vec{e}_2\right] = -\vec{e}_1, \left[\vec{e}_3, \vec{e}_3\right] = \vec{0}.
\end{aligned} \tag{2.2.10}$$

Из формул (2.2.10) следует, что ***операция векторного умножения, являясь алгебраической, не является коммутативной***. Иногда говорят, что операция векторного умножения векторов ***антикоммутативна***.

Как будет видно из дальнейшего изложения, операция векторного умножения векторов имеет большие приложения в геометрии. Однако приложениями в геометрии роль операции векторного умножения не исчерпывается — весьма широко данная операция применяется в теории физических полей.

Рассмотрим теперь функцию трёх векторных аргументов, определённую на каждой тройке векторов трёхмерного пространства $\left(\vec{x}, \vec{y}, \vec{z}\right)$, приложенных к общей точке O. Если векторы ***некомпланарны***, то есть связаны, например, соотношением $\vec{z} = \vec{x} + \vec{y}$, то такая тройка векторов определяет некоторый параллелепипед, объём которого обозначим $V\left(\vec{x}, \vec{y}, \vec{z}\right)$. Значение этой функции вычисляется в соответствии со следующим определением.

Определение 2.2.5. ***Ориентированным объёмом*** $V^{\pm}\left(\vec{x}, \vec{y}, \vec{z}\right)$ *параллелепипеда, построенного на приведённых к общему началу векторах* $\vec{x}, \vec{y}, \vec{z}$, *называется его объём, взятый со знаком* ***плюс****, если тройка* $\left(\vec{x}, \vec{y}, \vec{z}\right)$ *правая, и взятый со знаком* ***минус****, если тройка векторов* $\left(\vec{x}, \vec{y}, \vec{z}\right)$ *левая.* •

Определение 2.2.6. *Если сначала производится векторное умножение вектора* $\vec{x}$ *на вектор* $\vec{y}$, *а затем вычисляется скалярное произведение вектора* $\left[\vec{x}, \vec{y}\right]$ *на вектор* $\vec{z}$, *то полученное число* $\left(\left[\vec{x}, \vec{y}\right], \vec{z}\right)$ *называется* ***смешанным произведением векторов*** $\vec{x}, \vec{y}, \vec{z}$. •

Теорема 2.2.2. *Смешанное произведение* $\left(\left[\vec{x}, \vec{y}\right], \vec{z}\right)$ *равно ориентированному объёму* $V^{\pm}\left(\vec{x}, \vec{y}, \vec{z}\right)$ *параллелепипеда, построенного на приведённых к общему началу векторах* $\vec{x}, \vec{y}, \vec{z}$.

Д о к а з а т е л ь с т в о. Пусть векторы $\vec{x}$ и $\vec{y}$ неколлинеарны. Обозначим через S площадь параллелограмма, построенного на этих векторах. Тогда по определению скалярного произведения имеем:

$$\left(\left[\vec{x},\vec{y}\right],\vec{z}\right)=\left\|\left[\vec{x},\vec{y}\right]\right\|\cdot\left\|\vec{z}\right\|\cdot\cos\left\{\left[\vec{x},\vec{y}\right],\vec{z}\right\}=S\cdot\operatorname{Pr}_{\left[\vec{x},\vec{y}\right]}\left\{\vec{z}\right\}. \quad (2.2.11)$$

Если векторы $\vec{x}, \vec{y}, \vec{z}$ некомпланарны, проекция $\operatorname{Pr}_{\left[\vec{x},\vec{y}\right]}\left\{\vec{z}\right\}$ с точностью до знака равна высоте h параллелепипеда, построенного на приведённых к общему началу векторах $\vec{x}, \vec{y}, \vec{z}$, причём основанием последнего служит параллелограмм, построенный на векторах $\vec{x}, \vec{y}$ (см. рис. 2.10). Таким образом, правая часть выражения (2.2.11) с точностью до знака есть объём параллелепипеда, построенного на векторах $\vec{x}, \vec{y}, \vec{z}$. Когда

$$\operatorname{Pr}_{\left[\vec{x},\vec{y}\right]}\left\{\vec{z}\right\}=+h,$$

векторы $\left[\vec{x},\vec{y}\right]$ и $\vec{z}$ лежат по одну сторону от плоскости, образованной векторами $\vec{x}$ и $\vec{y}$, и тройка векторов $\left(\vec{x},\vec{y},\vec{z}\right)$ правая.

Если векторы $\left[\vec{x},\vec{y}\right]$ и $\vec{z}$ лежат по разные стороны от плоскости, определяемой векторами $\vec{x}$ и $\vec{y}$, то

$$\operatorname{Pr}_{\left[\vec{x},\vec{y}\right]}\left\{\vec{z}\right\}=-h$$

и тройка $\left(\vec{x},\vec{y},\vec{z}\right)$ левая.

Если векторы $\vec{x}, \vec{y}, \vec{z}$ компланарны, то

$$\operatorname{Pr}_{\left[\vec{x},\vec{y}\right]}\left\{\vec{z}\right\}=0$$

и искомый объём параллелепипеда полагается равным нулю. ••

Следствие из теоремы 2.2.2. *Для любых векторов* $\vec{x}, \vec{y}, \vec{z}$ *справедливо соотношение:*

$$\left(\left[\vec{x},\vec{y}\right],\vec{z}\right)=\left(\vec{x},\left[\vec{y},\vec{z}\right]\right). \quad (2.2.12)$$

Д о к а з а т е л ь с т в о. Из симметрии скалярного произведения имеем

$$\left(\vec{x},\left[\vec{y},\vec{z}\right]\right)=\left(\left[\vec{y},\vec{z}\right],\vec{x}\right).$$

Чтобы доказать справедливость равенства (2.2.12), достаточно показать, что

$$\left(\left[\vec{x},\vec{y}\right],\vec{z}\right)=\left(\left[\vec{y},\vec{z}\right],\vec{x}\right).$$

Но это очевидно, так как тройкам одного наименования $\left(\vec{x},\vec{y},\vec{z}\right)$ и $\left(\vec{y},\vec{z},\vec{x}\right)$ соответствует один параллелепипед. ••

Справедливо следующее, легко проверяемое утверждение.

Утверждение 2.2.1. *Пусть* $\vec{a}\in R^3$ *— произвольный фиксированный вектор, тогда если для двух векторов* $\vec{x}$ *и* $\vec{y}$ *выполнено условие*

$$\left(\vec{x},\vec{a}\right)=\left(\vec{y},\vec{a}\right),$$

то эти векторы равны: $\vec{x}=\vec{y}$.

Д о к а з а т е л ь с т в о. Добавляя к обеим частям указанного в условии равенства число $-\left(\vec{y},\vec{a}\right)$ и используя аддитивность функции скалярного произведения, получаем

$$\left(\vec{x}-\vec{y},\vec{a}\right)=0,$$

откуда с учётом аксиомы невырожденности скалярного произведения следует, что $\vec{x}-\vec{y}=\vec{0}$. ••

Теорема 2.2.3. *Для любых векторов* $\vec{x},\vec{y},\vec{z}\in R^3$ *и для любого действительного числа* α *имеют место следующие свойства векторного произведения*:

$$1)\ \left[\vec{x},\vec{y}\right]=-\left[\vec{y},\vec{x}\right];\qquad 2)\ \left[\alpha\cdot\vec{x},\vec{y}\right]=\alpha\cdot\left[\vec{x},\vec{y}\right];$$

$$3)\ \left[\vec{x}+\vec{y},\vec{z}\right]=\left[\vec{x},\vec{z}\right]+\left[\vec{y},\vec{z}\right];\quad 4)\ \left[\vec{x},\vec{x}\right]=\vec{0}.\qquad (2.2.13)$$

Доказательство этой теоремы проводится с использованием утверждения 2.2.1 и предлагается в качестве упражнения.

Из теоремы 2.2.3 вытекают такие следствия.

Следствие 1 из теоремы 2.2.3. *Ориентированный объём* $V^{\pm}\left(\vec{x}, \vec{y}, \vec{z}\right)$ *параллелепипеда, построенного на векторах* $\vec{x}, \vec{y}, \vec{z}$, *линеен по каждому из своих аргументов, то есть для любых векторов, например,* $\vec{x}', \vec{x}'' \in R^3$, *и для любых* $\alpha, \beta \in R$ *имеет место равенство*

$$V^{\pm}\left(\alpha \cdot \vec{x}' + \beta \cdot \vec{x}'', \vec{y}, \vec{z}\right) = \alpha \cdot V^{\pm}\left(\vec{x}', \vec{y}, \vec{z}\right) + \beta \cdot V^{\pm}\left(\vec{x}'', \vec{y}, \vec{z}\right).$$

Следствие 2 из теоремы 2.2.3. *Если система векторов* $\vec{x}, \vec{y}, \vec{z}$ *линейно зависима, то*

$$V^{\pm}\left(\vec{x}, \vec{y}, \vec{z}\right) = 0.$$

Следствие 3 из теоремы 2.2.3. *Ориентированный объём параллелепипеда, построенного на векторах любой ортонормированной системы, равен единице.*

Формулы для вычисления векторного и смешанного произведений. Используя свойства векторного и смешанного произведений, нетрудно получить формулы, выражающие их через координаты перемножаемых векторов.

Теорема 2.2.4. *Если в некоторой декартовой системе координат векторы* $\vec{x}$ *и* $\vec{y}$ *заданы, соответственно, разложениями*

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3, \quad \vec{y} = y^1 \vec{e}_1 + y^2 \vec{e}_2 + y^3 \vec{e}_3,$$

то справедлива формула

$$\left[\vec{x}, \vec{y}\right] = \left(x^2 y^3 - x^3 y^2\right)\vec{e}_1 + \left(x^3 y^1 - x^1 y^3\right)\vec{e}_2 + \left(x^1 y^2 - x^2 y^1\right)\vec{e}_3. \quad (2.2.14)$$

Д о к а з а т е л ь с т в о. Поскольку имеют место разложения

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3, \quad \vec{y} = y^1 \vec{e}_1 + y^2 \vec{e}_2 + y^3 \vec{e}_3,$$

то, проводя несложные выкладки с учётом свойств векторного произведения и выражения (2.2.10) для канонического (ортонормированного) базиса векторов $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$, получаем доказательство теоремы. ••

Следствие из теоремы 2.2.4. *Пусть в некоторой декартовой системе координат векторы* $\vec{x}$, $\vec{y}$ *и* $\vec{z}$ *определены своими разложениями*

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3,$$

$$\vec{y} = y^1 \vec{e}_1 + y^2 \vec{e}_2 + y^3 \vec{e}_3,$$

$$\vec{z} = z^1 \vec{e}_1 + z^2 \vec{e}_2 + z^3 \vec{e}_3.$$

Тогда в той же системе координат справедлива следующая формула для смешанного произведения:

$$\left(\left[\vec{x}, \vec{y}\right], \vec{z}\right) = x^1 y^2 z^3 + x^2 y^3 z^1 + x^3 y^1 z^2 - x^1 y^3 z^2 - x^2 y^1 z^3 - x^3 y^2 z^1 . \quad (2.2.15)$$

Д о к а з а т е л ь с т в о. Рекомендуется доказать формулу (2.2.15) самостоятельно, проводя выкладки непосредственно. ••

2.3. Прямая линия и плоскость в евклидовых пространствах *R*² и *R*³, уравнения и свойства

До сих пор понятия прямой линии и плоскости не были определены строго и рассматривались на интуитивном уровне. В этом параграфе понятия прямой линии и плоскости вводятся строго, но в простейшем случае евклидовых пространств R^2 и R^3. Дальнейшее обобщение теории плоскостей на случай векторных пространств высшей размерности будет изложено ниже.

Уравнения прямой линии на плоскости R^2. Фиксируя в пространстве R^2 канонический базис $\left\{O, \vec{e}_1, \vec{e}_2\right\}$, введём декартову систему координат. Каждые две точки $M_0\left(x_0^1; x_0^2\right)$ и $M_1\left(x_1^1; x_1^1\right)$ с известными (фиксированными) координатами определяют ***фиксированный*** вектор $\vec{a}$. Если мы умножим данный вектор на некоторый ***параметр*** t, который может принимать произвольные значения из множества действительных чисел, то есть $t \in (-\infty, \infty)$, то получим вектор

$$\overrightarrow{M_0 M} = t \cdot \vec{a},$$

начальная точка M_0 которого фиксирована, а конечная точка $M\left(x^1; x^2\right)$ является ***текущей***, то есть имеет меняющиеся в зависимости от значения параметра t координаты. Сам вектор $\overrightarrow{M_0 M_1}$ называется ***текущим вектором***. ***Бесконечное множество текущих точек*** $M\left(x^1; x^2\right)$ ***назовём прямой линией*** L***, проходящей через опорную точку*** $M_0\left(x_0^1; x_0^2\right)$ ***и имеющей направляющий вектор*** $\vec{a}$.

Получим уравнения для координат текущей точки $M\left(x^1; x^2\right)$, радиус-вектор которой $\vec{x}$ называется ***ведущим*** вектором точек прямой линии. Для этого обратимся к рис. 2.11. Из рисунка видно, что выполняются соотношения:

$$\vec{x} = \vec{x}_0 + \overrightarrow{M_0 M}, \ \overrightarrow{M_0 M} = t \cdot \vec{a},$$

$$\vec{x}_0 \equiv \overrightarrow{OM_0} = x_0^1 \cdot \vec{e}_1 + x_0^2 \cdot \vec{e}_2,$$

$$\vec{x} \equiv \overrightarrow{OM} = x^1 \cdot \vec{e}_1 + x^2 \cdot \vec{e}_2,$$

$$\vec{a} = a^1 \cdot \vec{e}_1 + a^2 \cdot \vec{e}_2.$$

Подставляя в первое соотношение все остальные, получаем следующее равенство:

$$x^1 \cdot \vec{e}_1 + x^2 \cdot \vec{e}_2 = x_0^1 \cdot \vec{e}_1 + x_0^2 \cdot \vec{e}_2 + t \cdot \left(a^1 \cdot \vec{e}_1 + a^2 \cdot \vec{e}_2\right).$$

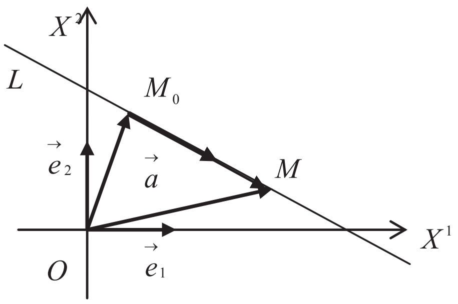


Рис. 2.11

Приравнивая координаты векторов в правой и левой частях этого равенства, получаем ***параметрические уравнения прямой линии на плоскости*** R^2:

$$\begin{cases} x^1 = x_0^1 + a^1 \cdot t, \\ x^2 = x_0^2 + a^2 \cdot t. \end{cases} \qquad (2.3.1)$$

Из параметрических уравнений (2.3.1) легко получить так называемое ***каноническое уравнение прямой линии*** на плоскости: прибавляя к обеим частям первого и второго уравнений соответственно $-x_0^1$ и $-x_0^2$ и деля обе части получившихся равенств на a^1 и a^2 соответственно, имеем:

$$\begin{cases} \dfrac{x^1 - x_0^1}{a^1} = t, \\ \dfrac{x^2 - x_0^2}{a^2} = t. \end{cases}$$

Приравнивая левые части, получаем следующее уравнение:

$$\frac{x^1 - x_0^1}{a^1} = \frac{x^2 - x_0^2}{a^2}, \qquad (2.3.2)$$

которое называется ***каноническим уравнением*** прямой линии на плоскости R^2.

Исключим из уравнений (2.3.1) параметр t, для чего умножим первое уравнение на a^2, а второе на a^1 и вычтем получившиеся равенства, получим

$$a^2 \cdot x^1 - a^1 \cdot x^2 = a^2 \cdot x_0^1 - a^1 \cdot x_0^2,$$

или

$$a^2 \cdot \left(x^1 - x_0^1\right) - a^1 \cdot \left(x^2 - x_0^2\right) = 0. \qquad (2.3.3)$$

Равенство (2.3.3) можно истолковать как условие ортогональности вектора

$$\vec{N} = a^2 \vec{e}_1 + \left(-a^1\right)\vec{e}_2$$

направляющему вектору прямой линии $\vec{a}$ и текущему вектору

$$\overrightarrow{M_0 M} = \left(x^1 - x_0^1\right)\vec{e}_1 + \left(x^2 - x_0^2\right)\vec{e}_2.$$

Вектор $\vec{N}$ называется ***нормальным вектором прямой линии L***. Таким образом, уравнение прямой линии (2.3.2) приводит к условию ортогональности нормального вектора прямой линии $\vec{N}$ и текущего вектора $\overrightarrow{M_0M}$, то есть

$$\left(\vec{N}, \overrightarrow{M_0M}\right) = 0.$$

Обозначая $a^2 \equiv A, -a^1 \equiv B$ и $-a^2 \cdot x_0^1 + a^1 \cdot x_0^2 \equiv C$, получаем ***неявное уравнение прямой линии***, записанное в стандартной форме

$$A \cdot x^1 + B \cdot x^2 + C = 0. \tag{2.3.4}$$

Из уравнения (2.3.4) при условии $C \neq 0$ следует уравнение ***прямой линии в отрезках***

$$\frac{x^1}{a} + \frac{x^2}{b} = 1, \tag{2.3.5}$$

где $a = -\dfrac{C}{A}$, $b = -\dfrac{C}{B}$. Условие $C \neq 0$, очевидно, означает, что прямая линия не проходит через начало координат. Геометрический смысл коэффициентов a и b легко выясняется из рис. 2.11 — эти коэффициенты равны, соответственно, ***величинам отрезков, отсекаемых прямой линией L на осях координат***.

Так как рассматриваются свободные векторы, для подсчёта координат которых положение начала системы координат не имеет значения, для вывода параметрических уравнений прямой линии в качестве направляющего вектора можно выбрать любой вектор $\vec{a}$, коллинеарный любому вектору, лежащему на прямой L. Аналогично, в качестве нормального вектора мы можем выбрать любой вектор, ортогональный любому вектору, лежащему на прямой L.

Пример 2.3.1. Получить уравнение прямой линии L, проходящей через точку $M_0(3; 2)$ и имеющей нормальный вектор $\vec{N} = \vec{e}_1 - 4\vec{e}_2$.

Р е ш е н и е. Уравнение прямой линии получается из условия ортогональности вектора $\vec{N}$ и текущего вектора прямой линии

$$\overrightarrow{M_0M} = \left(x^1 - 3\right)\vec{e}_1 + \left(x^2 - 2\right)\vec{e}_2,$$

то есть

$$\left(\vec{N}, \overrightarrow{M_0M}\right) = 0.$$

Используя выражение для скалярного произведения

$$\left(\vec{N}, \overrightarrow{M_0 M}\right) = x^1 - 4 \cdot x^2 + 5,$$

получаем неявное уравнение прямой линии $x^1 - 4 \cdot x^2 + 5 = 0$. ⊗

Уравнение прямой линии в трёхмерном пространстве R^3. Рассмотрим теперь случай пространства R^3 (рис. 2.12). Очевидно, что справедливы следующие соотношения:

$$\vec{x} = \vec{x} + \overrightarrow{M_0 M}, \ \overrightarrow{M_0 M} = t \cdot \vec{a},$$

$$\vec{x}_0 \equiv \overrightarrow{OM}_0 = x_0^1 \cdot \vec{e}_1 + x_0^2 \cdot \vec{e}_2 + x_0^3 \cdot \vec{e}_3,$$

$$\vec{x} \equiv \overrightarrow{OM} = x^1 \cdot \vec{e}_1 + x^2 \cdot \vec{e}_2 + x^3 \cdot \vec{e}_3,$$

$$\vec{a} = a^1 \cdot \vec{e}_1 + a^2 \cdot \vec{e}_2 + a^3 \cdot \vec{e}_3,$$

которые отличаются от соответствующих соотношений для плоскости R^2 только наличием третьей координаты.

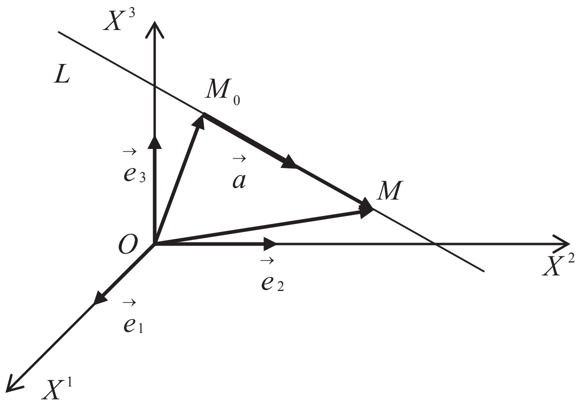


Рис. 2.12

Здесь снова точка $M\left(x^1, x^2, x^3\right)$ — ***текущая*** точка прямой линии, а её радиус-вектор $\vec{x}$ — ***ведущий*** вектор точек прямой линии.

Простые преобразования дают

$$x^1 \cdot \vec{e}_1 + x^2 \cdot \vec{e}_2 + x^3 \cdot \vec{e}_3 =$$

$$= x_0^1 \cdot \vec{e}_1 + x_0^2 \cdot \vec{e}_2 + x_0^3 \cdot \vec{e}_3 + t \cdot \left(a^1 \cdot \vec{e}_1 + a^2 \cdot \vec{e}_2 + a^3 \cdot \vec{e}_3\right),$$

откуда следуют ***параметрические уравнения прямой линии в евклидовом пространстве*** R^3:

$$\begin{cases} x^1 = x_0^1 + a^1 \cdot t, \\ x^2 = x_0^2 + a^2 \cdot t, \\ x^3 = x_0^3 + a^3 \cdot t. \end{cases} \tag{2.3.6}$$

Из уравнений (2.3.6) нетрудно получить канонические уравнения прямой линии в пространстве R^3, которые имеют вид:

$$\frac{x^1 - x_0^1}{a^1} = \frac{x^2 - x_0^2}{a^2} = \frac{x^3 - x_0^3}{a^3}. \tag{2.3.7}$$

Ниже показано, что канонические уравнения прямой линии (2.3.7) выражают тот факт, что прямая линия является множеством точек пересечения двух плоскостей в пространстве R^3.

Уравнения плоскости в трёхмерном пространстве R^3. На рис. 2.13 изображена фиксированная (***опорная***) точка M_0, от которой откладывается ***текущий*** вектор $\overrightarrow{M_0M}$, который является элементом линейной оболочки неколлинеарных векторов $\vec{a}_1$ и $\vec{a}_2$, закреплённых в точке M_0.

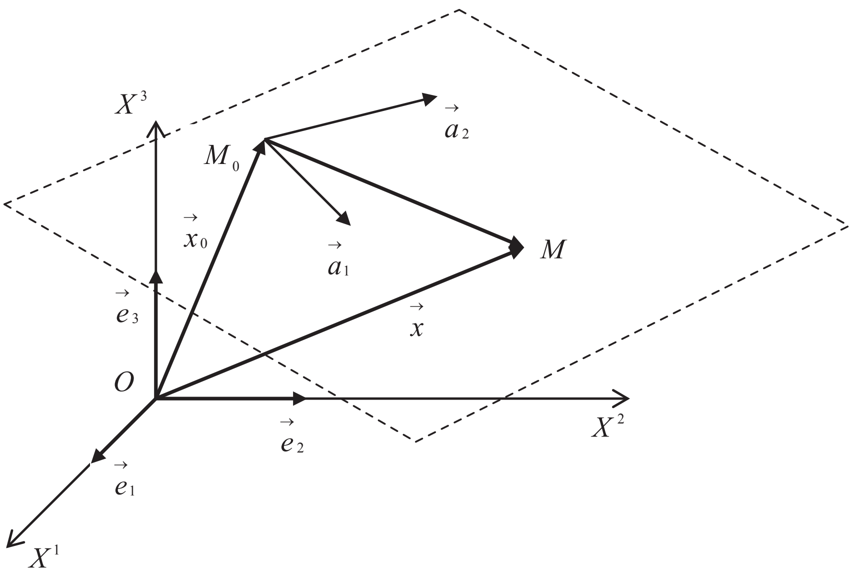


Рис. 2.13

Непосредственно из рисунка видно, что справедливы следующие соотношения и разложения векторов по базису декартовой системы координат:

$$\vec{x} \equiv \vec{x}_0 + \overrightarrow{M_0M};$$

$$\overrightarrow{M_0M} = t_1\vec{a}_1 + t_2\vec{a}_2;$$

$$\vec{x}_0 \equiv \overrightarrow{OM_0} = x_0^1\vec{e}_1 + x_0^2\vec{e}_2 + x_0^3\vec{e}_3;$$

$$\vec{x} \equiv \overrightarrow{OM} = x^1\vec{e}_1 + x^2\vec{e}_2 + x^3\vec{e}_3;$$

$$\vec{a}_1 = a_1^1\vec{e}_1 + a_1^2\vec{e}_2 + a_1^3\vec{e}_3;$$

$$\vec{a}_2 = a_2^1\vec{e}_1 + a_2^2\vec{e}_2 + a_2^3\vec{e}_3.$$

Вектор $\vec{x}$ называется ***ведущим*** вектором точек плоскости.

Истолковывая первое из приведённых соотношений аналогично тому, как это было сделано при выводе уравнения прямой линии, и записывая его с учётом второго соотношения в виде

$$\vec{x} = \vec{x}_0 + t_1\vec{a}_1 + t_2\vec{a}_2, \tag{2.3.8}$$

приходим к понятию плоскости.

Назовём ***плоскостью*** H^2 ***в пространстве*** R^3 ***множество текущих точек*** $M\left(x^1; x^2; x^3\right)$***, являющихся конечными точками вектора***

$$\overrightarrow{M_0M} = t_1 \cdot \vec{a}_1 + t_2 \cdot \vec{a}_2,$$

представленного разложением по паре неколлинеарных векторов $\left\{\vec{a}_1, \vec{a}_2\right\}$***, приложенных к точке*** $M_0\left(x_0^1; x_0^2; x_0^3\right)$***, при всевозможных значениях параметров*** t_1 и t_2. Векторы $\left\{\vec{a}_1, \vec{a}_2\right\}$ называются ***направляющими векторами плоскости*** H^2. Само уравнение (2.3.8) называется ***векторным параметрическим уравнением плоскости***.

Подстановка разложений всех векторов в векторное уравнение плоскости даёт:

$$x^1 \cdot \vec{e}_1 + x^2 \cdot \vec{e}_2 + x^3 \cdot \vec{e}_3 = x_0^1 \cdot \vec{e}_1 + x_0^2 \cdot \vec{e}_2 + x_0^3 \cdot \vec{e}_3 +$$
$$+ t_1 \cdot \left(a_1^1 \cdot \vec{e}_1 + a_1^2 \cdot \vec{e}_2 + a_1^3 \cdot \vec{e}_3\right) + t_2 \cdot \left(a_2^1 \cdot \vec{e}_1 + a_2^2 \cdot \vec{e}_2 + a_2^3 \cdot \vec{e}_3\right).$$

Приравнивая в последнем равенстве координаты при соответствующих базисных векторах, получаем параметрические уравнения плоскости H^2 в евклидовом пространстве R^3:

$$\begin{cases} x^1 = x_0^1 + a_1^1 \cdot t_1 + a_2^1 \cdot t_2, \\ x^2 = x_0^2 + a_1^2 \cdot t_1 + a_2^2 \cdot t_2, \\ x^3 = x_0^3 + a_1^3 \cdot t_1 + a_2^3 \cdot t_2. \end{cases} \tag{2.3.9}$$

Рассмотрим следующее преобразование уравнений (2.3.9). Исключим параметры t_1 и t_2 из системы уравнений (2.3.9), для чего рассмотрим первые два уравнения как систему относительно параметров

$$\begin{cases} a_1^1 \cdot t_1 + a_2^1 \cdot t_2 = x^1 - x_0^1, \\ a_1^2 \cdot t_1 + a_2^2 \cdot t_2 = x^2 - x_0^2. \end{cases}$$

Предполагая, что $a_1^1 \cdot a_2^2 - a_2^1 \cdot a_1^2 \neq 0$ (условие, следующее из того, что направляющие векторы линейно независимы), решим систему, умножая обе части каждого из уравнений на соответствующие величины и складывая уравнения почленно:

$$t_1 = \frac{\left(x^1 - x_0^1\right) \cdot a_2^2 - \left(x^2 - x_0^2\right) \cdot a_2^1}{a_1^1 \cdot a_2^2 - a_2^1 \cdot a_1^2}, \ t_2 = \frac{\left(x^2 - x_0^2\right) \cdot a_1^1 - \left(x^1 - x_0^1\right) \cdot a_1^2}{a_1^1 \cdot a_2^2 - a_2^1 \cdot a_1^2}.$$

Подстановка в третье уравнение системы (2.3.9) после простых преобразований приводит нас к следующему равенству:

$$\begin{aligned} &\left(a_2^3 \cdot a_1^2 - a_1^3 \cdot a_2^2\right) \cdot x^1 + \left(a_1^3 \cdot a_2^1 - a_2^3 \cdot a_1^1\right) \cdot x^2 + \left(a_1^1 \cdot a_2^2 - a_1^2 \cdot a_2^1\right) \cdot x^3 = \\ &= \left(a_2^3 \cdot a_1^2 - a_1^3 \cdot a_2^2\right) \cdot x_0^1 + \left(a_1^3 \cdot a_2^1 - a_2^3 \cdot a_1^1\right) \cdot x_0^2 + \left(a_1^1 \cdot a_2^2 - a_1^2 \cdot a_2^1\right) \cdot x_0^3, \end{aligned} \tag{2.3.10}$$

которое может быть переписано в виде

$$\begin{aligned} &\left(a_2^3 \cdot a_1^2 - a_1^3 \cdot a_2^2\right) \cdot \left(x^1 - x_0^1\right) + \left(a_1^3 \cdot a_2^1 - a_2^3 \cdot a_1^1\right) \cdot \left(x^2 - x_0^2\right) + \\ &+ \left(a_1^1 \cdot a_2^2 - a_1^2 \cdot a_2^1\right) \cdot \left(x^3 - x_0^3\right) = 0. \end{aligned}$$

Из этого равенства видно, что скалярное произведение, в качестве аргументов которого взяты векторное произведение

$$\left[\vec{a}_1, \vec{a}_2\right] =$$

$$= \left(a_2^3 \cdot a_1^2 - a_1^3 \cdot a_2^2\right) \cdot \vec{e}_1 + \left(a_1^3 \cdot a_2^1 - a_2^3 \cdot a_1^1\right) \cdot \vec{e}_2 + \left(a_1^1 \cdot a_2^2 - a_1^2 \cdot a_2^1\right) \cdot \vec{e}_3$$

и текущий вектор $\overrightarrow{M_0M}$, равно нулю. Поэтому вектор

$$\vec{N} = \left[\vec{a}_1, \vec{a}_2\right]$$

естественно назвать ***нормальным вектором плоскости*** $H^2 \subset R^3$. Соотношение (2.3.10) можно переписать в виде

$$\left(\left[\vec{a}_1, \vec{a}_2\right], \overrightarrow{OM}\right) = \left(\left[\vec{a}_1, \vec{a}_2\right], \overrightarrow{OM_0}\right), \qquad (2.3.11)$$

из которого видно, что из параметрических уравнений плоскости (2.3.9) следует ***постоянство ориентированного объёма параллелепипеда, построенного на приведённых к общему началу направляющих векторах и радиус-вектора текущей точки плоскости***. Обозначая в выражении (2.3.10) координаты вектора $\left[\vec{a}_1, \vec{a}_2\right]$ соответственно A, B и C, а правую часть – D, получаем неявное уравнение плоскости

$$A \cdot x^1 + B \cdot x^2 + C \cdot x^3 + D = 0. \qquad (2.3.12)$$

Из уравнения (2.3.12) при условии $D \neq 0$ получаем ***уравнение плоскости в отрезках***

$$\frac{x^1}{a} + \frac{x^2}{b} + \frac{x^3}{c} = 1, \qquad (2.3.13)$$

где $a = -\dfrac{D}{A}, b = -\dfrac{D}{B}, c = -\dfrac{D}{C}$. Нетрудно видеть, что эти коэффициенты равны величинам отрезков, отсекаемых плоскостью на осях координат. Очевидно, что уравнение плоскости в отрезках может быть получено только в том случае, если плоскость не проходит через начало координат.

В отношении нормального вектора плоскости $\vec{N} = \left[\vec{a}_1, \vec{a}_2\right]$ справедливо то же замечание, что и в отношении нормального вектора прямой линии на плоскости R^2: в качестве нормального вектора можно выбрать любой вектор, ортогональный произвольному вектору, лежащему в плоскости.

Взаимное расположение прямой линии и плоскости в пространстве R^3. Среди различных задач, решаемых средствами аналитической геометрии в пространстве R^3, есть несколько типичных, а именно: задача определения двугранного угла между плоскостями; задача получения уравнения прямой линии, заданной как пересечение двух плоскостей; задача нахождения координат точки пересечения прямой линии L и плоскости H^2; задача нахождения координат проекции точки $M_0\left(x_0^1; x_0^2; x_0^3\right)$ на плоскость H^2 и расстояния от данной точки до плоскости. Рассмотрим решение этих задач последовательно.

Угол между плоскостями $H^2_{(1)}$ ***и*** $H^2_{(2)}$**.** Пусть требуется найти угол между плоскостями, проходящими через три точки, не лежащие на одной прямой,

$$A_1\left(x_1^1; x_1^2, x_1^3\right), A_2\left(x_2^1; x_2^2, x_2^3\right), A_3\left(x_3^1; x_3^2, x_3^3\right),$$

и

$$B_1\left(y_1^1; y_1^2, y_1^3\right), B_2\left(y_2^1; y_2^2, y_2^3\right), B_1\left(y_3^1; y_3^2, y_3^3\right)$$

соответственно. Для решения этой задачи поступаем следующим образом:

1) за направляющие векторы плоскостей принимаем векторы

$$\vec{a}_1 \equiv \overrightarrow{A_1A_2} = a_1^1 \cdot \vec{e}_1 + a_1^2 \cdot \vec{e}_2 + a_1^3 \cdot \vec{e}_3,$$

$$\vec{a}_2 \equiv \overrightarrow{A_1A_3} = a_2^1 \cdot \vec{e}_1 + a_2^2 \cdot \vec{e}_2 + a_2^3 \cdot \vec{e}_3;$$

$$\vec{b}_1 \equiv \overrightarrow{B_1B_2} = b_1^1 \cdot \vec{e}_1 + b_1^2 \cdot \vec{e}_2 + b_1^3 \cdot \vec{e}_3,$$

$$\vec{b}_2 \equiv \overrightarrow{B_1B_3} = b_2^1 \cdot \vec{e}_1 + b_2^2 \cdot \vec{e}_2 + b_2^3 \cdot \vec{e}_3$$

соответственно;

2) находим нормальные векторы плоскостей как векторные произведения направляющих векторов:

$$\vec{N}_{(1)} = \left[\overrightarrow{A_1A_2}, \overrightarrow{A_1A_3}\right], \vec{N}_{(2)} = \left[\overrightarrow{B_1B_2}, \overrightarrow{B_1B_3}\right];$$

3) находим косинус угла между плоскостями как косинус угла между нормальными векторами плоскостей:

$$\cos\varphi = \frac{1}{\left\|\vec{N}_1\right\| \cdot \left\|\vec{N}_2\right\|}\left(\vec{N}_1, \vec{N}_2\right).$$

Если плоскости заданы своими параметрическими уравнениями

$$\begin{cases} x^1 = x_0^1 + a_1^1 \cdot t_1 + a_2^1 \cdot t_2, \\ x^2 = x_0^2 + a_1^2 \cdot t_1 + a_2^2 \cdot t_2, \\ x^3 = x_0^3 + a_1^3 \cdot t_1 + a_2^3 \cdot t_2, \end{cases} \quad \begin{cases} x^1 = x_0^1 + b_1^1 \cdot s_1 + b_2^1 \cdot s_2, \\ x^2 = x_0^2 + b_1^2 \cdot s_1 + b_2^2 \cdot s_2, \\ x^3 = x_0^3 + b_1^3 \cdot s_1 + b_2^3 \cdot s_2, \end{cases}$$

где $M_0\left(x_0^1, x_0^2, x_0^3\right)$ и $N_0\left(y_0^1, y_0^2, y_0^3\right)$ — точки, через которые проходят плоскости, а $a_j^i\left(i=1,2,3;\, j=1,2\right)$ и $b_j^i\left(i=1,2,3;\, j=1,2\right)$ — координаты направляющих векторов плоскостей, то находим направляющие векторы из параметрических уравнений. Далее находим угол через его косинус.

Если плоскости заданы неявными уравнениями

$$A_{(1)} \cdot x^1 + B_{(1)} \cdot x^2 + C_{(1)} \cdot x^3 + D_{(1)} = 0;$$

$$A_{(2)} \cdot x^1 + B_{(2)} \cdot x^2 + C_{(2)} \cdot x^3 + D_{(2)} = 0,$$

то, учитывая, что координаты нормальных векторов равны соответственно

$$A_{(1)}, B_{(1)}, C_{(1)} \text{ и } A_{(2)}, B_{(2)}, C_{(2)},$$

угол между плоскостями находим, исходя из формулы для косинуса:

$$\cos\varphi = \frac{\left(A_{(1)} \cdot A_{(2)} + B_{(1)} \cdot B_{(2)} + C_{(1)} \cdot C_{(2)}\right)}{\sqrt{\left(A_1\right)^2 + \left(B_1\right)^2 + \left(C_1\right)^2} \cdot \sqrt{\left(A_2\right)^2 + \left(B_2\right)^2 + \left(C_3\right)^2}}.$$

Уравнение прямой линии, заданной пересечением двух плоскостей. Пусть требуется найти параметрические или канонические уравнения прямой линии, заданной как множество точек пересечения двух плоскостей с неявными уравнениями:

$$A_{(1)} \cdot x^1 + B_{(1)} \cdot x^2 + C_{(1)} \cdot x^3 + D_{(1)} = 0,$$

$$A_{(2)} \cdot x^1 + B_{(2)} \cdot x^2 + C_{(2)} \cdot x^3 + D_{(2)} = 0.$$

1. Проверяем, что плоскости пересекаются, то есть что нормальные векторы неколлинеарны.

2. Ищем направляющий вектор прямой линии. Так как плоскости пересекаются по прямой, то её направляющий вектор ортогонален нормальным векторам обеих плоскостей. Поэтому

$$\vec{a} = \left[\vec{N}_1, \vec{N}_2\right] = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ A_{(1)} & B_{(1)} & C_{(1)} \\ A_{(2)} & B_{(2)} & C_{(2)} \end{vmatrix}.$$

3. Ищем какую-либо точку, которую можно выбрать в качестве опорной точки. Так как направляющий вектор прямой линии непараллелен хотя бы одной из координатных плоскостей, то в качестве опорной точки прямой линии выбираем точку пересечения этой прямой и какой-либо координатной плоскости.

4. Подставляя координаты точки прямой и направляющего вектора в параметрические или канонические уравнения прямой, получаем решение задачи.

Координаты точки пересечения прямой линии L и плоскости H^2**.** Пусть требуется найти координаты точки пересечения прямой линии L и плоскости H^2, заданных соответственно уравнениями

$$\frac{x^1 - x_0^1}{a^1} = \frac{x^2 - x_0^2}{a^2} = \frac{x^3 - x_0^3}{a^3}, \; A \cdot x^1 + B \cdot x^2 + C \cdot x^3 + D = 0.$$

Для решения этой задачи поступаем так.

1. Проверяем, что прямая линия не параллельна плоскости, то есть что направляющий вектор прямой и нормальный вектор плоскости

$$\vec{a} = a^1 \cdot \vec{e}_1 + a^2 \cdot \vec{e}_2 + a^3 \cdot \vec{e}_3, \; \vec{N} = A \cdot \vec{e}_1 + B \cdot \vec{e}_2 + C \cdot \vec{e}_3$$

не ортогональны, то есть проверяем выполнение неравенства:

$$\left(\vec{a}, \vec{N}\right) = A \cdot a^1 + B \cdot a^2 + C \cdot a^3 \neq 0.$$

2. Параметрические уравнения прямой линии имеют вид

$$\begin{cases} x^1 = x_0^1 + a^1 \cdot t, \\ x^2 = x_0^2 + a^2 \cdot t, \\ x^3 = x_0^3 + a^3 \cdot t. \end{cases}$$

Подставляя эти выражения для координат точек прямой x^1, x^2, x^3 в неявное уравнение плоскости, получаем уравнение относительно параметра t, решая которое, находим значение параметра $t = t_0$, соответствующее точке пересечения прямой линии и плоскости.

3. Подставляя найденное значение параметра $t = t_0$ в параметрические уравнения прямой линии, находим координаты точки пересечения.

Координаты проекции точки $M_0\left(x_0^1; x_0^2; x_0^3\right)$ ***на плоскость*** H^2 ***и расстояние от данной точки до плоскости.*** Пусть требуется найти координаты проекции точки $M_0\left(x_0^1; x_0^2; x_0^3\right)$ на плоскость с уравнением

$$A \cdot x^1 + B \cdot x^2 + C \cdot x^3 + D = 0$$

и расстояние от точки M_0 до плоскости. Для решения этой задачи поступаем так.

1. Так как проекция точки на плоскость является основанием $M'\left(x^{1'}, x^{2'}, x^{3'}\right)$ перпендикуляра, опущенного из данной точки на плоскость, составляем уравнения прямой линии, проходящей через точку $M_0\left(x_0^1; x_0^2; x_0^3\right)$ перпендикулярно данной плоскости. Для этого в качестве направляющего вектора прямой выбираем нормальный вектор плоскости, то есть полагаем

$$\vec{a} = \vec{N} = A \cdot \vec{e}_1 + B \cdot \vec{e}_2 + C \cdot \vec{e}_3.$$

Тогда параметрические уравнения прямой линии принимают вид:

$$\begin{cases} x^1 = x_0^1 + A \cdot t, \\ x^2 = x_0^2 + B \cdot t, \\ x^3 = x_0^3 + C \cdot t. \end{cases}$$

2. Подставляя x^1, x^2, x^3 в неявное уравнение плоскости и решая получающееся уравнение относительно параметра t, находим значение параметра $t = t_0$, при котором прямая пересекается с плоскостью.

3. Подставляя найденное значение параметра $t = t_0$ в параметрические уравнения прямой линии, находим координаты проекции $x^{1'}, x^{2'}, x^{3'}$.

4. Находим расстояние от точки $M_0\left(x_0^1; x_0^2; x_0^3\right)$ до данной плоскости, то есть норму вектора $\overrightarrow{M_0 M'}$:

$$\left\| \overrightarrow{M_0 M'} \right\| = \sqrt{\left(x^{1'} - x_0^1\right)^2 + \left(x^{2'} - x_0^2\right)^2 + \left(x^{3'} - x_0^3\right)^2}.$$

2.4. Абстрактные векторные пространства и системы линейных алгебраических уравнений

Абстрактные векторные пространства *n* измерений. Во многих разделах математики, физики и других естественных наук требуется дальнейшее обобщение понятия трёхмерного векторного пространства. Рассмотрим это обобщение, для чего сначала дадим развёрнутое определение ***абстрактного векторного пространства***.

Определение 2.4.1. *Множество X абстрактных элементов* (***векторов***) $\vec{x}, \vec{y}, \vec{z}, \ldots$ *называется* ***векторным пространством*** *над полем P, если для его элементов выполнены перечисленные ниже аксиомы.*

1^0. *Для любых векторов* $\vec{x}, \vec{y} \in X$ *однозначно определена* ***операция сложения***, *результатом которой является вектор, обозначаемый* $\vec{x} + \vec{y} \in X$ *и называемый* ***суммой*** *векторов* $\vec{x}$ *и* $\vec{y}$, *причём операция сложения обладает следующими свойствами:*

1) $\left(\forall \vec{x}, \vec{y}, \vec{z} \in X\right)\left(\vec{x} + \vec{y}\right) + \vec{z} = \vec{x} + \left(\vec{y} + \vec{z}\right)$ — ***ассоциативность;***

2) $\left(\forall \vec{x}, \vec{y} \in X\right)\vec{x} + \vec{y} = \vec{x} + \vec{y}$ — ***коммутативность***.

2^0. *Существует однозначно определённый элемент* $\vec{0} \in X$, *такой, что* $\left(\forall \vec{x} \in X\right)$

$$\vec{x} + \vec{0} = \vec{0} + \vec{x} = \vec{x},$$

который называется ***нуль-вектор***.

3^0. *Существует однозначно определённый вектор* $-\vec{x} \in X$, *такой, что* $\left(\forall \vec{x} \in X\right)$

$$\vec{x} + \left(-\vec{x}\right) = \left(-\vec{x}\right) + \vec{x} = \vec{0},$$

который называется ***обратным к вектору*** $\vec{x} \in X$.

4^0. *Для любого числа* $\alpha \in P$ *и для любого вектора* $\vec{x} \in X$ *определена операция* ***умножения вектора на число***, *результатом которой является вектор* $\alpha \cdot \vec{x} \in X$, *называемый* ***произведением вектора*** $\vec{x}$ ***на число*** α, *причём* *операция умножения векторов на числа обладает следующими свойствами:* $(\forall \alpha, \beta \in P)$ и $\left(\forall \vec{x}, \vec{y} \in X\right)$

1) $1 \cdot \vec{x} = \vec{x};$

2) $\alpha \cdot \left(\vec{x} + \vec{y}\right) = \alpha \cdot \vec{x} + \alpha \cdot \vec{y};$

3) $(\alpha + \beta) \cdot \vec{x} = \alpha \cdot \vec{x} + \beta \cdot \vec{x};$

4) $(\alpha \cdot \beta) \cdot \vec{x} = \alpha \cdot \left(\beta \cdot \vec{x}\right)$. •

Из приведённых аксиом следует существование операции, обратной операции сложения векторов — операции ***вычитания векторов***.

Вычесть вектор $\vec{y}$ *из вектора* $\vec{x}$ *— это значит найти вектор* $\vec{z}$*, такой, что*

$$\vec{z} + \vec{y} = \vec{x}.$$

Вектор $\vec{z}$ называется ***разностью векторов*** $\vec{x}$ и $\vec{y}$**.** Обозначается разность векторов так:

$$\vec{z} = \vec{x} - \vec{y}.$$

Докажем однозначность выполнения операции вычитания векторов. Действительно, пусть разность $\vec{z}$ существует, то есть выполнено равенство

$$\vec{z} + \vec{y} = \vec{x}.$$

К обеим частям этого равенства добавим $-\vec{y}$. Получим:

$$\vec{z} + \vec{y} + \left(-\vec{y}\right) = \vec{x} + \left(-\vec{y}\right).$$

Используя ассоциативность сложения, получаем

$$\vec{z} = \vec{x} + \left(-\vec{y}\right).$$

Таким образом, если разность $\vec{z}$ существует, то она имеет полученный вид.

Покажем, что $\vec{x}+\left(-\vec{y}\right)$ действительно разность. Для этого сложим $\vec{x}+\left(-\vec{y}\right)$ с вектором $\vec{y}$, в результате получим вектор $\vec{x}$. То есть $\vec{x}-\vec{y}=\vec{x}+(-\vec{y})$.

Имеют место два важных утверждения.

Утверждение 2.4.1. *При любом* $\vec{x}\in X$ *выполняется равенство* $0\cdot\vec{x}=\vec{0}$.

Доказательство. Для произвольного числа α составим выражение:

$$\alpha\cdot\vec{x}+0\cdot\vec{x}=(\alpha+0)\cdot\vec{x}=\alpha\cdot\vec{x}.$$

Здесь использована аксиома 4^0. Итак,

$$\alpha\cdot\vec{x}+0\cdot\vec{x}=\alpha\cdot\vec{x}\Rightarrow 0\cdot\vec{x}=\alpha\cdot\vec{x}-\alpha\cdot\vec{x}=\vec{0},$$

что и доказывает утверждение. ••

Утверждение 2.4.2. *Для любого* $\alpha\in P$ *выполняется равенство* $\alpha\cdot\vec{0}=\vec{0}$.

Доказательство. Доказательство этого утверждения проводится аналогично доказательству предыдущего утверждения, для чего рассматривается следующая цепочка преобразований:

$$\alpha\cdot\vec{x}+\alpha\cdot\vec{0}=\alpha\cdot\left(\vec{x}+\vec{0}\right)=\alpha\cdot\vec{x}\Rightarrow\alpha\cdot\vec{0}=\alpha\cdot\vec{x}-\alpha\cdot\vec{x}=\vec{0}.\ \bullet\bullet$$

В абстрактном векторном пространстве, аналогично случаю пространства R^3, вводятся понятия ***линейно зависимых*** и ***линейно независимых систем векторов*** и ***размерности пространства***.

Система векторов $\left\{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n\right\}\subset X$ называется ***линейно независимой***, если

$$\left(\alpha_1\cdot\vec{x}_1+\alpha_2\cdot x_2+\ldots+\alpha_n\cdot\vec{x}_n=\vec{0}\right)\Leftrightarrow\left(\alpha_1=\alpha_2=\ldots=\alpha_n=0\right),$$

и ***линейно зависимой*** в противном случае.

Если векторное пространство содержит линейно независимую систему из n векторов, а любая система, содержащая $n+1$ векторов, уже линейно зависима (система является максимальной по числу векторов линейно независимой системой в пространстве X), то говорят, что пространство X является ***n*-мерным векторным пространством**, и обозначают его X^n. Максимальная по числу векторов линейно независимая система $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset X^n$ (содержащая n векторов) называется ***базисом пространства*** X^n. Сформулируем соответствующую аксиому — ***аксиому размерности***.

5^0. *В пространстве X^n существует линейно независимая система, содержащая только конечное число n векторов, а система из любых $n+1$ векторов линейно зависима.*

Теорема 2.4.1. *Любой вектор $\vec{x} \in X^n$ единственным образом можно представить в виде разложения*

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{e}_i = x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^n \vec{e}_n \tag{2.4.1}$$

по векторам как-либо выбранной линейно независимой системе n векторов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, образующей базис пространства.

Доказательство. Система векторов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ по условию теоремы является базисом, то есть максимальной по числу векторов линейно независимой системой в пространстве X^n. Поэтому, добавляя к этой системе вектор $\vec{x}$, получим линейно зависимую систему из $n+1$ векторов

$$\left\{\vec{x}, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}.$$

Но тогда по определению линейной зависимости можно подобрать такие неравные одновременно нулю коэффициенты $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_n$, что линейная комбинация векторов системы будет иметь своим значением нуль-вектор:

$$\alpha \cdot \vec{x} + \alpha_1 \cdot \vec{e}_1 + \alpha_2 \cdot \vec{e}_2 + \ldots + \alpha_n \cdot \vec{e}_n = \vec{0}. \tag{2.4.2}$$

В этом равенстве коэффициент $\alpha \neq 0$, так как в противном случае получили бы, что

$$\alpha_1 \cdot \vec{e}_1 + \alpha_2 \cdot \vec{e}_2 + \ldots + \alpha_n \cdot \vec{e}_n = \vec{0},$$

причём в последнем равенстве не все коэффициенты равны нулю одновременно, а это невозможно, так как система векторов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$

по условию линейно независима. Прибавляя к обеим частям равенства (2.4.2) вектор

$$-\alpha_1 \cdot \vec{e}_1 - \alpha_2 \cdot \vec{e}_2 - \ldots - \alpha_n \cdot \vec{e}_n,$$

умножая обе части получившегося равенства на величину $1/\alpha$, получаем

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^n \vec{e}_n,$$

где введены обозначения $x^k = -\dfrac{\alpha_k}{\alpha},\ k = 1, 2, \ldots, n.$

Единственность полученного разложения доказывается методом от противного. Действительно, пусть имеют место два разложения

$$\vec{x} = x_1^1 \vec{e}_1 + x_1^2 \vec{e}_2 + \ldots + x_1^n \vec{e}_n,\ \vec{x} = x_2^1 \vec{e}_1 + x_2^2 \vec{e}_2 + \ldots + x_2^n \vec{e}_n.$$

Так как левые части равны, то получаем

$$\left(x_1^1 - x_2^1\right)\vec{e}_1 + \left(x_1^2 - x_2^2\right)\vec{e}_2 + \ldots + \left(x_1^n - x_2^n\right)\vec{e}_n = \vec{0}.$$

Из последнего равенства в силу линейной независимости системы базисных векторов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ получаем:

$$x_1^1 = x_2^1,\ x_1^2 = x_2^2,\ \ldots,\ x_1^n = x_2^n,$$

что и доказывает единственность разложения вектора по базису. ••

Коэффициенты разложения в формуле (2.4.1) называются ***координатами*** вектора относительно базиса $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$.

Легко видеть, что если система векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\} \subset X^n$ линейно зависима, то любой вектор системы выражается в виде линейной комбинации остальных векторов системы. Действительно, пусть система линейно зависима. Тогда можно подобрать неравные нулю одновременно коэффициенты так, чтобы выполнялось равенство

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \ldots + \alpha_m \vec{x}_m = \vec{0}.$$

Пусть, например, $\alpha_1 \neq 0$, тогда, умножая обе части векторного равенства на величину, обратную α_1, получим

$$\vec{x}_1 = -\frac{\alpha_2}{\alpha_1}\cdot\vec{x}_2 - \ldots - \frac{\alpha_m}{\alpha_1}\cdot\vec{x}_m = \beta_2\cdot\vec{x}_2 + \ldots + \beta_m\cdot\vec{x}_m,$$

где $\beta_k = -\dfrac{\alpha_k}{\alpha}$, $k = 1, 2, \ldots, m-1$.

Справедливы также следующие свойства систем векторов:

1) если система векторов содержит линейно зависимую подсистему, то и вся система линейно зависима, в частности, любая система, содержащая нуль-вектор, линейно зависима;

2) если система векторов линейно независима, то и любая её подсистема также линейно независима.

Из аксиом $1^0 - 4^0$ непосредственно следует, что для нахождения суммы векторов $\vec{x}, \vec{y} \in X^n$ нужно сложить их координаты, а чтобы умножить вектор $\vec{x}$ на число, нужно все его координаты умножить на это число. Таким образом, в абстрактном векторном пространстве все операции с векторами производятся ***по тем же правилам***, что и в пространстве R^3.

Пусть, например, в пространстве X^n относительно фиксированного базиса $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset X$ своими разложениями заданы два вектора

$$\vec{x} = x^1\vec{e}_1 + x^2\vec{e}_2 + \ldots + x^n\vec{e}_n,\ \vec{y} = y^1\vec{e}_1 + y^2\vec{e}_2 + \ldots + y^n\vec{e}_n.$$

Утверждается, что эти векторы равны в том и только в том случае, если

$$(\forall j = 1, 2, \ldots, n)\ x^j = y^j.$$

Действительно, естественно считать векторы $\vec{x}$ и $\vec{y}$ равными, если для них выполняется условие

$$\vec{x} + \left(-\vec{y}\right) = \vec{0},$$

Но тогда из аксиомы $3^0, 4^0$ и существования операции вычитания следует, что

$$\vec{x} + \left(-\vec{y}\right) = \vec{x} - \vec{y} = \left(x^1 - y^1\right)\vec{e}_1 + \left(x^2 - y^2\right)\vec{e}_2 + \ldots + \left(x^n - y^n\right)\vec{e}_n = \vec{0}.$$

В силу линейной независимости системы векторов базиса (или единственности нуль-вектора) получаем:

$$(\forall j = 1, 2, \ldots, n)\, x^j = y^j.$$

Обратно, если $(\forall j = 1, 2, \ldots, n)\, x^j = y^j$, то $\vec{x} - \vec{y} = \vec{0}$ и $\vec{x} = \vec{y}$.

Имеются, однако, и существенные различия в понятиях евклидова пространства R^3 и абстрактного векторного пространства X^n.

Действительно, в абстрактном векторном пространстве вводится понятие ***линейного многообразия*** (***линейной оболочки***) — множества всевозможных линейных комбинаций

$$L\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} = \left\{\vec{x} \in X : \vec{x} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_m \vec{a}_m\right\}$$

элементов линейно независимой системы векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \in X^n$

с произвольно изменяющимися коэффициентами $\alpha_1, \alpha_2, \ldots, \alpha_m$. Ввиду отсутствия понятия точки линейное многообразие ***не может быть теперь названо плоскостью***.

В силу отсутствия в абстрактном векторном пространстве X^n понятия точки его элементы лишены своей ***генетической связи с геометрией*** и не являются, вообще говоря, многомерным аналогом ***геометрических векторов*** трёхмерного пространства в модели R^3. Однако относительно понятия линейной зависимости и независимости системы векторов и свойств операций, выполняемых над абстрактными векторами, абстрактное векторное пространство X^n ничем не отличается от евклидова пространства R^3. По этой причине почти все теоретические построения в абстрактных векторных пространствах проводятся в большой аналогии со случаем пространства R^3. Отметим, что элементами X^n могут быть самые различные математические объекты, например, функции. По этой причине стрелка в обозначении абстрактных векторов часто не используется (элементы абстрактного векторного пространства обозначаются жирным шрифтом). Мы, однако, сохраним за векторами обозначение со стрелкой независимо от природы векторов векторного пространства.

Системы линейных алгебраических уравнений (СЛАУ). Выше выяснили, что любой вектор линейно зависимой системы $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\} \subset X^n$

может быть представлен в виде линейной комбинации остальных векторов системы. Сформулируем следующую задачу:

пусть в векторном пространстве X^n, определённом над полем действительных чисел R^1, задана система $m+1$ векторов $\left\{\vec{b}, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$, требуется подобрать числа $x^1, x^2, \ldots, x^m \in R^1$ так, чтобы выполнялось равенство

$$x^1 \vec{a}_1 + x^2 \vec{a}_2 + \ldots + x^m \vec{a}_m = \vec{b}, \tag{2.4.3}$$

то есть нужно выяснить вопрос, является ли система векторов

$$\left\{\vec{b}, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$$

линейно зависимой и, если система линейно зависимая, найти разложение вектора $\vec{b}$ по векторам $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$.

Фиксируем в пространстве X^n некоторый базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$. Векторы $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ и $\vec{b}$ разложим по векторам базиса:

$$\begin{cases} \vec{a}_1 = a_1^1 \vec{e}_1 + a_1^2 \vec{e}_2 + \ldots + a_1^n \vec{e}_n, \\ \vec{a}_2 = a_2^1 \vec{e}_1 + a_2^2 \vec{e}_2 + \ldots + a_2^n \vec{e}_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \vec{a}_m = a_m^1 \vec{e}_1 + a_m^2 \vec{e}_2 + \ldots + a_m^n \vec{e}_n, \\ \vec{b} = b^1 \vec{e}_1 + b^2 \vec{e}_2 + \ldots + b^n \vec{e}_n. \end{cases} \tag{2.4.4}$$

Подставляя разложения (2.4.4) в разложение (2.4.3), проводя несложные преобразования и приравнивая коэффициенты при соответствующих векторах базиса, получаем:

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_m^1 x^m = b^1, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_m^2 x^m = b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n x^1 + a_2^n x^2 + \ldots + a_m^n x^m = b^n. \end{cases} \tag{2.4.5}$$

Итак, имеем следующий важнейший факт: *векторное уравнение* (2.4.3) *в фиксированном базисе* $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ *векторного простран-*

*ства X^n свелось к **системе линейных алгебраических уравнений*** (2.4.5) *относительно системы **неизвестных величин*** $x^1, x^2, \ldots, x^m \in R$.

Ниже для термина «***система линейных алгебраических уравнений***» мы будем применять общепринятую аббревиатуру — ***СЛАУ***. Введём следующее определение.

Определение 2.4.2. *Упорядоченная система* $\{c^1, c^2, \ldots, c^m\}$ *значений неизвестных* $\{x^1, x^2, \ldots, x^m\}$, *удовлетворяющая каждому из уравнений СЛАУ (2.4.5) (то есть обращающая каждое уравнение в тождество), называется **решением** СЛАУ. Если СЛАУ имеет хотя бы одно решение, она называется **совместной,** и **несовместной** — в противном случае. Если СЛАУ совместна и имеет единственное решение, то она называется **определённой,** и **неопределённой** — в противном случае (то есть когда СЛАУ имеет не одно решение).*•

В процессе исследования СЛАУ вида (2.4.5) решаются следующие ***основные задачи:***

1) *выяснить, является СЛАУ совместной или несовместной;*

2) *если СЛАУ совместна, то выяснить, является ли она определённой;*

3) *если СЛАУ совместна и определённа, то найти её единственное решение;*

4) *если СЛАУ совместна и неопределённа, то описать совокупность всех её решений.*

Эквивалентные СЛАУ. Метод Гаусса. Среди всевозможных СЛАУ особое значение имеют так называемые ***эквивалентные СЛАУ***.

Определение 2.4.3. *Две СЛАУ относительно одной и той же системы неизвестных называются **эквивалентными**, если каждое решение одной СЛАУ является решением другой или они обе несовместны.*•

Из определения 2.4.3 следует ***общий метод решения СЛАУ, основанный на её преобразовании к такой эквивалентной системе, решение которой находится достаточно просто***. Этот метод называется ***методом Гаусса*** и состоит не более чем из $n-1$ этапов. Коэффициенты при неизвестных и правые части на каждом этапе метода обозначим левым верхним индексом. При использовании таких обозначений исходная система (2.4.5) принимает вид:

$$\begin{cases} {}^{(0)}a_1^1 x^1 + {}^{(0)}a_2^1 x^2 + \ldots + {}^{(0)}a_m^1 x^m = {}^{(0)}b^1, \\ {}^{(0)}a_1^2 x^1 + {}^{(0)}a_2^2 x^2 + \ldots + {}^{(0)}a_m^2 x^m = {}^{(0)}b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ {}^{(0)}a_1^n x^1 + {}^{(0)}a_2^n x^2 + \ldots + {}^{(0)}a_m^n x^m = {}^{(0)}b^n. \end{cases} \tag{2.4.6}$$

На первом этапе рассмотрим первое уравнение. Если все коэффициенты при неизвестных и правая часть равны нулю, то этому уравнению удовлетворяет *любая совокупность значений неизвестных*. Поэтому, если первое уравнение исключить из рассмотрения, мы получим *эквивалентную СЛАУ*. Если же все коэффициенты при неизвестных равны нулю, а правые части неравны нулю, то такому уравнению *не удовлетворяет никакая совокупность значений неизвестных*. В последнем случае *СЛАУ несовместна* и её исследование закончено.

Пусть среди коэффициентов при неизвестных в первом уравнении есть хотя бы один отличный от нуля, например, ${}^{(0)}a_1^1 \neq 0$ (если отличен от нуля другой коэффициент, его можно перестановкой неизвестных поместить на первое место). Первый отличный от нуля коэффициент, в данном случае ${}^{(0)}a_1^1$, ***назовём ведущим коэффициентом***. Выразим неизвестное x^1 из первого уравнения через остальные неизвестные и правую часть, полученное выражение подставим во все оставшиеся уравнения СЛАУ, причём ***первое уравнение оставим без изменения***. Приведя подобные, получим:

$$\begin{cases} {}^{(0)}a_1^1 x^1 + {}^{(0)}a_2^1 x^2 + \ldots + {}^{(0)}a_m^1 x^m = {}^{(0)}b^1, \\ \qquad\qquad {}^{(1)}a_2^2 x^2 + \ldots + {}^{(1)}a_m^2 x^m = {}^{(1)}b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \qquad\qquad {}^{(1)}a_2^n x^2 + \ldots + {}^{(1)}a_m^n x^m = {}^{(1)}b^n. \end{cases} \tag{2.4.7}$$

СЛАУ (2.4.7) и (2.4.6) ***эквивалентны***. На этом первый этап исследования СЛАУ (2.4.6) закончен.

На втором этапе исключаем неизвестное x^2 из всех уравнений, кроме первых двух.

На третьем этапе исключаем неизвестное x^3 из всех уравнений, кроме первых трёх. И так далее ...

Если в процессе преобразований мы не встретим уравнений, в которых все коэффициенты при неизвестных равны нулю, то после $n-1$ этапов придём к СЛАУ следующего вида:

$$\begin{cases} {}^{(0)}a_1^1 x^1 + {}^{(0)}a_2^1 x^2 + \ldots + {}^{(0)}a_n^1 x^n + {}^{(0)}a_{n+1}^1 x^{n+1} + \ldots + {}^{(0)}a_m^1 x^m = {}^{(0)}b^1, \\ \qquad\qquad {}^{(1)}a_2^2 x^2 + \ldots + {}^{(1)}a_n^2 x^n + {}^{(1)}a_{n+1}^2 x^{n+1} + \ldots + {}^{(1)}a_m^2 x^m = {}^{(1)}b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \qquad\qquad\qquad {}^{(n-1)}a_n^n x^n + {}^{(n-1)}a_{n+1}^n x^{n+1} + \ldots + {}^{(n-1)}a_m^n x^m = {}^{(n-1)}b^n, \end{cases}$$

которая эквивалентна исходной СЛАУ (2.4.6).

Может случиться, что в процессе преобразований нам встретятся уравнения, удовлетворяющиеся тождественно. Тогда СЛАУ $(n-1)$-го этапа будет состоять из меньшего числа уравнений.

Определение 2.4.4. *Неизвестные* $x^{n+1}, \ldots, x^m$ *называются* ***свободными неизвестными***.•

Свободным неизвестным мы можем придать любые действительные значения. При этом, какие бы значения этим неизвестным мы ни приписали, можно последовательно определить и все остальные неизвестные из СЛАУ $(n-1)$-го этапа путём обратной подстановки, начиная с x^n. При этом все коэффициенты ${}^{(0)}a_1^1, {}^{(1)}a_2^2, \ldots, {}^{(n-1)}a_n^n$ являются ведущими и поэтому отличны от нуля. На каждом этапе исследования СЛАУ коэффициенты при неизвестных достаточно сложным образом выражаются через коэффициенты эквивалентной СЛАУ предыдущего этапа. Не выписывая общие формулы для коэффициентов, покажем процедуру реализации метода Гаусса на примере, показывающем, как с его помощью выяснить вопрос о линейной зависимости или независимости системы векторов.

Пример 2.4.1. В трёхмерном евклидовом пространстве R^3 в некотором базисе $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ заданы векторы

$$\vec{x}_1 = \vec{e}_1 + \vec{e}_2,\ \vec{x}_2 = \vec{e}_1 - \vec{e}_2 + \vec{e}_3,\ \vec{x}_3 = -3 \cdot \vec{e}_1 + 5 \cdot \vec{e}_2 - 6 \cdot \vec{e}_3.$$

Показать, что эти векторы также образуют базис.

Р е ш е н и е. 1. Составим линейную комбинацию заданных векторов и проверим, при каких условиях на её коэффициенты выполняется следующее основное векторное равенство:

$$\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2 + \alpha_3 \cdot \vec{x}_3 = \vec{0}.$$

2. Вычисляем линейную комбинацию в левой части предыдущего равенства, используя правила умножения вектора на число и сложения векторов:

$$(\alpha_1 + \alpha_2 - 3 \cdot \alpha_3) \cdot \vec{e}_1 + (\alpha_1 - \alpha_2 + 5 \cdot \alpha_3) \cdot \vec{e}_2 + (\alpha_2 - 6 \cdot \alpha_3) \cdot \vec{e}_3 = \vec{0}.$$

3. Приравнивая координаты векторов в обеих частях последнего равенства

$$\begin{cases} \alpha_1 + \alpha_2 - 3\cdot\alpha_3 = 0, \\ \alpha_1 - \alpha_2 + 5\cdot\alpha_3 = 0, \\ \qquad \alpha_2 - 6\cdot\alpha_3 = 0, \end{cases}$$

получили СЛАУ относительно коэффициентов линейной комбинации исходных векторов.

4. Ищем решение СЛАУ методом Гаусса. Для этого выражаем неизвестное α_1 из первого уравнения системы и подставляем в остальные уравнения, а затем выражаем второе неизвестное из второго уравнения и подставляем в третье:

$$\begin{cases} \alpha_1 + \quad \alpha_2 - 3\cdot\alpha_3 = 0, \\ \qquad -2\cdot\alpha_2 + 8\cdot\alpha_3 = 0, \\ \qquad \alpha_2 - 6\cdot\alpha_3 = 0; \end{cases} \Rightarrow \begin{cases} \alpha_1 + \quad \alpha_2 - 3\cdot\alpha_3 = 0, \\ \qquad -2\cdot\alpha_2 + 8\cdot\alpha_3 = 0, \\ \qquad\qquad 2\cdot\alpha_3 = 0. \end{cases}$$

Из третьего уравнения последней СЛАУ находим $\alpha_3 = 0$. Из второго и первого уравнений имеем $\alpha_2 = 0$, $\alpha_1 = 0$. Итак, основное векторное равенство выполняется только при нулевых значениях коэффициентов $\alpha_1, \alpha_2, \alpha_3$. Последнее и означает, что векторы $\vec{x}_1, \vec{x}_2, \vec{x}_3$ образуют линейно независимую систему, которая, следовательно, является одним из возможных базисов трёхмерного пространства. ⊗

2.5. Аффинное и евклидово пространства

Аффинное и евклидово пространства *n* измерений. В абстрактных векторных пространствах можно ввести понятие скалярного произведения векторов и построить евклидово пространство. Чтобы на базе абстрактного векторного пространства можно было построить евклидово пространство, необходимо, однако, дополнить определение аксиомами, постулирующими наличие в пространстве X^n аналога понятию точки в пространствах R^3. Сформулируем эти аксиомы.

6^0. *В пространстве X^n существует, по меньшей мере, одна точка.*

7^0. *Каждой упорядоченной паре точек $\{A, B\}$ поставлен в соответствие один и только один вектор* (*обозначение вектора* $\vec{AB}, \vec{x}, \ldots$).

8^0. *Для каждой точки А и каждого вектора* $\vec{x}$ *существует одна и только одна точка В, такая, что* $\vec{AB} = \vec{x}$.

9^0. *Если* $\vec{AB} = \vec{CD}$, то $\vec{AC} = \vec{BD}$ (*аксиома параллелограмма*).

Множество точек и векторов, удовлетворяющих аксиомам 1^0 — 9^0, называется ***n-мерным аффинным пространством*** и обозначается A^n.

В отличие от пространства R^3 в аффинном пространстве A^n нет понятия расстояния и по этой причине оно лишено метрических свойств. Определим на аффинном пространстве A^n скалярную функцию двух векторных аргументов таким образом, чтобы через неё можно было бы определить расстояние.

Определение 2.5.1. *Пусть A^n — некоторое n-мерное аффинное пространство, заданное над полем действительных чисел R^1.* ***Скалярной функцией двух векторных переменных*** $\vec{x}, \vec{y}$ *называется отображение* $\varphi : A^n \times A^n \to R^1$, *ставящее в соответствие каждой упорядоченной паре векторов* $\left\{\vec{x}, \vec{y}\right\} \in A^n \times A^n$ *вполне определённое действительное число* $\varphi = \varphi\left(\vec{x}, \vec{y}\right)$.•

Основываясь на опытных фактах, требуют, чтобы введённая функция $\varphi = \varphi\left(\vec{x}, \vec{y}\right)$ удовлетворяла следующей ***системе аксиом***:

$$\left(\forall \vec{x}, \vec{y}, \vec{z} \in A^n\right) \wedge \left(\forall \alpha \in R^1\right)$$

1) $\varphi\left(\vec{x}, \vec{y}\right) = \varphi\left(\vec{y}, \vec{x}\right)$;

2) $\varphi\left(\alpha \cdot \vec{x}, \vec{y}\right) = \alpha \cdot \varphi\left(\vec{x}, \vec{y}\right), \quad \varphi\left(\vec{x}, \beta \cdot \vec{y}\right) = \beta \cdot \varphi\left(\vec{x}, \vec{y}\right)$;

3) $\varphi\left(\vec{x} + \vec{y}, \vec{z}\right) = \varphi\left(\vec{x}, \vec{z}\right) + \varphi\left(\vec{y}, \vec{z}\right), \quad \varphi\left(\vec{x}, \vec{y} + \vec{z}\right) = \varphi\left(\vec{x}, \vec{y}\right) + \varphi\left(\vec{x}, \vec{z}\right)$;

4) $\left(\forall \vec{x} \neq \vec{0}\right)\left(\exists \vec{y} \neq \vec{0}\right) : \varphi\left(\vec{x}, \vec{y}\right) \neq 0$.

Первая аксиома означает, что функция $\varphi\left(\vec{x}, \vec{y}\right)$ является ***симметрической***. Вторая и третья аксиомы означают, что эта функция является ***линейной функцией по первому и второму аргументу***. Четвёртая аксиома утверждает, что функция φ является ***невырожденной***.

Можно показать, что четвёртая аксиома может быть записана в виде

$$\varphi\left(\vec{x}, \vec{x}\right) > 0 \Leftrightarrow \vec{x} \neq \vec{0} \text{ и } \varphi\left(\vec{x}, \vec{x}\right) = 0 \Leftrightarrow \vec{x} = \vec{0}.$$

При условии выполнения четвёртой аксиомы евклидово пространство называется ещё ***собственно евклидовым пространством***. Следует отметить, что существуют евклидовы пространства, в которых четвёртая аксиома не выполняется. Такие пространства по своим геометрическим свойствам существенно отличаются от собственно евклидова пространства.

На основе аксиомы 1 нетрудно сделать вывод о том, что если аксиомы линейности 2 и 3 выполняются в отношении первого аргумента функции φ, то они выполняются и в отношении второго аргумента, то есть часть этих аксиом относительно второго аргумента является следствием первой части относительно первого аргумента функции φ. Покажем, например, что

$$\varphi\left(\vec{x}, \vec{y}+\vec{z}\right) = \varphi\left(\vec{x}, \vec{y}\right) + \varphi\left(\vec{x}, \vec{z}\right).$$

Действительно, справедливость этой аксиомы вытекает из следующей цепочки преобразований:

$$\varphi\left(\vec{x}, \vec{y}+\vec{z}\right) = \{\text{аксиома } 1\} = \varphi\left(\vec{y}+\vec{z}, \vec{x}\right) =$$

$$= \varphi\left(\vec{y}, \vec{x}\right) + \varphi\left(\vec{z}, \vec{x}\right) = \{\text{аксиома } 1\} = \varphi\left(\vec{x}, \vec{y}\right) + \varphi\left(\vec{x}, \vec{z}\right).$$

Таким образом, функция φ из определения 2.5.1, согласно аксиомам, является линейной по каждому аргументу. Поэтому функция φ называется ***билинейной функцией***.

Определение 2.5.2. ***Евклидовым пространством n измерений*** *E^n называется n-мерное аффинное пространство A^n, на векторах которого задана фиксированная, билинейная, скалярная функция* $\varphi = \varphi\left(\vec{x}, \vec{y}\right)$ *двух векторных аргументов* $\vec{x}$ *и* $\vec{y}$, *удовлетворяющая аксиомам* ***симметрии*** *и* ***невырожденности***. •

За фигурирующей в определении 2.5.2 функцией φ исторически закрепилось название «***скалярное произведение***» векторов $\vec{x}$ и $\vec{y}$. Следует, однако, чётко понимать, что это именно функция, не имеющая никакого отношения к какой-либо операции умножения.

Аксиомы 1–4 называются ***аксиомами скалярного произведения***.

Для скалярного произведения используются различные обозначения:

$$\varphi\left(\vec{x}, \vec{y}\right) \equiv \left(\vec{x}, \vec{y}\right) \equiv \vec{x}\,\vec{y} \equiv \vec{x}\cdot\vec{y}. \tag{2.5.1}$$

Понятие базиса в евклидовом пространстве E^n вводится аналогично случаю абстрактного векторного пространства. Дополняя базис фиксированной точкой — началом O, получим аналог репера в трёхмерном евклидовом пространстве — ***евклидов репер*** $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$.

Пусть в евклидовом пространстве E^n фиксирован некоторый репер

$$\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n.$$

Тогда для векторов $\vec{x}, \vec{y} \in E^n$ справедливы разложения:

$$\vec{x} = x^1\vec{e}_1 + x^2\vec{e}_2 + \ldots + x^n\vec{e}_n = \sum_{i=1}^{n} x^i\vec{e}_i; \tag{2.5.2}$$

$$\vec{y} = y^1\vec{e}_1 + y^2\vec{e}_2 + \ldots + y^n\vec{e}_n = \sum_{j=1}^{n} y^j\vec{e}_j. \tag{2.5.3}$$

Используя аксиомы скалярного произведения, свойства конечных сумм и разложения (2.5.2) и (2.5.3), легко выразить скалярное произведение векторов $\vec{x}$ и $\vec{y}$ через их координаты:

$$\left(\vec{x}, \vec{y}\right) = \left(\sum_{i=1}^{n} x^i\vec{e}_i, \sum_{j=1}^{n} y^j\vec{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j \left(\vec{e}_i, \vec{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x^i y^j. \tag{2.5.4}$$

В выражении (2.5.4) введены обозначения $g_{ij} \equiv \left(\vec{e}_i, \vec{e}_j\right)$ для скалярных произведений базисных векторов. Из полученного равенства видно, что скалярное произведение двух векторов имеет довольно сложный вид и выражается через скалярные произведения векторов базиса. Поэтому следует попытаться выбрать базис таким образом, чтобы скалярное произведение его векторов имело наиболее простой вид. Прежде чем выбрать требуемый базис, определим некоторые понятия.

Величины g_{ij} называются ***метрическими коэффициентами***, а само выражение (2.5.4) называется ***метрической билинейной*** (если $\vec{x} = \vec{y}$ — ***квадратичной***) ***формой***. Легко видеть, что в силу условия симметрии скалярного произведения (аксиома 1) метрические коэффициенты подчинены условию симметрии

$$g_{ij} = g_{ji},$$

где индексы принимают значения $i, j = 1, 2, \ldots, n$ независимо друг от друга.

Выражение типа (2.5.4)

$$\varphi\left(\vec{x}, \vec{y}\right) = \sum_{i-1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i y^j$$

в общем случае называется просто ***билинейной*** (если $\vec{x} = \vec{y}$ — ***квадратичной***) ***формой***, а его коэффициенты называются коэффициентами этой формы и обозначаются другой буквой, например φ_{ij}, так как обозначение g_{ij} зарезервировано именно за попарными произведениями векторов базиса коэффициентами метрической билинейной формы.

Отметим, что метрические коэффициенты являются ***координатами*** (или ***компонентами***) вполне определённого геометрического объекта, так называемого ***метрического тензора***. Значение этого объекта весьма велико — его строение определяет геометрию окружающего нас мира. Компоненты метрического тензора можно расположить в виде следующей таблицы:

$$G = \left(g_{ij}\right) = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}, \tag{2.5.5}$$

которая называется ***матрицей метрических коэффициентов*** (***матрицей Грама системы базисных векторов***).

Ортонормированный базис в собственно евклидовом пространстве E^n. В собственно евклидовом пространстве E^n вводятся аналогичные случаю пространства R^3 понятия, которые в общем случае имеют намного более сложный вид.

Определение 2.5.3. *Величина*

$$\vec{x}^{\,2} \equiv \left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i x^j$$

называется ***скалярным квадратом****, а величина*

$$\left\|\vec{x}\right\| \stackrel{def}{=} \left(\vec{x}, \vec{x}\right)^{\frac{1}{2}} \equiv \sqrt{\left(\vec{x}, \vec{x}\right)}$$

— ***евклидовой нормой*** (*дальше просто нормой*) *вектора* $\vec{x}$. *Вектор* $\vec{x} \in E^n$ *называется* ***нормированным****, если*

$$\left(\vec{x}, \vec{x}\right) = \left\|\vec{x}\right\|^2 = 1 \Rightarrow \left\|\vec{x}\right\| = 1. \bullet$$

Справедливы следующие утверждения.

Лемма 2.5.1. *Любой ненулевой вектор* $\vec{x} \in E^n$ $\left(\vec{x} \neq \vec{0}\right)$ *можно нормировать, умножив его на некоторое действительное число* λ.

Д о к а з а т е л ь с т в о. Потребуем, чтобы

$$\left(\lambda \cdot \vec{x}, \lambda \cdot \vec{x}\right) = \lambda^2 \cdot \left(\vec{x}, \vec{x}\right) = 1.$$

Откуда видно, что

$$\lambda = \left(\vec{x}, \vec{x}\right)^{-\frac{1}{2}} = \left\|\vec{x}\right\|^{-1}.$$

То, что при таком выборе множителя λ вектор $\lambda \cdot \vec{x}$ нормирован, проверяется прямой выкладкой:

$$\left(\lambda \cdot \vec{x}, \lambda \cdot \vec{x}\right) = \frac{1}{\left(\vec{x}, \vec{x}\right)}\left(\vec{x}, \vec{x}\right) \equiv \left\|\vec{x}\right\|^{-2}\left\|\vec{x}\right\|^{2} = 1. \bullet\bullet$$

Определение 2.5.4. *Два вектора* $\vec{x}$ *и* $\vec{y}$ *(любого векторного пространства) называются* ***коллинеарными****, если*

$$\left(\exists \lambda, \mu \in R^1\right) : \vee\ \vec{x} = \lambda \cdot \vec{y} \vee \vec{y} = \mu \cdot \vec{x}. \bullet$$

Так как для любого вектора $\vec{x}$ всегда $\vec{0} = 0 \cdot \vec{x}$, то два вектора заведомо коллинеарны, если один из них нулевой.

Теорема 2.5.1. *Для любых векторов* $\vec{x}, \vec{y} \in E^n$ *справедливо* ***неравенство Коши — Буняковского***

$$\left(\vec{x}, \vec{y}\right)^2 \le \left(\vec{x}, \vec{x}\right) \cdot \left(\vec{y}, \vec{y}\right). \tag{2.5.6}$$

Д о к а з а т е л ь с т в о. Так как для $\vec{y} = \vec{0}$ утверждение теоремы очевидно, предположим, что $\vec{y} \ne \vec{0}$. Рассмотрим вектор $\vec{x} - \lambda \cdot \vec{y}$, где $\lambda \in R^1$. Для него имеем:

$$\left(\vec{x} - \lambda \cdot \vec{y}, \vec{x} - \lambda \cdot \vec{y}\right) = \left(\vec{x}, \vec{x}\right) - 2 \cdot \lambda \cdot \left(\vec{x}, \vec{y}\right) + \lambda^2 \cdot \left(\vec{y}, \vec{y}\right).$$

В левой части стоит скалярный квадрат вектора $\vec{x} - \lambda \cdot \vec{y}$, поэтому в соответствии с аксиомой 4 скалярного произведения квадратный трёхчлен относительно λ в правой части неотрицателен при любых значениях λ. Следовательно, для дискриминанта этого трёхчлена имеем:

$$D = 4 \cdot \left(\vec{x}, \vec{y}\right)^2 - 4 \cdot \left(\vec{x}, \vec{x}\right) \cdot \left(\vec{y}, \vec{y}\right) \le 0.$$

Откуда и получаем неравенство (2.5.6). ••

Теорема 2.5.2. *Неравенство Коши — Буняковского обращается в равенство в том и только в том случае, если векторы* $\vec{x}$ *и* $\vec{y}$ *коллинеарны.*

Д о к а з а т е л ь с т в о. *Докажем необходимость.* Пусть векторы коллинеарны, то есть, например, справедливо равенство $\vec{x} = \lambda \cdot \vec{y}$. Нетрудно видеть, что в этом случае

$$\left(\vec{x}, \vec{y}\right)^2 = \left(\lambda \cdot \vec{y}, \vec{y}\right)^2 = \lambda^2 \cdot \left(\vec{y}, \vec{y}\right)^2.$$

Далее имеем

$$\left(\vec{x}, \vec{x}\right) \cdot \left(\vec{y}, \vec{y}\right) = \left(\lambda \cdot \vec{y}, \lambda \cdot \vec{y}\right) \cdot \left(\vec{y}, \vec{y}\right) = \lambda^2 \cdot \left(\vec{y}, \vec{y}\right)^2.$$

Сравнивая два последних равенства, имеем

$$\left(\vec{x}, \vec{y}\right)^2 = \left(\vec{x}, \vec{x}\right) \cdot \left(\vec{y}, \vec{y}\right).$$

Докажем достаточность. Пусть выполняется равенство

$$\left(\vec{x},\vec{y}\right)^2=\left(\vec{x},\vec{x}\right)\cdot\left(\vec{y},\vec{y}\right).$$

Если $\vec{y}=\vec{0}$, то очевидно, что векторы коллинеарны. Положим $\vec{y}\neq\vec{0}$. Выберем

$$\lambda=\left(\vec{y},\vec{y}\right)^{-1}\left(\vec{x},\ \vec{y}\right).$$

Тогда из предположения имеем:

$$\left(\vec{x},\vec{y}\right)^2=\left(\vec{x},\vec{x}\right)\cdot\left(\vec{y},\vec{y}\right)\Rightarrow\left(\vec{x},\vec{x}\right)\cdot\left(\vec{y},\vec{y}\right)-\left(\vec{x},\vec{y}\right)^2=0\Rightarrow$$

$$\left(\vec{x},\vec{x}\right)-\frac{\left(\vec{x},\vec{y}\right)}{\left(\vec{y},\vec{y}\right)}\left(\vec{x},\vec{y}\right)=0\Rightarrow\left(\vec{x},\vec{x}\right)-\lambda\cdot\left(\vec{x},\vec{y}\right)=0\Rightarrow$$

$$\left(\vec{x},\vec{x}-\lambda\cdot\vec{y}\right)=0\Rightarrow\vec{x}-\lambda\cdot\vec{y}=\vec{0}\Rightarrow\vec{x}=\lambda\cdot\vec{y}.\ \bullet\bullet$$

Определение 2.5.5. *Два вектора* $\vec{x},\vec{y}\in E^n$ *называются* ***ортогональными**, если их скалярное произведение равно нулю, то есть* $(\vec{x},\vec{y})=0.\bullet$

Определение 2.5.6. *Система векторов* $\left\{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_m\right\}$ *евклидова пространства* E^n *называется* ***ортогональной**, если она состоит из попарно ортогональных векторов. Если, кроме этого, векторы системы нормированы, то она называется* ***ортонормированной**.*•

Лемма 2.5.2. *Любая ортогональная (ортонормированная) система векторов евклидова пространства линейно независима.*

Д о к а з а т е л ь с т в о. Пусть система $\left\{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_m\right\}$ ортогональна и $\vec{x}_i\neq\vec{0}$ для любого $i=1,2,\ldots,m$. Тогда по определению ортогональности

$$\left(\vec{x}_i,\vec{x}_j\right)=0\text{ для }i\neq j;\qquad(2.5.7)$$

$$\left(\vec{x}_i, \vec{x}_j\right) \neq 0 \text{ для } i = j. \tag{2.5.8}$$

Составим равенство

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \ldots + \alpha_i \vec{x}_i + \ldots + \alpha_m \vec{x}_m = \vec{0}$$

и найдём скалярное произведение, выбирая в качестве его аргументов обе части данного равенства и вектор $\vec{x}_i$ для каждого номера $i = 1, 2, \ldots, m$. Так как, очевидно, что $\left(\vec{x}_i, \vec{0}\right) = 0$, то с учётом условий (2.5.7), (2.5.8) получаем

$$\alpha_1 \cdot \left(\vec{x}_1, \vec{x}_i\right) + \alpha_2 \cdot \left(\vec{x}_2, \vec{x}_i\right) + \ldots + \alpha_i \cdot \left(\vec{x}_i, \vec{x}_i\right) + \ldots + \alpha_m \cdot \left(\vec{x}_m, \vec{x}_i\right) = 0,$$

следовательно, $\alpha_i \cdot \left(\vec{x}_i, \vec{x}_i\right) = 0$, откуда вытекает, что $\left(\forall i = \overline{1, m}\right) \alpha_i = 0$.

Учитывая, что подобные действия произведены последовательно для каждого $i = 1, 2, \ldots, m$, получаем, что все коэффициенты α_i равны нулю одновременно. Отсюда в соответствии с определением линейной независимости приходим к доказательству теоремы. ••

Определение 2.5.7. *Базис, векторы которого образуют ортонормированную систему, называется* ***ортонормированным***. •

Нетрудно видеть, что ***скалярное произведение векторов***

$$\vec{x} = \sum_{k=1}^{n} x^k \vec{e}_k, \quad \vec{y} = \sum_{k=1}^{n} y^k \vec{e}_k$$

в случае, когда базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ ортонормированный, приобретает простой вид

$$\left(\vec{x}, \vec{y}\right) = \sum_{k=1}^{n} x^k y^k.$$

Действительно, ортонормированный базис характеризуется условиями

$$g_{ij} = \left(\vec{e}_i, \vec{e}_j\right) = \delta_{ij} = \begin{cases} 1, i = j; \\ 0, i \neq j. \end{cases}$$

С учётом этих условий получаем:

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j \left(\vec{e}_i, \vec{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i y^j = \sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} x^i y^j = \sum_{k=1}^{n} x^k y^k.$$

По этой причине вопрос существования ортонормированного базиса в евклидовом пространстве E^n имеет большое значение.

Отметим, что безотносительно к геометрическому смыслу величины

$$\delta_{ij} = \begin{cases} 1, i = j, \\ 0, i \neq j \end{cases}$$

образуют так называемый ***символ Кронекера***, который будет ещё использован многократно в процессе изучения линейной алгебры.

Теорема 2.5.3. *В любом конечномерном евклидовом пространстве E^n существует ортонормированный базис.*

Доказательство. Для доказательства теоремы применим ***процедуру ортогонализации Шмидта***.

Зафиксируем в пространстве E^n некоторый, в общем случае неортогональный и ненормированный базис

$$\left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset E^n$$

и построим на его основе сначала ортогональный базис, обозначив его, например, $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$. Положим

$$\vec{a}_1 = \vec{g}_1, \ \vec{a}_2 = \vec{g}_2 + \alpha_1 \vec{a}_1.$$

Коэффициент α_1 найдём, требуя ортогональности векторов $\vec{a}_1$ и $\vec{a}_2$:

$$\left(\vec{a}_1, \vec{a}_2\right) = 0 \Rightarrow \left(\vec{a}_1, \vec{g}_2 + \alpha_1 \cdot \vec{a}_1\right) = 0 \Rightarrow \left(\vec{a}_1, \vec{g}_2\right) + \alpha_1 \cdot \left(\vec{a}_1, \vec{a}_1\right) = 0.$$

Отсюда имеем

$$\alpha_1 = -\frac{\left(\vec{a}_1, \vec{g}_2\right)}{\left(\vec{a}_1, \vec{a}_1\right)} = -\frac{\left(\vec{a}_1, \vec{g}_2\right)}{\left\|\vec{a}_1\right\|^2}.$$

Итак, получаем:

$$\vec{a}_1 = \vec{g}_1;$$

$$\vec{a}_2 = \vec{g}_2 + \alpha_1 \vec{a}_1 = \vec{g}_2 - \frac{\left(\vec{a}_1, \vec{g}_2\right)}{\left(\vec{a}_1, \vec{a}_1\right)} \vec{a}_1 = \vec{g}_2 - \frac{\left(\vec{a}_1, \vec{g}_2\right)}{\left\|\vec{a}_1\right\|^2} \vec{a}_1. \qquad (2.5.9)$$

Положим далее

$$\vec{a}_3 = \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2.$$

Замечая, что $\left(\vec{a}_1, \vec{a}_2\right) = 0$, потребуем выполнения условий

$$\left(\vec{a}_1, \vec{a}_3\right) = 0, \left(\vec{a}_2, \vec{a}_3\right) = 0.$$

Получаем для коэффициента α_1:

$$\left(\vec{a}_1, \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2\right) = 0 \Rightarrow \left(\vec{a}_1, \vec{g}_3\right) + \alpha_1 \left(\vec{a}_1, \vec{a}_1\right) = 0.$$

Теперь видно, что

$$\alpha_1 = -\frac{\left(\vec{a}_1, \vec{g}_3\right)}{\left(\vec{a}_1, \vec{a}_1\right)} = -\frac{\left(\vec{a}_1, \vec{g}_3\right)}{\left\|\vec{a}_1\right\|^2}.$$

Аналогично для коэффициента α_2 имеем:

$$\left(\vec{a}_2, \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2\right) = 0 \Rightarrow \left(\vec{a}_2, \vec{g}_3\right) + \alpha_2 \left(\vec{a}_2, \vec{a}_2\right) = 0.$$

Учитывая, что $\left(\vec{a}_1, \vec{a}_2\right) = 0$, получаем

$$\alpha_2 = -\frac{\left(\vec{a}_2, \vec{g}_3\right)}{\left(\vec{a}_2, \vec{a}_2\right)} = -\frac{\left(\vec{a}_2, \vec{g}_3\right)}{\left\|\vec{a}_2\right\|^2}.$$

Итак, аналогично выражению (2.5.9), имеем:

$$\vec{a}_3 = \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 = \vec{g}_3 - \frac{\left(\vec{a}_1, \vec{g}_3\right)}{\left\|\vec{a}_1\right\|^2}\vec{a}_1 - \frac{\left(\vec{a}_2, \vec{g}_3\right)}{\left\|\vec{a}_2\right\|^2}\vec{a}_2. \qquad (2.5.10)$$

Продолжая этот процесс, для некоторого номера $k < n$ приходим к такому результату:

$$\vec{a}_k = \vec{g}_k - \frac{\left(\vec{a}_1, \vec{g}_k\right)}{\left\|\vec{a}_1\right\|^2}\vec{a}_1 - \frac{\left(\vec{a}_2, \vec{g}_k\right)}{\left\|\vec{a}_2\right\|^2}\vec{a}_2 - \ldots - \frac{\left(\vec{a}_{k-1}, \vec{g}_k\right)}{\left\|\vec{a}_{k-1}\right\|^2}\vec{a}_{k-1}. \qquad (2.5.11)$$

В выражении (2.5.11) вектор $\vec{a}_k \neq \vec{0}$. Действительно, вектор $\vec{a}_k$ является линейной комбинацией векторов

$$\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}, \vec{g}_k.$$

Вектор $\vec{a}_{k-1}$ является линейной комбинацией векторов

$$\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-2}, \vec{g}_{k-1},$$

и так далее. Окончательно имеем

$$\vec{a}_k = \beta_1 \cdot \vec{g}_1 + \beta_2 \cdot \vec{g}_2 + \ldots + \beta_{k-1} \cdot \vec{g}_{k-1} + \vec{g}_k,$$

откуда и следует, что $\vec{a}_k \neq \vec{0}$.

Продолжая процесс до номера $k = n$, получим ортогональный базис $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset E^n$, нормируя который, получим ортонормированный базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$. ••

Следствие из теоремы 2.5.3. *Любую ортонормированную систему векторов евклидова пространства можно дополнить до базиса.*

Два типа координат в евклидовом пространстве. Пусть в пространстве E^n линейно независимая система векторов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ является базисом, то есть число векторов n в системе максимально. Тогда $\left(\forall \vec{x} \in E^n\right)$ имеем единственное представление

$$\vec{x} = \sum_{i=1}^{n} x^k \vec{e}_k.$$

Значение скалярного произведения, в качестве аргументов которого выбираются левая и правая части последнего равенства и базисный вектор $\vec{e}_i$, обозначим

$$x_i = \left(\vec{x}, \vec{e}_i\right).$$

Выясним смысл обозначения величин x_i при помощи нижних индексов, для чего развернём скалярное произведение, используя свойства симметрии и билинейности:

$$x_i = \left(\vec{x}, \vec{e}_i\right) = \left(x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^i \vec{e}_i + \ldots + x^n \vec{e}_n, \vec{e}_i\right) =$$

$$= x^1\left(\vec{e}_1, \vec{e}_i\right) + x^2\left(\vec{e}_2, \vec{e}_i\right) + \ldots + x^i\left(\vec{e}_i, \vec{e}_i\right) + \ldots + x^n\left(\vec{e}_n, \vec{e}_i\right) =$$

$$= \sum_{j=1}^{n} x^j\left(\vec{e}_j, \vec{e}_i\right) = \sum_{j=1}^{n}\left(\vec{e}_i, \vec{e}_j\right)x^j = \sum_{j=1}^{n} g_{ij}x^j.$$

Очевидно, результат вычисления зависит только от нижнего индекса. Мы получили, вообще говоря, некоторый новый тип величин, связанных с вектором $\vec{x}$, которые называются ***ковариантными (евклидовыми) координатами*** вектора. Поясним, почему эти величины называются координатами вектора. Если базис ортогонален, то

$$x_i = \left(\vec{x}, \vec{e}_i\right) = \sum_{j=1}^{n}\left(\vec{e}_i, \vec{e}_j\right)x^j = \sum_{j=1}^{n} g_{ij}x^j = g_{ii}x^i,$$

где $g_{ij} = 0$, если $i \neq j$, и $g_{ii} \neq 0$. Откуда получаем, что $\left(\forall i = 1, 2, \ldots, n\right)$

$$x^i = \frac{1}{g_{ii}} x_i.$$

Подставляя это выражение в разложение вектора $\vec{x}$ по базису, получаем

$$\vec{x} = \frac{x_1}{g_{11}}\vec{e}_1 + \frac{x_2}{g_{22}}\vec{e}_2 + \ldots + \frac{x_n}{g_{nn}}\vec{e}_n.$$

Если базис к тому же нормирован, то

$$g_{11} = g_{22} = \ldots g_{nn} = 1$$

и получаем $x_i = x^i$ и

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \ldots + x_n\vec{e}_n.$$

Видим, что аффинные и ковариантные координаты вектора совпадают только в фиксированном ***ортонормированном базисе*** (в этом случае координаты вектора иногда называются ***физическими координатами***).

Итак, ***в евклидовом пространстве для произвольного вектора в фиксированном ортонормированном базисе*** справедливо следующее ***ортогональное разложение***:

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{e}_i = \sum_{i=1}^{n} x_i \vec{e}_i = \sum_{i=1}^{n} \left(\vec{x}, \vec{e}_i\right) \vec{e}_i \equiv$$

$$\equiv \left(\vec{x}, \vec{e}_1\right) \vec{e}_1 + \left(\vec{x}, \vec{e}_2\right) \vec{e}_2 + \ldots + \left(\vec{x}, \vec{e}_n\right) \vec{e}_n. \qquad (2.5.12)$$

Подчеркнём, что евклидовы (ковариантные) координаты вектора ***существуют только в евклидовом пространстве***. В аффинном пространстве у любого вектора существуют только координаты с верхними индексами, которые выше были названы ***аффинными координатами***. Аффинные координаты называются также ***контравариантными*** координатами.

2.6. Координатное пространство R^n

Пространство вектор-столбцов. В качестве примера многомерного пространства рассмотрим так называемое ***координатное пространство***, или пространство ***вектор-столбцов***.

Пусть n — натуральное число. Рассмотрим вектор-столбцы вида

$$|x\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^n \end{pmatrix}, \qquad (2.6.1)$$

где ***координаты*** x^k вектор-столбца являются действительными числами, то есть $\left(\forall k = 1, 2, \ldots, n\right) x^k \in R^1$. Рассматриваются также ***транспонированные векторы строки*** $\langle x| = \begin{pmatrix} x^1 & x^2 & \ldots & x^n \end{pmatrix}$.

Отношение равенства и операции между вектор-столбцами (2.6.1) определяются аналогично случаю трёх измерений.

Определение 2.6.1. 1) *Два вектор-столбца считаются* ***равными*** *и пишут*

$$\left|x\right\rangle = \left|y\right\rangle \Rightarrow \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} = \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix},$$

если

$$\left(\forall k = 1, 2, \dots, n\right) x^k = y^k. \qquad (2.6.2)$$

2) ***Суммой*** *вектор-столбцов* $\left|x\right\rangle$, $\left|y\right\rangle$ *и* ***произведением*** *вектор-столбца* $\left|x\right\rangle$ *на действительное число* α *называются вектор-столбцы, определяемые по правилам:*

$$\left|x\right\rangle + \left|y\right\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} + \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix} = \begin{pmatrix} x^1 + y^1 \\ x^2 + y^2 \\ \dots \\ x^n + y^n \end{pmatrix}, \qquad (2.6.3)$$

$$\alpha \cdot \left|x\right\rangle = \begin{pmatrix} \alpha \cdot x^1 \\ \alpha \cdot x^2 \\ \dots \\ \alpha \cdot x^n \end{pmatrix}. \bullet \qquad (2.6.4)$$

Операция сложения вектор-столбцов ***коммутативна*** и ***ассоциативна***, то есть

$$1) \left|x\right\rangle + \left|y\right\rangle = \left|y\right\rangle + \left|x\right\rangle, \; 2) \left(\left|x\right\rangle + \left|y\right\rangle\right) + \left|z\right\rangle = \left|x\right\rangle + \left(\left|y\right\rangle + \left|z\right\rangle\right). \qquad (2.6.5)$$

Вектор-столбец, составленный из нулей, называется ***нуль-вектором***:

$$\left|0\right\rangle \overset{def}{=} \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}. \qquad (2.6.6)$$

Очевидно, что для любого вектор-столбца (2.6.1) справедливо соотношение

$$\left|x\right\rangle + \left|0\right\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} = \left|x\right\rangle. \qquad (2.6.7)$$

Можно показать, что нуль-вектор (2.6.5) единственен.

Вектор-столбец вида

$$-|x\rangle = \begin{pmatrix} -x^1 \\ -x^2 \\ \dots \\ -x^n \end{pmatrix}$$

называется ***обратным*** (***противоположным***) вектору (2.6.1).

Разность $|z\rangle = |x\rangle - |y\rangle$ двух вектор-столбцов определяется как сумма вектор-столбца $|x\rangle$ и вектор-столбца $(-1)\cdot|y\rangle$:

$$|z\rangle = |x\rangle - |y\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} - \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix} \overset{def}{=} \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} + \begin{pmatrix} -y^1 \\ -y^2 \\ \dots \\ -y^n \end{pmatrix} = \begin{pmatrix} x^1 - y^1 \\ x^2 - y^2 \\ \dots \\ x^n - y^n \end{pmatrix}.$$

Путём непосредственных выкладок из выражения (2.6.3) можно заключить, что справедливы следующие правила выполнения операции умножения вектор-столбца на число: $\left(\forall\, |x\rangle, |y\rangle \wedge \forall \alpha, \beta \in R^1\right)$

1) $1\cdot|x\rangle = |x\rangle$

2) $\alpha\cdot\left(|x\rangle + |y\rangle\right) = \alpha\cdot|x\rangle + \alpha\cdot|y\rangle$ (2.6.8)

3) $(\alpha+\beta)\cdot|x\rangle = \alpha\cdot|x\rangle + \beta\cdot|x\rangle$

4) $\alpha\cdot\left(\beta\cdot|x\rangle\right) = (\alpha\cdot\beta)\cdot|x\rangle$

Определение 2.6.2. *Множество вектор-столбцов с введёнными выше операциями (2.6.2) — (2.6.4), удовлетворяющими условиям (2.6.5) — (2.6.8), называется* ***действительным координатным векторным пространством*** R^n. •

Скалярное произведение в координатном пространстве. Во множестве вектор-столбцов аналогично случаю трёх измерений вводится понятие ***скалярного произведения — действительной функции двух векторных переменных***, превращающего его в ***евклидово координатное пространство n измерений***.

Определение 2.6.3. *Пусть* $|x\rangle$ *и* $|y\rangle$ *два вектор-столбца вида (2.6.1). Действительная функция* $\varphi : R^n \times R^n \to R^1$ *двух векторных переменных, действие которой определено правилом*

$$\varphi(|x\rangle, |y\rangle) = \begin{pmatrix} x^1 & x^2 & \dots & x^n \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix} \overset{def}{=} \sum_{k=1}^{n} x^k y^k, \qquad (2.6.9)$$

называется ***скалярным произведением*** *векторов* $|x\rangle$ *и* $|y\rangle$. •

Для скалярного произведения используется также следующее обозначение:

$$\varphi(|x\rangle, |y\rangle) \equiv \langle x|y\rangle.$$

В пространстве R^n справедливы аналогичные случаю трёх измерений свойства скалярного произведения векторов, устанавливаемые следующей теоремой.

Теорема 2.6.1. *Скалярное произведение (2.6.8) обладает следующими свойствами:* $\left(\forall |x\rangle, |y\rangle, |z\rangle \in R^n\right) \wedge \left(\forall \alpha \in R^1\right)$

$$\begin{aligned} &1)\ \langle x|y\rangle = \langle y|x\rangle; \\ &2)\ \langle x+y|z\rangle = \langle x|z\rangle + \langle y|z\rangle; \\ &3)\ \langle \alpha \cdot x|y\rangle = \alpha \cdot \langle x|y\rangle; \\ &4)\ \langle x|x\rangle > 0; \langle x|x\rangle = 0 \Leftrightarrow |x\rangle \equiv |0\rangle. \end{aligned} \qquad (2.6.10)$$

Доказательство этих свойств не составляет труда.

Определение 2.6.4. *Множество всех вектор-столбцов вида (2.6.1), для которых определены действия сложения (2.6.3), умножения на действительные числа (2.6.4) и функция скалярного произведения (2.6.9), называется* ***n-мерным (координатным) евклидовым пространством*** R^n. •

Вводится понятие ортогональных векторов.

Определение 2.6.5. *Два вектор-столбца* $|x\rangle, |y\rangle \in R^n$ *называются* ***ортогональными****, если*

$$\langle x|y\rangle = \sum_{k=1}^{n} x^k y^k = 0. \bullet \qquad (2.6.11)$$

Линейная зависимость и линейная независимость системы вектор-столбцов, базис. Важнейшим является ***свойство системы векторов***, сформулированное в следующем определении.

Определение 2.6.6. *Система вектор-столбцов*

$$\left\{|x_1\rangle, |x_2\rangle, \dots, |x_m\rangle\right\} \in R^n$$

*называется **линейно зависимой системой**, если можно подобрать числа* $\alpha_1, \alpha_2, \ldots, \alpha_m \in R^1$*, не все равные нулю одновременно, таким образом, чтобы выполнялось равенство*

$$\alpha_1 |x_1\rangle + \alpha_2 |x_2\rangle + \ldots + \alpha_m |x_m\rangle = |0\rangle. \quad (2.6.12)$$

Если таких чисел подобрать нельзя, то есть равенство (2.6.12) выполняется лишь при условии $\alpha_1 = \alpha_2 = \ldots = \alpha_m = 0$*, то система вектор-столбцов называется **линейно независимой**.* •

Если система n векторов $\{|x_1\rangle, |x_2\rangle, \ldots, |x_n\rangle\} \in R^n$ линейно независима, а любая система $\{|x_1\rangle, |x_2\rangle, \ldots, |x_n\rangle, |x_{n+1}\rangle\}$, где $|x_{n+1}\rangle \in R^n$, является линейно зависимой, то говорят, что система $\{|x_1\rangle, |x_2\rangle, \ldots, |x_n\rangle\}$ является ***базисом пространства*** R^n.

Лемма 2.6.1. *Пусть* $\{|a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle\} \in R^n$ *— как-либо выбранный базис пространства* R^n*. Тогда любой вектор* $|y\rangle \in R^n$ *можно единственным образом представить в виде **линейной комбинации** базисных векторов*

$$|y\rangle = y^1 |a_1\rangle + y^2 |a_2\rangle + \ldots + y^n |a_n\rangle. \quad (2.6.13)$$

Д о к а з а т е л ь с т в о. По определению базиса система

$$\{|a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle, |y\rangle\}$$

является линейно зависимой. Поэтому можно подобрать не равные нулю одновременно числа $\alpha_1, \alpha_2, \ldots, \alpha_m \in R^n$ так, чтобы выполнялось равенство

$$\alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle + \ldots + \alpha_n |a_n\rangle + \alpha |y\rangle = |0\rangle. \quad (2.6.14)$$

Здесь обязательно $\alpha \neq 0$, в противном случае получаем

$$\alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle + \ldots + \alpha_n |a_n\rangle = |0\rangle,$$

где не все числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ равны нулю одновременно. А это противоречит линейной независимости системы векторов $\{|a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle\}$. Прибавляя к обеим частям равенства (2.6.14) вектор $-(\alpha_1 |a_1\rangle + \alpha_2 |a_2\rangle + \ldots + \alpha_n |a_n\rangle)$ и умножая результат на $\frac{1}{\alpha}$, имеем:

$$|y\rangle = -\frac{\alpha_1}{\alpha}|a_1\rangle - \frac{\alpha_2}{\alpha}|a_2\rangle - \ldots - \frac{\alpha_n}{\alpha}|a_n\rangle.$$

Обозначив $y^k = -\dfrac{\alpha_k}{\alpha}$ $(k = 1, 2, \ldots, n)$, получим требуемый результат (2.6.13).

Пусть вектор-столбец $|y\rangle$ имеет два представления вида (2.6.12):

$$|y\rangle = \xi^1 |a_1\rangle + \xi^2 |a_2\rangle + \ldots + \xi^n |a_n\rangle;$$

$$|y\rangle = \eta^1 |a_1\rangle + \eta^2 |a_2\rangle + \ldots + \eta^n |a_n\rangle.$$

Вычитая эти два равенства, получаем

$$\left(\xi^1 - \eta^1\right) \cdot |a_1\rangle + \left(\xi^2 - \eta^2\right) \cdot |a_2\rangle + \ldots + \left(\xi^n - \eta^n\right) \cdot |a_n\rangle = |0\rangle,$$

что в силу линейной независимости системы векторов базиса даёт

$$\xi^1 = \eta^1, \xi^2 = \eta^2, \ldots, \xi^n = \eta^n.$$

Таким образом, представление (2.6.12) единственно.••

Числа $y^1, y^2, \ldots, y^n$ в разложении (2.6.12) вектор-столбца $|y\rangle$ по векторам базиса $\{|a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle\}$ называются его ***координатами относительно данного базиса***.

Рассмотрим в пространстве R^n вектор-столбцы вида

$$|e_1\rangle = \begin{pmatrix} 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix}, |e_2\rangle = \begin{pmatrix} 0 \\ 1 \\ \ldots \\ 0 \end{pmatrix}, \ldots, |e_n\rangle = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 1 \end{pmatrix}. \qquad (2.6.15)$$

Лемма 2.6.2. *Система вектор-столбцов (2.6.15) линейно независима.*

Д о к а з а т е л ь с т в о. Составим линейную комбинацию векторов (2.6.12) с некоторыми коэффициентами $\alpha_1, \alpha_2, \ldots, \alpha_n \in R^1$ и потребуем выполнения равенства

$$\alpha_1 |e_1\rangle + \alpha_2 |e_2\rangle + \ldots + \alpha_n |e_n\rangle = |0\rangle. \qquad (2.6.16)$$

Пользуясь правилами сложения, умножения на число и определением равенства векторов-столбцов, легко видеть, что выражение (2.6.15) для векторов (2.6.14) сводится к следующим равенствам:

$$\alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_n = 0.$$

Следовательно, условием выполнения системы (2.6.15) является одновременное обращение коэффициентов $\alpha_1, \alpha_2, \ldots, \alpha_n$ в нуль, что доказывает линейную независимость системы векторов (2.6.14). ••

Нетрудно видеть, что для любого вектор-столбца $|x\rangle \in R^n$ справедливо разложение

$$|x\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^n \end{pmatrix} = x^1 \begin{pmatrix} 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix} + x^2 \begin{pmatrix} 0 \\ 1 \\ \ldots \\ 0 \end{pmatrix} + \ldots + x^n \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 1 \end{pmatrix} = x^1 |e_1\rangle + x^2 |e_2\rangle + \ldots + x^n |e_n\rangle.$$

Нетрудно показать, что векторы (2.6.15) ***попарно ортогональны***.

Норма вектор-столбца в координатном пространстве. Определение нормы (длины) вектора в пространстве R^n вводится в полной аналогии со случаем векторов трёхмерного евклидова пространства.

Определение 2.6.7. *Отображение пространства R^n во множество действительных чисел R, ставящее в соответствие каждому вектор-столбцу $|x\rangle \in R^n$ действительное число $\||x\rangle\|$ по правилу*

$$\||x\rangle\| \overset{def}{=} \sqrt{\langle x|x\rangle} \equiv \langle x|x\rangle^{1/2}, \tag{2.6.17}$$

*называется **евклидовой нормой**.*•

Теорема 2.6.2 (неравенство Коши — Буняковского). *Для любых вектор-столбцов $|x\rangle, |y\rangle \in R^n$ всегда выполняется неравенство*

$$\langle x|y\rangle^2 \le \langle x|x\rangle \cdot \langle y|y\rangle, \tag{2.6.18}$$

причём равенство имеет место в том и только в том случае, если система векторов $\{|x\rangle, |y\rangle\}$ линейно зависима.

Теорема 2.6.3. *Евклидова норма обладает следующими свойствами:*

$$\begin{aligned} &1)\ \||x\rangle\| > 0 \Leftrightarrow |x\rangle \ne |0\rangle \wedge \||0\rangle\| = 0; \\ &2)\ \left(\forall |x\rangle \in R^n \wedge \forall \beta \in R^1\right) \|\beta \cdot |x\rangle\| = |\beta| \cdot \||x\rangle\|; \\ &3)\ \left(\forall |x\rangle, |y\rangle \in R^n\right) |\langle x|y\rangle| \le \||x\rangle\| \cdot \||y\rangle\|; \\ &4)\ \left(\forall |x\rangle, |y\rangle \in R^n\right) \||x\rangle + |y\rangle\| \le \||x\rangle\| + \||y\rangle\|. \end{aligned} \tag{2.6.19}$$

Д о к а з а т е л ь с т в о. Свойства 1 и 2 очевидным образом следуют из свойств скалярного произведения и определения евклидовой нормы. Свойство 3 является следствием теоремы 2.6.2. Для доказательства свойств 4 рассмотрим квадрат нормы:

$$\left\| |x\rangle + |y\rangle \right\|^2 = \left(|x\rangle + |y\rangle, |x\rangle + |y\rangle \right) = \langle x|x\rangle + 2\langle x|y\rangle + \langle y|y\rangle \le$$

$$\le \langle x|x\rangle + 2\left|\langle x|y\rangle\right| + \langle y|y\rangle \le \langle x|x\rangle + 2\langle x|x\rangle^{1/2} \cdot \langle y|y\rangle^{1/2} + \langle y|y\rangle =$$

$$= \left(\langle x|x\rangle^{1/2} + \langle x|x\rangle^{1/2} \right)^2.$$

Здесь использовано неравенство Коши — Буняковского (2.6.17). Извлекая из обеих частей последнего равенства квадратный корень, получаем свойство 4.••

Определение 2.6.8. *Пусть $|x\rangle \in R^n$. Если $\langle x|x\rangle = 1$, то вектор-столбец $|x\rangle$ называется* ***нормированным****. Система, состоящая из попарно ортогональных нормированных вектор-столбцов, называется* ***ортонормированной****.•*

Справедливо следующее утверждение, доказательство которого для случая абстрактного евклидова пространства E^n приведено ниже.

Лемма 2.6.3. *Система попарно ортогональных (ортонормированных) вектор-столбцов $\{|x_1\rangle, |x_2\rangle, \dots, |x_m\rangle\}$ евклидова пространства R^n линейно независима.*

Нетрудно проверить, что система векторов (2.6.15) ортонормированна. Говорят, что векторы (2.6.15) образуют ***ортонормированный базис пространства*** R^n, который называется ***каноническим*** (простейшим) или ***стандартным***.

Определение 2.6.9. ***Евклидовым расстоянием*** *между точками (векторами) $|x\rangle, |y\rangle \in R^n$ называется величина*

$$\rho\left(|x\rangle, |y\rangle\right) \equiv dist\left(|x\rangle, |y\rangle\right) \overset{def}{=} \left\| |x\rangle - |y\rangle \right\| = \langle x-y|x-y\rangle^{1/2}, \qquad (2.6.20)$$

где $\left\| |x\rangle - |y\rangle \right\|$ — евклидова норма, определённая формулой (2.6.16).•

Теорема 2.6.4. *Евклидово расстояние в пространстве R^n обладает следующими свойствами:*

1) $\rho\left(|x\rangle, |y\rangle\right) > 0 \Leftrightarrow |x\rangle \ne |y\rangle \wedge \rho\left(|x\rangle, |y\rangle\right) = 0 \Leftrightarrow |x\rangle = |y\rangle$;

2) $\rho\left(|x\rangle, |y\rangle\right) = \rho\left(|y\rangle, |x\rangle\right)$;

3) $\rho\left(|x\rangle, |y\rangle\right) \le \rho\left(|x\rangle, |z\rangle\right) + \rho\left(|z\rangle, |y\rangle\right)$.

Доказательство теоремы не представляет особого труда и предлагается в качестве упражнения.

Расстояние, угол и проекция в координатном пространстве. С учётом определения (2.6.17) евклидово расстояние (2.6.20), очевидно, записывается в виде:

$$\rho\left(|x\rangle, |y\rangle\right) = \sqrt{\left(x^1 - y^1\right)^2 + \left(x^2 - y^2\right)^2 + \ldots + \left(x^n - y^n\right)^2}.$$

Определение 2.6.10. ***Углом*** $\{|x\rangle, |y\rangle\}$ ***между ненулевыми векторами*** $|x\rangle$ ***и*** $|y\rangle$ *евклидова пространства* R^n *называется величина, определяемая соотношением*

$$\cos\{|x\rangle, |y\rangle\} = \frac{\langle x|y\rangle}{\langle x|x\rangle^{1/2} \cdot \langle y|y\rangle^{1/2}}, \tag{2.6.21}$$

причём предполагается, что $0 \le \{|x\rangle, |y\rangle\} \le \pi$. •

Нетрудно видеть, что

$$\cos\{|x\rangle, |y\rangle\} = \frac{x^1 y^1 + x^2 y^2 + \ldots + x^n y^n}{\sqrt{\left(x^1\right)^2 + \ldots + \left(x^n\right)^2} \cdot \sqrt{\left(y^1\right)^2 + \ldots + \left(y^n\right)^2}}.$$

По аналогии с трёхмерным случаем величина

$$\mathrm{Pr}_{|y\rangle}\{|x\rangle\} = \langle x|x\rangle^{\frac{1}{2}} \cdot \cos\{|x\rangle, |y\rangle\} = \frac{1}{\langle y|y\rangle^{1/2}} \cdot \langle x|y\rangle \tag{2.6.22}$$

называется ***проекцией вектор-столбца*** $|x\rangle$ ***на направление вектор-столбца*** $|y\rangle$. В формуле (2.6.21) косинус угла между вектор-столбцами в пространстве R^n определяется в соответствии с формулой (2.6.20).

2.7. Комплексное евклидово пространство

Унитарное пространство. Векторные пространства (в том числе аффинные и евклидовы) можно определять не только над полем действительных чисел. Все метрические понятия для евклидовых пространств распространяются и на ***комплексное евклидово пространство***.

Определение 2.7.1. *Векторное пространство U, определённое над полем комплексных чисел C, называется* ***унитарным****, если каждой паре векторов* $\vec{x}$ *и* $\vec{y}$ *этого пространства ставится в соответствие комплексное число* $\left(\vec{x}, \vec{y}\right)$, *которое называется* ***скалярным произведением*** *и удовлетворяет следующим аксиомам:*

$$\left(\forall \vec{x}, \vec{y}, \vec{z} \in U\right) \wedge (\forall \lambda \in C)$$

1) $\left(\vec{x}, \vec{y}\right) = \overline{\left(\vec{y}, \vec{x}\right)}$;

2) $\left(\vec{x}, \lambda \cdot \vec{y}\right) = \lambda \cdot \left(\vec{x}, \vec{y}\right)$;

3) $\left(\vec{x} + \vec{y}, \vec{z}\right) = \left(\vec{x}, \vec{z}\right) + \left(\vec{y}, \vec{z}\right)$;

4) $\left(\vec{x}, \vec{x}\right) > 0$ при $\vec{x} \neq \vec{0}$ и $\left(\vec{0}, \vec{0}\right) = 0$. •

В первой аксиоме черта над правой частью равенства означает комплексное сопряжение. Это изменение по сравнению с вещественным евклидовым пространством не влечёт за собой никаких существенных отличий, но некоторые особенности имеются. Так, в отличие от вещественного случая, в унитарном пространстве имеем:

$$\left(\lambda \cdot \vec{x}, \vec{y}\right) = \overline{\left(\vec{y}, \lambda \cdot \vec{x}\right)} = \bar{\lambda} \cdot \overline{\left(\vec{y}, \vec{x}\right)} = \bar{\lambda} \cdot \left(\vec{x}, \vec{y}\right).$$

В унитарном пространстве (как и в вещественном) вводится ***норма (длина) вектора***

$$\left\|\vec{x}\right\| \stackrel{def}{=} +\left(\vec{x}, \vec{x}\right)^{\frac{1}{2}}.$$

Норма каждого ненулевого вектора положительна, норма нулевого вектора равна нулю. Для любого комплексного числа λ справедливо равенство

$$\left\|\lambda \cdot \vec{x}\right\| = |\lambda| \cdot \left\|\vec{x}\right\|,$$

где $|\lambda|$ — ***модуль*** комплексного числа λ.

Справедливо также неравенство Коши — Буняковского

$$\left|\left(\vec{x},\vec{y}\right)\right|^2 \le \left(\vec{x},\vec{x}\right)\cdot\left(\vec{y},\vec{y}\right),$$

где в левой части записан модуль скалярного произведения. Неравенство Коши — Буняковского доказывается аналогично вещественному случаю.

Понятие угла в унитарном пространстве не вводится, рассматривается только случай ортогональных векторов. По аналогии с вещественным случаем ***векторы называют ортогональными, если их скалярное произведение равно нулю, то есть***

$$\left(\vec{x},\vec{y}\right)=0.$$

Вся теория евклидовых пространств без существенных изменений переносится на случай унитарного пространства.

Комплексное координатное пространство. В качестве типичного примера комплексного евклидова (унитарного) пространства рассмотрим ***координатное*** (***арифметическое***) ***пространство*** C^n, вектор-столбцы которого обозначим

$$|x\rangle \overset{def}{=} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \dots \\ \alpha^n \end{pmatrix}.$$

В пространстве C^n вместо операции транспонирования используется двойная операция — сначала транспонирование, а потом комплексное сопряжение. Поэтому вектор-строка называется (***комплексно***) ***сопряжённым вектором*** к вектор-столбцу и обозначается так:

$$|x\rangle^* = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \dots \\ \alpha^n \end{pmatrix}^* = \left(\overline{\alpha^1}\ \overline{\alpha^2} \dots \overline{\alpha^n}\right) = \langle x|.$$

Скалярное произведение двух векторов

$$|x\rangle = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \dots \\ \alpha^n \end{pmatrix}, |y\rangle = \begin{pmatrix} \beta^1 \\ \beta^2 \\ \dots \\ \beta^n \end{pmatrix}$$

определяется правилом

$$\langle x|y\rangle \overset{def}{=} \begin{pmatrix} \overline{\alpha^1} & \overline{\alpha^2} & \dots & \overline{\alpha^n} \end{pmatrix} \begin{pmatrix} \beta^1 \\ \beta^2 \\ \dots \\ \beta^n \end{pmatrix} = \sum_{i=1}^{n} \overline{\alpha^i} \cdot \beta^i .$$

Предположим, что мы определили скалярное произведение без использования комплексного сопряжения. Тогда, например, для вектора

$$|x\rangle = \begin{pmatrix} 0 \\ 3 \\ 3i \end{pmatrix}$$

при вычислении скалярного квадрата получаем

$$\langle x|x\rangle = \begin{pmatrix} 0 & 3 & 3i \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 3i \end{pmatrix} = 0 + 9 + 9i^2 = 0.$$

Видим, что аксиома невырожденности не выполняется.

Вычислим теперь скалярный квадрат, используя определение с комплексным сопряжением:

$$\langle x|x\rangle = \begin{pmatrix} \overline{0} & \overline{3} & \overline{3i} \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ 3i \end{pmatrix} = 0 \cdot 0 + 3 \cdot 3 + (-3i)3i = 9 + 9 = 18.$$

Таким образом, использование комплексного сопряжения при определении скалярного произведения в унитарном пространстве существенно.

2.8. Строение векторных пространств. Изоморфизм

В этом параграфе приведены некоторые общие понятия теории конечномерных векторных пространств, которые существенны для понимания материала последующих глав. Кроме этого, излагаемые понятия имеют широкие приложения в различных предметных областях.

Подпространства векторного пространства. Напомним введённое выше понятие линейной оболочки системы векторов $\{\vec{x}_1, \vec{x}_2, ..., \vec{x}_m\}$.

Определение 2.8.1. ***Линейной оболочкой*** *системы векторов*

$$\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$$

n-мерного векторного (абстрактного, аффинного или евклидова) пространства A^n называется множество значений линейной комбинации

$$\vec{y} = \alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2 + \ldots + \alpha_m \cdot \vec{x}_m \tag{2.8.1}$$

векторов этой системы при всевозможных значениях, принимаемых коэффициентами $\alpha_1, \alpha_2, \ldots, \alpha_m$ из поля P. •

Для линейной оболочки применяется обозначение

$$L\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}.$$

Лемма 2.8.1. *Линейная оболочка любой системы векторов из любого векторного (абстрактного, аффинного или евклидова) пространства сама является векторным (абстрактным векторным, аффинным или евклидовым) пространством.*

Доказательство. Выполнение аксиом $1^0 - 9^0$ аффинного пространства проверяется без особого труда. Поэтому мы дадим только разъяснения относительно аксиом, постулирующих существование нулевого и обратного векторов.

Нулевой вектор заведомо принадлежит любой линейной оболочке и получается при нулевых значениях её коэффициентов, то есть

$$\vec{0} = 0\vec{x}_1 + 0\vec{x}_2 + \ldots + 0\vec{x}_m.$$

Вектор, обратный вектору (2.8.1), имеет следующий вид:

$$-\vec{y} = (-\alpha_1)\vec{x}_1 + (-\alpha_2)\vec{x}_2 + \ldots + (-\alpha_m)\vec{x}_m.$$

Действительно, используя правило сложения и умножения векторов на числа, получаем

$$\vec{0} = \vec{y} + (-\vec{y}) = (\alpha_1 - \alpha_1)\vec{x}_1 + (\alpha_2 - \alpha_2)\vec{x}_2 + \ldots + (\alpha_m - \alpha_m)\vec{x}_m = \vec{0}. \bullet\bullet$$

Из леммы 2.8.1 следует, что каждое векторное пространство содержит в себе бесконечное множество других векторных пространств — линейных оболочек всевозможных систем векторов из данного пространства. Само ***вмещающее*** векторное пространство является линейной оболочкой векторов некоторого своего базиса $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$.

Определение 2.8.2. *Пусть A^n — некоторое n-мерное векторное (абстрактное, аффинное или евклидово — всё равно) пространство и $L \subset A^n$ — некоторое подмножество векторов из пространства A^n. Если при тех же операциях над векторами, что и во всём пространстве A^n, множество L само является векторным пространством, то оно называется* ***подпространством пространства*** *A^n.* •

Наименьшим подпространством пространства A^n является подпространство, состоящее только из нулевого вектора. ***Наибольшим*** подпространством векторного пространства A^n является само это пространство. Эти два подпространства пространства называются ***тривиальными подпространствами*** рассматриваемого пространства.

Как определить, является ли некоторое множество векторов $L \subset A^n$ подпространством? Ответ на этот вопрос содержится в следующей теореме.

Теорема 2.8.1. *Если подмножество L векторов векторного (абстрактного, аффинного или евклидова) пространства A^n вместе с каждой парой своих векторов $\vec{x}$ и $\vec{y}$ содержит и все их линейные комбинации вида $\alpha \cdot \vec{x} + \beta \cdot \vec{y}$, то оно является подпространством пространства A^n.*

Д о к а з а т е л ь с т в о. Пусть $\alpha = \beta = 0$. Тогда

$$0 \cdot \vec{x} + 0 \cdot \vec{y} = \vec{0},$$

то есть нулевой вектор $\vec{0} \in L$.

Пусть $\alpha = -1, \beta = 0$. Тогда имеем:

$$(-1) \cdot \vec{x} + 0 \cdot \vec{y} = (-1) \cdot \vec{x}.$$

Поэтому во множество L вместе с каждым вектором $\vec{x}$ входит и обратный ему вектор. ••

Пусть A^n есть n-мерное векторное пространство. В любом его подпространстве можно построить базис. Если в пространстве A^n выбран некоторый базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, то, вообще говоря, базисные векторы подпространства $L \subset A^n$ нельзя выбрать прямо из числа векторов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, так как последние просто могут не содержаться в подпространстве L. Однако, справедлива следующая лемма.

Лемма 2.8.2. *Пусть L^s является s-мерным подпространством n-мерного векторного пространства A^n. Если в подпространстве L^s выбран некоторый (произвольный) базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s\right\}$, то его всегда можно дополнить до базиса всего пространства A^n*

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \ldots, \vec{e}_n\right\},$$

присоединив выбранные соответствующим образом векторы $\vec{e}_{s+1}, \ldots, \vec{e}_n$.

Доказательство. Рассмотрим лишь те линейно независимые системы векторов из пространства A^n, которые содержат в себе векторы $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s$. Среди этих систем векторов есть система векторов

$$L_p = \left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \ldots, \vec{e}_p\right\},$$

содержащая максимальное число векторов. Но тогда по аксиоме 5^0 система векторов

$$L_{p+1} = \left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \ldots, \vec{e}_p, \vec{x}\right\},$$

где $\vec{x}$ — любой вектор из пространства A^n, не содержащийся в системе

$$L_p = \left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \ldots, \vec{e}_p\right\},$$

является линейно зависимой. Последнее означает, что вектор $\vec{x}$ линейно выражается через векторы системы $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \ldots, \vec{e}_p$, то есть

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^s \vec{e}_s + x^{s+1} \vec{e}_{s+1} + \ldots + x^p \vec{e}_p.$$

Следовательно, система

$$L_p = \left\{ \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{e}_{s+1}, \ldots, \vec{e}_p \right\}$$

является базисом пространства A^n, причём $p \equiv n$. ••

Над подпространствами определяются ***операции***.

Определение 2.8.3. ***Суммой*** *$L_1 + L_2$ подпространств L_1, L_2 n-мерного векторного пространства A^n* (абстрактного, аффинного или евклидова — всё равно) *называется множество всех векторов вида*

$$\vec{z} = \vec{x} + \vec{y},$$

где $\vec{x} \subset L_1$, $\vec{y} \subset L_2$, то есть

$$L_1 + L_2 = \left\{ \vec{z} = \vec{x} + \vec{y} \in A^n : \vec{x} \in L_1 \wedge \vec{y} \in L_2 \right\}. \bullet$$

Определение 2.8.4. ***Пересечением*** *$L_1 \cap L_2$ подпространств $L_1, L_2 \subset A^n$ называется множество векторов, одновременно принадлежащих как L_1, так и L_2, то есть множество векторов*

$$L_1 \cap L_2 = \left\{ \vec{z} \in A^n : \vec{z} \in L_1 \wedge \vec{z} \in L_2 \right\}. \bullet$$

Теорема 2.8.2. *Если L_1, L_2 — подпространства пространства A^n, то $L_1 + L_2$ и $L_1 \cap L_2$ не пусты и также являются подпространствами пространства A^n.*

Доказательство. Так как L_1 и L_2 являются подпространствами пространства A^n, то они содержат нулевой вектор $\vec{0}$. Очевидно, что $\vec{0} \in L_1 + L_2$ и $\vec{0} \in L_1 \cap L_2$.

Покажем, что $L_1 + L_2$ и $L_1 \cap L_2$ являются подпространствами.

Пусть $\vec{z}_1, \vec{z}_2 \in L_1 + L_2$. Тогда

$$\vec{z}_1 = \vec{x}_1 + \vec{y}_1,\ \vec{z}_2 = \vec{x}_2 + \vec{y}_2,$$

где $\vec{x}_1, \vec{x}_2 \in L_1$ и $\vec{y}_1, \vec{y}_2 \in L_2$. Рассмотрим произвольную линейную комбинацию

$$\alpha\vec{z}_1+\beta\vec{z}_2=(\alpha\vec{x}_1+\beta\vec{x}_2)+(\alpha\vec{y}_1+\beta\vec{y}_2).$$

Так как L_1 и L_2 являются подпространствами, то $\alpha\vec{x}_1+\beta\vec{x}_2\in L_1$ и $\alpha\vec{y}_1+\beta\vec{y}_2\in L_2$ и линейная комбинация

$$\alpha\vec{z}_1+\beta\vec{z}_2\in L_1+L_2.$$

Поэтому L_1+L_2 — подпространство A^n.

Пусть теперь $\vec{z}_1, \vec{z}_2\in L_1\cap L_2$, то есть $\vec{z}_1, \vec{z}_2\in L_1$ и $\vec{z}_1, \vec{z}_2\in L_2$. Тогда, так как L_1 и L_2 являются подпространствами, выполняются включения

$$\alpha\vec{z}_1+\beta\vec{z}_2\in L_1,\ \alpha\vec{z}_1+\beta\vec{z}_2\in L_2,$$

то есть

$$\alpha\vec{z}_1+\beta\vec{z}_2\in L_1\cap L_2.$$

Следовательно, $L_1\cap L_2$ тоже является подпространством A^n. ••

Из доказанной теоремы следует, что операции сложения подпространств и их пересечения являются ***алгебраическими операциями***. Они ***коммутативны*** и ***ассоциативны***. Кроме этого,

$$\left(\forall L\subset A^n\right)\ L+\vec{0}=L,\ L\cap A^n=L.$$

Однако дистрибутивные законы отсутствуют.

Прежде чем сформулировать следующую теорему, заметим, что дальше мы будем обозначать размерность n-мерного векторного пространства при помощи записи $\dim A=n$.

Теорема 2.8.3. *Для любых двух подпространств L_1 и L_2 векторного пространства A^n имеет место согласование размерностей следующего вида*:

$$\dim(L_1\cap L_2)+\dim(L_1+L_2)=\dim L_1+\dim L_2.$$

Д о к а з а т е л ь с т в о. Обозначим $L_1+L_2=R$ и $L_1\cap L_2=T$. Размерности подпространств обозначим малыми буквами:

$$\dim(L_1+L_2)=r,\ \dim(L_1\cap L_2)=t,\ \dim L_1=p,\ \dim L_2=q.$$

Выберем базис пересечения подпространств

$$T=L_1\cap L_2,$$

пусть его составляют векторы $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_t\right\}$. По определению 2.8.4 имеем: $T \subset L_1$ и $T \subset L_2$. По лемме 2.8.2 базис подпространства T можно дополнить как до базиса подпространства L_1, так и до базиса подпространства L_2. Поэтому предположим, что

$$L_p = \left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_s, \vec{c}_{t+1}, \ldots, \vec{c}_p\right\}$$

есть базис подпространства L_1, а

$$L_q = \left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_t, \vec{s}_{t+1}, \ldots, \vec{s}_q\right\}$$

есть базис подпространства L_2.

Покажем, что система

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_t, \vec{c}_{t+1}, \ldots, \vec{c}_p, \vec{s}_{t+1}, \ldots, \vec{s}_q\right\}$$

является базисом подпространства $R = L_1 + L_2$.

По определению 2.8.3 для любого вектора $\vec{z} \in R$ имеем представление $\vec{z} = \vec{x} + \vec{y}$, где $\vec{x} \in L_1$, $\vec{y} \in L_2$. Поэтому получаем:

$$\vec{z} = \vec{x} + \vec{y} = \left(x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^t \vec{e}_t + x^{t+1} \vec{c}_{t+1} + \ldots + x^p \vec{c}_p\right) +$$
$$+\left(y^1 \vec{e}_1 + y^2 \vec{e}_2 + \ldots + y^t \vec{e}_t + y^{t+1} \vec{s}_{t+1} + \ldots + y^q \vec{s}_q\right).$$

Видим, что подпространство $R = L_1 + L_2$ является линейной оболочкой системы векторов

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_t, \vec{c}_{t+1}, \ldots, \vec{c}_p, \vec{s}_{t+1}, \ldots, \vec{s}_q\right\}$$

(говорят, что R порождается данной системой векторов). Если мы докажем линейную независимость этой системы, то тем самым покажем, что эта система векторов является базисом подпространства $R = L_1 + L_2$. Предположим, что выполняется векторное равенство

$$\alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \ldots + \alpha_t \vec{e}_t + \beta_{t+1} \vec{c}_{t+1} + \ldots +$$
$$+\beta_p \vec{c}_p + \gamma_{t+1} \vec{s}_{t+1} + \ldots + \gamma_q \vec{s}_q = \vec{0}.$$

Тогда, обозначив

$$\vec{u}=\alpha_1\vec{e}_1+\ldots+\alpha_t\vec{e}_t+\beta_{t+1}\vec{c}_{t+1}+\ldots+\beta_p\vec{c}_p,$$

получим

$$\vec{u}=\alpha_1\vec{e}_1+\ldots+\alpha_t\vec{e}_t+\beta_{t+1}\vec{c}_{t+1}+\ldots+\beta_p\vec{c}_p=-\gamma_{t+1}\vec{s}_{t+1}-\ldots-\gamma_q\vec{s}_q.$$

Вектор $\vec{u}\in L_1$, так как он является линейной комбинацией векторов базиса L_1, а именно векторов системы

$$\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_t,\vec{c}_{t+1},\ldots,\vec{c}_p\right\}.$$

Однако, $\vec{u}\in L_2$, так как он является линейной комбинацией части базиса

$$\left\{\vec{s}_{t+1},\ldots,\vec{s}_q\right\}\subset\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_t,\vec{s}_{t+1},\ldots,\vec{s}_q\right\}$$

подпространства L_2. Поэтому по определению пересечения подпространств $\vec{u}\in L_1\cap L_2$. Но тогда этот вектор может быть представлен в виде линейной комбинации векторов базиса подпространства

$$T=L_1\cap L_2\colon \vec{u}=\alpha_1\vec{e}_1+\alpha_2\vec{e}_2+\ldots+\alpha_t\vec{e}_t.$$

Приравняем это представление к представлению через векторы базиса L_2

$$\alpha_1\vec{e}_1+\alpha_2\vec{e}_2+\ldots+\alpha_t\vec{e}_t=-\gamma_{t+1}\vec{s}_{t+1}-\ldots-\gamma_q\vec{s}_q,$$

откуда имеем

$$\alpha_1\vec{e}_1+\alpha_2\vec{e}_2+\ldots+\alpha_t\vec{e}_t+\gamma_{t+1}\vec{s}_{t+1}+\ldots+\gamma_q\vec{s}_q=\vec{0}.$$

Но система векторов

$$\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_t,\vec{s}_{t+1},\ldots,\vec{s}_q\right\},$$

являясь базисом подпространства L_2, линейно независима. Следовательно,

$$\alpha_1=\alpha_2=\ldots=\alpha_t=\gamma_{t+1}=\ldots=\gamma_q=0.$$

Поэтому получаем

$$\alpha_1 \vec{e}_1 + \ldots + \alpha_t \vec{e}_t + \beta_{t+1} \vec{c}_{t+1} + \ldots + \beta_p \vec{c}_p = \vec{0},$$

где $\alpha_1 = \alpha_2 = \ldots = \alpha_t = 0$.

Система векторов

$$\left\{ \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_t, \vec{c}_{t+1}, \ldots, \vec{c}_p \right\}$$

образует базис векторного подпространства L_1, поэтому имеем:

$$\alpha_1 = \ldots = \alpha_t = \beta_{t+1} = \ldots = \beta_p = 0.$$

Видим, что из предположения о равенстве нуль-вектору линейной комбинации векторов

$$\left\{ \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_t, \vec{c}_{t+1}, \ldots, \vec{c}_p, \vec{s}_{t+1}, \ldots, \vec{s}_q \right\}$$

следует равенство нулю всех коэффициентов этой линейной комбинации. Поэтому по определению линейной независимости получаем, что система векторов линейно независима. Но векторы системы порождают подпространство $R = L_1 + L_2$ (R является их линейной оболочкой) и, следовательно, они составляют базис суммы

$$R = L_1 + L_2.$$

Их число определяет размерность $\dim(L_1 + L_2)$ и равно

$$\dim(L_1 + L_2) = t + (p - t) + (q - t) =$$

$$= p + q - t = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2),$$

откуда и следует, что

$$\dim(L_1 + L_2) + \dim(L_1 \cap L_2) = \dim L_1 + \dim L_2. \bullet\bullet$$

Особенно важным примером подпространств являются ***плоскости*** в аффинном пространстве, ***проходящие через начало системы координат***. В этом параграфе дадим только инвариантное определение плоскости, то есть определение, не зависящее от выбора репера в пространстве A^n. Теория плоскостей в n-мерном собственно евклидовом пространстве будет подробно рассмотрена ниже.

Определение 2.8.5. *Пусть A, B, C — некоторые произвольные точки множества* $H \subset A^n$ *(рис. 2.14). Построим вектор* $\vec{AB}$ *(аксиома* 2^0*), умножим его на некоторое (действительное) число* α *(аксиома* 8^0*) и отложим от точки C. Получим вектор* $\vec{CD} = \alpha \cdot \vec{AB}$*. Тогда, если точка D принадлежит множеству H, то это множество называется* ***плоскостью*** *в аффинном пространстве* A^n. •

Для плоскости все аксиомы аффинного пространства остаются в силе. Может лишь измениться аксиома размерности 5^0 (теперь число векторов, образующих максимальную линейно независимую систему векторов плоскости H, может быть меньше размерности пространства A^n). Если такая система содержит m векторов (размерность плоскости равна m), то плоскость является m-мерным линейным многообразием в аффинном пространстве A^n. Поэтому на плоскости H можно выбрать новый аффинный репер

$$\left\{O^*, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}.$$

Если $M = n-1$, то плоскость называется ***гиперплоскостью***.

Если плоскость проходит через начало системы координат, то согласно определению 2.8.2 она является подпространством размерности m вмещающего пространства A^n.

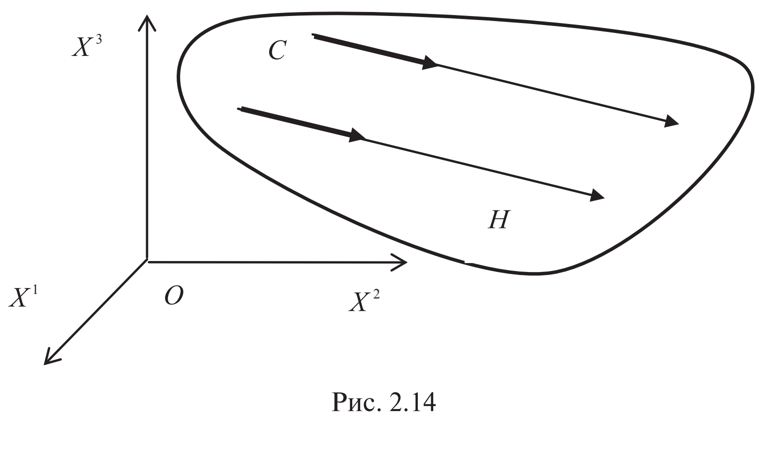


Рис. 2.14

Прямая сумма подпространств. Пусть даны векторные подпространства $L_1, L_2, \ldots, L_m$ некоторого векторного пространства, причём

$$A^n = L_1 + L_2 + \ldots + L_m.$$

Тогда всякий вектор $\vec{x} \in A^n$ по определению может быть представлен в виде

$$\vec{x} = \vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_m, \tag{2.8.2}$$

где $\vec{x}_1 \in L_1$, $\vec{x}_2 \in L_2$, …, $\vec{x}_m \in L_m$.

Определение 2.8.6. *Если каждый вектор из пространства A^n допускает единственное представление вида (2.8.2), то сумма*

$$A^n = L_1 + L_2 + \ldots + L_m$$

называется ***прямой суммой*** *и обозначается*

$$A^n = L_1 \oplus L_2 \oplus \ldots \oplus L_m. \bullet \tag{2.8.3}$$

Если векторное пространство A^n может быть разложено в прямую сумму своих подпространств (2.8.3), то в силу единственности разложения (2.8.2) систему подпространств $\{L_1, L_2, \ldots, L_m\}$ можно рассматривать как ***обобщённый базис пространства*** A^n. Таким образом, система подпространств $\{L_1, L_2, \ldots, L_m\}$, при выполнении условия (2.8.3), является (обобщённо) ***линейно независимой***.

Пусть A^n — есть n-мерное векторное пространство и $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ — некоторый его базис. Построим совокупность линейных оболочек

$$L_1 = L\left\{\vec{e}_1\right\}, L_2 = L\left\{\vec{e}_2\right\}, \ldots, L_n = L\left\{\vec{e}_n\right\}.$$

Очевидно, что

$$A^n = L_1 \oplus L_2 \oplus \ldots \oplus L_n.$$

Полученные подпространства одномерны, но можно, в общем случае, разложить векторное пространство A^n и в прямую сумму подпространств иных размерностей.

Теорема 2.8.4. *Векторное пространство A^n является прямой суммой своих подпространств $L_1, L_2, \ldots, L_m$ в том и только в том случае, если объединение базисов этих подпространств составляет базис всего пространства.*

Д о к а з а т е л ь с т в о. Пусть A^n есть прямая сумма своих подпространств $L_1, L_2, \ldots, L_m$, то есть

$$A^n = L_1 \oplus L_2 \oplus \ldots \oplus L_m,$$

а векторы

$$\left\{\vec{e}_1^{\,1}, \vec{e}_2^{\,1}, \ldots, \vec{e}_{s_1}^{\,1}\right\} \subset L_1,$$

$$\left\{\vec{e}_1^{\,2}, \vec{e}_2^{\,2}, \ldots, \vec{e}_{s_2}^{\,2}\right\} \subset L_2,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\left\{\vec{e}_1^{\,m}, \vec{e}_2^{\,m}, \ldots, \vec{e}_{s_m}^{\,m}\right\} \subset L_m$$

составляют соответственно базисы этих подпространств. Тогда для любого вектора $\vec{x}$ из пространства A^n имеет место разложение вида (2.8.2):

$$\vec{x} = \vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_m;\ \vec{x}_1 \in L_1,\ \vec{x}_2 \in L_2,\ \ldots,\ \vec{x}_m \in L_m.$$

Представляя каждый из векторов в этом разложении, в свою очередь, в виде разложения по векторам базиса соответствующего подпространства, получаем

$$\begin{gathered} \vec{x} = x_1^1 \vec{e}_1^{\,1} + x_1^2 \vec{e}_2^{\,1} + \ldots + x_1^{s_1} \vec{e}_{s_1}^{\,1} + \\ + x_2^1 \vec{e}_1^{\,2} + x_2^2 \vec{e}_2^{\,2} + \ldots + x_2^{s_2} \vec{e}_{s_2}^{\,2} + \ldots + \\ + x_m^1 \vec{e}_1^{\,m} + x_m^2 \vec{e}_2^{\,m} + \ldots + x_m^{s_m} \vec{e}_{s_m}^{\,m}. \end{gathered} \tag{2.8.4}$$

Здесь нижний индекс при координатах векторов и верхний индекс при векторах базисов принимают значения, соответствующие номеру подпространства. Таким образом, каждый вектор из пространства A^n представляется в виде линейной комбинации векторов

$$\left\{\vec{e}_1^{\,1}, \vec{e}_2^{\,1}, \ldots, \vec{e}_{s_1}^{\,1}; \vec{e}_1^{\,2}, \vec{e}_2^{\,2}, \ldots, \vec{e}_{s_2}^{\,2}; \ldots; \vec{e}_1^{\,m}, \vec{e}_2^{\,m}, \ldots, \vec{e}_{s_m}^{\,m}\right\}.$$

Если мы покажем, что эта система векторов линейно независима, то сможем утверждать, что векторы системы составляют базис пространства A^n. Потребуем, чтобы значением линейной комбинации векторов системы с числовыми коэффициентами $a_1^1, a_1^2, \ldots, a_1^{s_1}, \ldots, a_m^1, a_m^2, \ldots, a_m^{s_m}$ был нуль-вектор

$$a_1^1 \vec{e_1^1} + a_1^2 \vec{e_2^1} + \ldots + a_1^{s_1} \vec{e_{s_1}^1} + \ldots + a_m^1 \vec{e_1^m} + a_m^2 \vec{e_2^m} + \ldots + a_m^{s_m} \vec{e_{s_m}^m} = \vec{0}. \qquad (2.8.5)$$

Обозначим символами $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ линейные комбинации

$$\begin{gathered} a_1^1 \vec{e_1^1} + a_1^2 \vec{e_2^1} + \ldots + a_1^{s_1} \vec{e_{s_1}^1} = \vec{x}_1, \\ a_2^1 \vec{e_1^2} + a_2^2 \vec{e_2^2} + \ldots + a_2^{s_2} \vec{e_{s_2}^2} = \vec{x}_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \qquad (2.8.6) \\ a_m^1 \vec{e_1^m} + a_m^2 \vec{e_2^m} + \ldots + a_m^{s_m} \vec{e_{s_m}^m} = \vec{x}_m. \end{gathered}$$

Понятно, что $\vec{x}_i \in L_i$, а из разложения (2.8.5) следует, что

$$\vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_m = \vec{0}.$$

Так как все подпространства содержат нулевой вектор, то в силу единственности представления нулевого вектора имеем

$$\vec{0} + \vec{0} + \ldots + \vec{0} = \vec{0},$$

откуда получаем

$$\vec{x}_1 = \vec{x}_2 = \ldots = \vec{x}_m = \vec{0}.$$

Отсюда вытекает равенство нулю коэффициентов линейных комбинаций (2.8.6) и, следовательно, (2.8.5). Поэтому система векторов

$$\left\{ \vec{e_1^1}, \vec{e_2^1}, \ldots, \vec{e_{s_1}^1}; \vec{e_1^2}, \vec{e_2^2}, \ldots, \vec{e_{s_2}^2}; \ldots; \vec{e_1^m}, \vec{e_2^m}, \ldots, \vec{e_{s_m}^m} \right\}$$

линейно независима.

Предположим теперь, что система векторов

$$\left\{ \vec{e_1^1}, \vec{e_2^1}, \ldots, \vec{e_{s_1}^1}; \vec{e_1^2}, \vec{e_2^2}, \ldots, \vec{e_{s_2}^2}; \ldots; \vec{e_1^m}, \vec{e_2^m}, \ldots, \vec{e_{s_m}^m} \right\}$$

является базисом пространства A^n. Тогда для любого вектора $\vec{x}$ из пространства A^n имеет место единственное разложение вида (2.8.4). Обозначая

$$a_1^1 \vec{e_1^1} + a_1^2 \vec{e_2^1} + \ldots + a_1^{s_1} \vec{e_{s_1}^1} = \vec{x}_1,$$

$$a_2^1 \vec{e}_1^{\,2} + a_2^2 \vec{e}_2^{\,2} + \ldots + a_2^{s_1} \vec{e}_{s_1}^{\,2} = \vec{x}_2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \tag{2.8.7}$$

$$a_m^1 \vec{e}_1^{\,m} + a_m^2 \vec{e}_2^{\,m} + \ldots + a_m^{s_m} \vec{e}_{s_m}^{\,m} = \vec{x}_m,$$

получаем, что для каждого вектора $\vec{x}$ имеет место, по крайней мере, одно представление вида (2.8.2). Каждый вектор из (2.8.7) есть линейная комбинация векторов базиса соответствующего подпространства. Из единственности разложения (2.8.4) следует и единственность представления (2.8.2) для вектора $\vec{x}$. Теорема доказана. ••

Пример 2.8.1. Нетрудно видеть, что пространство R^3 трёхмерных геометрических векторов может быть представлено как прямая сумма, например, двумерного подпространства — координатной плоскости $X^1 O X^2$, и одномерного подпространства — оси $O X^3$. ⊗

Изоморфизм векторных пространств. В каждом векторном пространстве над его векторами (элементами) можно производить операции. В принципе, от природы векторов (силы, скорости, ускорения, векторов различных физических полей) можно абстрагироваться и рассматривать векторы только с точки зрения производимых над ними операций сложения и умножения на числа. Два векторных пространства, устроенные по отношению к операциям над векторами одинаково, можно считать обладающими одинаковыми свойствами или ***изоморфными***. Сформулируем понятие изоморфизма векторных пространств.

Определение 2.8.7. *Векторные пространства A_1 и A_2, заданные над одним и тем же полем P, называются* ***изоморфными****, если существует такое взаимно однозначное отображение $\hat{f} : A_1 \to A_2$, при котором сумме любых двух векторов пространства A_1 ставится в соответствие сумма соответствующих векторов пространства A_2, а произведению любого вектора пространства A_1 на некоторое число из поля P ставится в соответствие произведение соответствующего вектора пространства A_2 на то же самое число.*•

Таким образом, имеем: $\left(\forall \vec{x}, \vec{y} \in A_1\right)$ и $\left(\forall \alpha \in P\right)$

$$\hat{f}\left(\vec{x}+\vec{y}\right)=\hat{f}\left(\vec{x}\right)+\hat{f}\left(\vec{y}\right)\in A_2,\ \hat{f}\left(\alpha\vec{x}\right)=\alpha\cdot\hat{f}\left(\vec{x}\right)\in A_2.$$

Взаимная однозначность указанного отображения означает, что если $\vec{x},\vec{y}\in A_1$, причём $\vec{x}\neq\vec{y}$, то и $\hat{f}\left(\vec{x}\right)\neq\hat{f}\left(\vec{y}\right)$.

В изоморфных пространствах много общего. В частности, ***в изоморфных пространствах нуль-вектору соответствует нуль-вектор***. Заметим, например, что если обозначить $\vec{0}\in A_1$ и $\vec{0}'\in A_2$, то

$$\hat{f}\left(\vec{0}\right)=\hat{f}\left(0\cdot\vec{x}\right)=0\cdot\hat{f}\left(\vec{x}\right)=0\cdot\vec{x}'=\vec{0}'\in A_2.$$

Важнейшее свойство изоморфных пространств устанавливается следующей теоремой.

Теорема 2.8.5. *Пусть A_1 и A_2 — изоморфные векторные пространства. Тогда каждой линейно независимой системе векторов векторного пространства A_1 соответствует снова линейно независимая система векторов векторного пространства A_2.*

Доказательство. Пусть векторы $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n\subset A_1$ образуют линейно независимую систему. Рассмотрим линейную комбинацию

$$\alpha_1\cdot\hat{f}\left(\vec{x}_1\right)+\alpha_2\cdot\hat{f}\left(\vec{x}_2\right)+\alpha_n\cdot\hat{f}\left(\vec{x}_n\right)$$

и потребуем, чтобы эта линейная комбинация имела своим значением нулевой вектор $\vec{0}'$. В силу определения изоморфизма имеем:

$$\alpha_1\cdot\hat{f}\left(\vec{x}_1\right)+\alpha_2\cdot\hat{f}\left(\vec{x}_2\right)+\ldots+\alpha_n\cdot\hat{f}\left(\vec{x}_n\right)=\vec{0}'\Rightarrow$$
$$\hat{f}\left(\alpha_1\cdot\vec{x}_1+\alpha_2\cdot\vec{x}_2+\alpha_n\cdot\vec{x}_n\right)=\vec{0}'.$$

Далее, по сделанному выше замечанию получаем, что $\vec{0}'=\hat{f}\left(\vec{0}\right)$, где $\vec{0}\in A_1$ и $\vec{0}'\in A_2$. Поэтому

$$\alpha_1\cdot\vec{x}_1+\alpha_2\cdot\vec{x}_2+\alpha_n\cdot\vec{x}_n=\vec{0}.$$

Так как система векторов

$$\left\{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_n\right\}$$

по условию линейно независима, то все коэффициенты этой линейной комбинации должны быть равными нулю. Но тогда и система векторов

$$\left\{\hat{f}\left(\vec{x}_1\right), \hat{f}\left(\vec{x}_2\right), \ldots, \hat{f}\left(\vec{x}_n\right)\right\}$$

является линейно независимой. ••

Следствие из теоремы 2.8.5. *Если любые два конечномерных векторных пространства изоморфны, то они имеют одинаковую размерность.*

Справедлива теорема, имеющая «*обратный*» характер.

Теорема 2.8.6. *Любые два конечномерных пространства, имеющие одинаковые размерности и заданные над одним и тем же полем, изоморфны.*

Д о к а з а т е л ь с т в о. Пусть A_1 и A_2 — два векторных пространства размерности n, заданные над одним полем P. Предположим, что

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset A_1, \left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset A_2$$

— соответствующие друг другу при отображении $\hat{f}$ базисы этих пространств, то есть

$$\vec{g}_i = \hat{f}\left(\vec{e}_i\right)(i = 1, 2, \ldots, n).$$

Таким образом, каждому вектору

$$\vec{x} = x^1 \vec{e}_1 + \ldots + x^n \vec{e}_n \in A_1$$

ставится в соответствие вектор

$$\hat{f}\left(\vec{x}\right) = \hat{f}\left(x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^n \vec{e}_n\right) =$$

$$= x^1 \hat{f}\left(\vec{e}_1\right) + x^2 \hat{f}\left(\vec{e}_2\right) + \ldots + x^n \hat{f}\left(\vec{e}_n\right) = x^1 \vec{g}_1 + x^2 \vec{g}_2 + \ldots + x^n \vec{g}_n$$

из A_2. В силу единственности разложения по базису построенное соответствие будет взаимно однозначным.

Возьмём два любых вектора $\vec{x}$ и $\vec{y}$ из пространства A_1 и произвольное число $\alpha \in P$. Пусть

$$\vec{x} = x^1 \vec{e}_1 + \ldots + x^n \vec{e}, \vec{y} = y^1 \vec{e}_1 + \ldots + y^n \vec{e}_n.$$

Получаем:

$$\hat{f}\left(\vec{x}+\vec{y}\right)=\hat{f}\left(\left(x^1+y^1\right)\vec{e}_1+\left(x^2+y^2\right)\vec{e}_2+\ldots+\left(x^n+y^n\right)\vec{e}_n\right)=$$

$$=\left(x^1+y^1\right)\vec{g}_1+\ldots+\left(x^n+y^n\right)\vec{g}_n=$$

$$=\left(x^1\vec{g}_1+\ldots+x^n\vec{g}_n\right)+\left(y^1\vec{g}_1+\ldots+y^n\vec{g}_n\right)=\hat{f}\left(\vec{x}\right)+\hat{f}\left(\vec{y}\right);$$

$$\hat{f}\left(\alpha\cdot\vec{x}\right)=\hat{f}\left(\alpha\cdot x^1\vec{e}_1+\ldots+\alpha\cdot x^n\vec{e}_n\right)=\alpha\cdot x^1\vec{g}_1+\ldots+\alpha\cdot x^n\vec{g}_n=$$

$$=\alpha\cdot\left(x^1\vec{g}_1+\ldots+x^n\vec{g}_n\right)=\alpha\cdot\hat{f}\left(\vec{x}\right).$$

Две последние цепочки выкладок доказывают теорему. ••

Из доказанной теоремы следует, что с точки зрения всех аксиом векторного пространства и следствий из них любые два заданные над одним полем пространства одинаковой размерности неразличимы. Поэтому формально можно построить одно, в определённом смысле слова, простейшее пространство над заданным полем и изучать на его примере свойства, общие для всех пространств данной размерности, заданных над тем же полем. Таковым является, например, пространство R^n.

Ортогональная сумма подпространств евклидова пространства. Понятие подпространства остаётся неизменным и для евклидова пространства с учётом того, что в подпространстве, как и во всём пространстве, существует скалярное произведение.

Определение 2.8.8. *Два множества X и Y векторов евклидова пространства E^n называются* ***ортогональными*** *и пишут $X \perp Y$, если каждый вектор из множества X ортогонален каждому вектору из множества Y.* •

Справедливо следующее важное утверждение.

Лемма 2.8.3. *Вектор $\vec{x}\in E^n$ ортогонален к некоторому подпространству $L^m \subset E^n$ в том и только в том случае, если он ортогонален ко всем векторам какого-либо базиса этого подпространства.*

Доказательство. Пусть $\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_m\right\}$ — некоторый базис в подпространстве L^m. Если $\vec{x}\perp L^m$, то он ортогонален и ко всем векторам подпространства, в частности, ко всем векторам базиса.

Пусть теперь

$$(\forall i = 1, 2, \ldots, m)\left(\vec{x}, \vec{a}_i\right) = 0.$$

Возьмём некоторый вектор $\vec{z} \in L^m$. Для этого вектора имеем

$$\vec{z} = z^1 \vec{a}_1 + z^2 \vec{a}_2 + \ldots + z^m \vec{a}_m.$$

Находя скалярное произведение с аргументами $\vec{x}$ и $\vec{z}$, получаем:

$$\left(\vec{x}, \vec{z}\right) = z^1\left(\vec{x}, \vec{a}_1\right) + z^2\left(\vec{x}, \vec{a}_2\right) + \ldots + z^m\left(\vec{x}, \vec{a}_m\right) = 0.$$

Так как вектор $\vec{z}$ произволен, то $\vec{x} \perp L^m$. ••

Определение 2.8.9. *Сумма некоторой системы подпространств $L_1, L_2, \ldots, L_m$ евклидова пространства E^n называется **ортогональной**, если*

$$(\forall i, j = 1, 2, \ldots, m)\ L_i \perp L_j. \bullet$$

Лемма 2.8.4. *Ортогональная сумма ненулевых подпространств*

$$L_1, L_2, \ldots, L_m \subset E^n$$

является прямой суммой.

Доказательство. Выберем в каждом подпространстве ортонормированный базис, учитывая, что подпространства попарно ортогональны:

$$\begin{gathered} \left\{\vec{a}_1^1, \vec{a}_2^1, \ldots, \vec{a}_{k_1}^1\right\} \subset L_1, \\ \left\{\vec{a}_1^2, \vec{a}_2^2, \ldots, \vec{a}_{k_2}^2\right\} \subset L_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \left\{\vec{a}_1^m, \vec{a}_2^m, \ldots, \vec{a}_{k_m}^m\right\} \subset L_m. \end{gathered} \tag{2.8.8}$$

По определению суммы подпространств

$$\left(\forall \vec{x} \in L_1 + L_2 + \ldots + L_m\right) \vec{x} = \vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_m,$$

где $(\forall i = 1, 2, \ldots, m)\ \vec{x}_i \in L_i$. Раскладывая каждый вектор $\vec{x}_i$ по векторам базиса соответствующего подпространства, получаем:

$$\vec{x} = \vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_m =$$

$$= x_1^1 \vec{a}_1^1 + x_1^2 \vec{a}_2^1 + \ldots + x_1^{k_1} \vec{a}_{k_1}^1 +$$

$$+ x_2^1 \vec{a}_1^2 + x_2^2 \vec{a}_2^2 + \ldots + x_2^{k_2} \vec{a}_{k_2}^2 +$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots +$$

$$+ x_m^1 \vec{a}_1^m + x_m^2 \vec{a}_2^m + \ldots + x_m^{k_m} \vec{a}_{k_m}^m .$$

Теперь видно, что каждый вектор из суммы $L_1 + L_2 + \ldots + L_m$ линейно выражается через векторы (2.8.8). Но система векторов

$$\left\{ \vec{a}_1^1, \vec{a}_2^1, \ldots, \vec{a}_{k_1}^1, \vec{a}_1^2, \vec{a}_2^2, \ldots, \vec{a}_{k_2}^2, \ldots, \vec{a}_1^m, \vec{a}_2^m, \ldots, \vec{a}_{k_m}^m \right\}$$

линейно независима, так как система ортогональная и состоит из ненулевых векторов. Следовательно, эта система составляет базис в $L_1 + L_2 + \ldots + L_m$. Теперь доказательство леммы следует из теоремы 2.8.4. ••

Определение 2.8.10. *Если $L^m \subset E^n$ — непустое подмножество векторов евклидова пространства E^n, то совокупность всех векторов, ортогональных подмножеству L^m, называется* ***ортогональным дополнением подмножества*** *L^m и обозначается $L^{m\perp}$.* •

Теорема 2.8.7. *Евклидово пространство E является ортогональной суммой любого своего подпространства L и его ортогонального дополнения $L^\perp$, то есть $E = L \oplus L^\perp$.*

Д о к а з а т е л ь с т в о. Пусть $\dim L = p$, а $\dim L^\perp = q$. В подпространстве L выберем некоторый ортонормированный базис

$$\left\{ \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p \right\} \subset L,$$

а в подпространстве $L^\perp$ — некоторый ортонормированный базис

$$\left\{ \vec{\mathrm{v}}_1, \vec{\mathrm{v}}_2, \ldots, \vec{\mathrm{v}}_q \right\} \subset L^\perp.$$

Система векторов

$$\left\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{\mathrm{v}}_1, \vec{\mathrm{v}}_2, \ldots, \vec{\mathrm{v}}_q\right\}$$

ортонормированная и, следовательно, линейно независимая.

Предположим, что она не является базисом пространства E. В этом случае её можно дополнить до базиса, применив процедуру ортогонализации Шмидта.

Пусть $\vec{w}$ — один из дополняющих векторов. Он ортогонален векторам $\left\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\right\}$, а поэтому $\vec{w} \perp L$. Следовательно, $\vec{w} \in L^{\perp}$. Но вектор $\vec{w}$ ортогонален также и векторам $\left\{\vec{\mathrm{v}}_1, \vec{\mathrm{v}}_2, \ldots, \vec{\mathrm{v}}_q\right\}$ и поэтому $\vec{w} \perp L^{\perp}$. Итак, вектор $\vec{w}$ одновременно принадлежит $L^{\perp}$ и ортогонален $L^{\perp}$. Поэтому в силу аксиомы невырожденности $\vec{w} = \vec{0}$. Следовательно,

$$\left\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p, \vec{\mathrm{v}}_1, \vec{\mathrm{v}}_2, \ldots, \vec{\mathrm{v}}_q\right\}$$

— базис в пространстве E и $p+q=n$.••

Глава 3. Линейные операторы в абстрактных векторных пространствах

3.1. Линейные операторы и матрицы. Алгебра линейных операторов

Линейные операторы. В абстрактных векторных пространствах, в частности в пространствах R^3 и R^n, вводится понятие линейного оператора.

Оператор — это отображение $\hat{A}: X \to Y$, ставящее в соответствие по определённому правилу каждому вектору $\vec{x} \in L \subset X$ некоторый, в общем случае не единственный, вектор $\vec{y} \in Q \subset Y$. Здесь X и Y — векторные пространства с различными, вообще говоря, размерностями. Множество $L \subset X$ называется ***множеством определения***, а множество $Q \subset Y$ — ***множеством значений оператора*** $\hat{A}$. Действие оператора записывается в префиксной форме

$$\vec{y} = \hat{A}\left(\vec{x}\right) = \hat{A}\vec{x}.$$

В этой записи вектор $\vec{y}$ называется ***образом вектора*** $\vec{x}$, а вектор $\vec{x}$ — ***прообразом вектора*** $\vec{y}$ при действии оператора $\hat{A}$.

Дальше мы будем изучать только операторы специального вида — ***линейные операторы***.

Определение 3.1.1. *Пусть X и Y — два векторных пространства и P — некоторое числовое поле. Тогда отображение $\hat{A}: X \to Y$ называется **линейным оператором из X в Y**, если:* $\left(\forall \vec{x}_1, \vec{x}_2 \in X\right)$ *и* $\left(\forall \alpha_1, \alpha_2 \in P\right)$.

$$\hat{A}\left(\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2\right) \stackrel{def}{=} \alpha_1 \cdot \hat{A}\vec{x}_1 + \alpha_2 \cdot \hat{A}\vec{x}_2. \bullet \qquad (3.1.1)$$

В определении 3.1.1, как и в общем случае произвольного оператора, действие линейного оператора $\hat{A}: X \to Y$ на векторы пространства X записано в префиксной форме:

$$\vec{y} = \hat{A}\vec{x}. \tag{3.1.2}$$

Как видно из определения, линейный оператор переводит линейную комбинацию прообразов в линейную же комбинацию образов. Следовательно, линейный оператор обладает двумя свойствами — ***аддитивностью*** и ***однородностью***, а именно:

1) $\left(\forall \vec{x}_1, \vec{x}_2 \in X\right) \hat{A}\left(\vec{x}_1 + \vec{x}_2\right) = \hat{A}\vec{x}_1 + \hat{A}\vec{x}_2;$

2) $\left(\forall \vec{x} \in X\right) \wedge \left(\forall \alpha \in P\right) \hat{A}\left(\alpha \cdot \vec{x}\right) = \alpha \cdot \hat{A}\left(\vec{x}\right).$

Аналогично общему случаю, для случая линейного оператора вектор $\vec{x} \in X$ в записи (3.1.2) называется ***прообразом*** вектора $\vec{y}$, а вектор $\vec{y} \in Y$ — ***образом*** вектора $\vec{x}$. Множество $\hat{A}(X) = \left\{\vec{y} \in Y : \vec{y} = \hat{A}\vec{x}, \vec{x} \in L \subset X\right\}$ называется ***множеством значений*** $\hat{A}(L)$ (или $\hat{A}(X)$), а множество $L \subset X$ всех векторов $\vec{x} \in X$, для которых выполняется равенство (3.1.2), — ***множеством определения*** линейного оператора $\hat{A}$.

Частным, но весьма важным случаем линейного оператора является оператор, действующий в одном пространстве, то есть оператор $\hat{A}: X \to X$. Такой оператор является ***преобразованием***. У операторов $\hat{A}: X \to X$ имеется ряд свойств, которых нет у операторов $\hat{A}: X \to Y$.

Очень важными для теории линейных операторов являются понятия ***ранга*** и ***ядра***.

Определение 3.1.2. *Пусть* $\hat{A}: X \to \hat{A}(X)$. *Размерность*

$$\dim \hat{A}(X) \equiv r\left(\hat{A}\right)$$

множества значений $\hat{A}(X)$ *называется* ***рангом линейного оператора*** $\hat{A}$. •

Определение 3.1.3. *Пусть* $\hat{A}: X \to \hat{A}(X)$. *Множество*

$$K\left(\hat{A}\right) = \left\{\vec{x} \in X : \hat{A}\vec{x} = \vec{0}\right\}$$

называется ***ядром оператора*** $\hat{A}$. *Размерность*

$$\dim K\left(\hat{A}\right) \equiv d\left(\hat{A}\right)$$

ядра называется ***дефектом линейного оператора*** $\hat{A}$. *Если* $K\left(\hat{A}\right)$ *содержит только* $\vec{0}$, *то оператор* $\hat{A}: X \to \hat{A}(X)$ *называется* ***невырожденным***.•

Теорема 3.1.1. *Линейный оператор* $\hat{A}: X \to X$ *взаимно однозначен (биективен) в том и только в том случае, если он невырожденный.*

Доказательство. Пусть оператор биективный. Из биективности оператора следует, что

$$\vec{x}_1 = \vec{x}_2 \Rightarrow \hat{A}\vec{x}_1 = \hat{A}\vec{x}_2.$$

Тогда в силу линейности оператора получаем

$$\hat{A}\left(\vec{x}_1 - \vec{x}_2\right) = \vec{0}.$$

Последнее и означает, что ядро оператора $K\left(\hat{A}\right) = \vec{0}$.

Обратно, пусть оператор невырожденный, то есть $K\left(\hat{A}\right) = \vec{0}$. Тогда

$$\hat{A}\vec{x}_1 = \hat{A}\vec{x}_2 \Rightarrow \hat{A}\left(\vec{x}_1 - \vec{x}_2\right) = \vec{0} \Rightarrow \vec{x}_1 = \vec{x}_2. \bullet\bullet$$

Следствие из теоремы 3.1.1. *Если* $\hat{A}: X \to X$ *— невырожденный линейный оператор, то* $\left(\forall \vec{y} \in X\right)$ *полный прообраз* $\hat{A}^{-1}\vec{y}$ *состоит из единственного вектора.*

Доказательство. Пусть для некоторого $\vec{y} \in X$ прообраз не единственен, то есть $\left(\exists \vec{x}_1, \vec{x}_2 \in X\right): \hat{A}\vec{x}_1 = \vec{y}$ и $\hat{A}\vec{x}_2 = \vec{y}$, но тогда

$$\hat{A}\left(\vec{x}_1 - \vec{x}_2\right) = \vec{0}.$$

Так как оператор $\hat{A}$ невырожденный, то $\vec{x}_1 - \vec{x}_2 = \vec{0} \Rightarrow \vec{x}_1 = \vec{x}_2$. ●●

Пусть $\hat{A}: X \to X$ — невырожденный оператор. В силу единственности прообраза вектора $\vec{y} = \hat{A}\vec{x}$ любому вектору $\vec{y} \in X$ можно поставить в соответствие единственный вектор $\vec{x} \in X$. Такое соответствие определяет ***обратный оператор***, *обозначаемый* $\hat{A}^{-1}: X \to X$.

Полагая $\vec{x} = \hat{A}^{-1}\vec{y}$, действуя на обе части равенства $\vec{y} = \hat{A}\vec{x}$ оператором $\hat{A}^{-1}$, а на обе части равенства $\vec{x} = \hat{A}^{-1}\vec{y}$ оператором $\hat{A}$, получаем, что

$$\vec{x} = \hat{A}^{-1}\hat{A}\vec{x},\ \vec{y} = \hat{A}\hat{A}^{-1}\vec{y}.$$

Из этих формул следует, что операторы $\hat{A}^{-1}\hat{A}$ и $\hat{A}\hat{A}^{-1}$ не влияют на соответствующий прообраз, то есть

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I},$$

где $\hat{I}$ — ***единичный*** (***тождественный***) ***оператор***, обладающий свойством

$$\left(\forall \vec{x} \in X\right) \hat{I}\vec{x} = \vec{x}.$$

Теорема 3.1.2. *Если $\hat{A}: X \to X$ — невырожденный линейный оператор, то обратный оператор $\hat{A}^{-1}$ линейный.*

Д о к а з а т е л ь с т в о. Пусть $\vec{x}_1, \vec{x}_2 \in X$ и $\alpha_1, \alpha_2 \in P$. Рассмотрим вектор

$$\vec{z} = \hat{A}^{-1}\left(\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2\right) - \alpha_1 \cdot \hat{A}^{-1}\vec{x}_1 - \alpha_2 \cdot \hat{A}^{-1}\vec{x}_2.$$

Применяя к этому вектору оператор $\hat{A}$, получаем, что образ $\hat{A}\vec{z} = \vec{0}$. В силу невырожденности $\hat{A}$ вектор $\vec{z} = \vec{0}$, то есть

$$\hat{A}^{-1}\left(\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2\right) - \alpha_1 \cdot \hat{A}^{-1}\vec{x}_1 - \alpha_2 \cdot \hat{A}^{-1}\vec{x}_2 = \vec{0},$$

откуда следует линейность обратного оператора:

$$\hat{A}^{-1}\left(\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2\right) = \alpha_1 \cdot \hat{A}^{-1}\vec{x}_1 + \alpha_2 \cdot \hat{A}^{-1}\vec{x}_2. ●●$$

Теорема 3.1.3. *Если оператор* $\hat{A}$ *невырожденный, то оператор* $\hat{A}^{-1}$ *также невырожденный.*

Доказательство. Пусть $\hat{A}^{-1}\vec{y}=\vec{0}$. Действуя на обе части этого равенства оператором $\hat{A}$, получаем $\hat{I}\,\vec{y}=\vec{0}$, то есть $\vec{y}=\vec{0}$. ••

Справедлива следующая важная теорема.

Теорема 3.1.4. *Пусть* $\hat{A}:X\to X$ *— некоторый линейный оператор. Если* $\dim X=m$*, то ранг и дефект оператора связаны соотношением*

$$r\left(\hat{A}\right)+d\left(\hat{A}\right)=m.$$

Доказательство. Пусть

$$d\left(\hat{A}\right)\equiv\dim K\left(\hat{A}\right)=k<m$$

и $\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_k\right\}\subset K\left(\hat{A}\right)$ — базис ядра оператора. Дополним его до базиса всего пространства

$$\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_k,\vec{e}_{k+1},\ldots,\vec{e}_m\right\}\subset X.$$

Тогда каждый вектор $\vec{x}\in X$ можно представить в виде

$$\vec{x}=x^1\vec{e}_1+x^2\vec{e}_2+\ldots+x^k\vec{e}_k+x^{k+1}\vec{e}_{k+1}+\ldots+x^m\vec{e}_m=\vec{x}_1+\vec{x}_2,$$

где вектор

$$\vec{x}_1=x^1\vec{e}_1+\ldots+x^k\vec{e}_k\in K\left(\hat{A}\right),$$

а вектор

$$\vec{x}_2=x^{k+1}\vec{e}_{k+1}+\ldots+x^m\vec{e}_m\notin K\left(\hat{A}\right),$$

причём $k<m\Rightarrow\vec{x}_2\neq\vec{0}$.

Так как $\hat{A}\vec{x}_1=\vec{0}$, то для каждого вектора $\vec{y}\in\hat{A}(X)$ справедливо представление

$$\vec{y}=\hat{A}\vec{x}=\hat{A}\left(\vec{x}_1+\vec{x}_2\right)=\hat{A}\vec{x}_2=x^{k+1}\hat{A}\vec{e}_{k+1}+x^{k+2}\hat{A}\vec{e}_{k+2}+\ldots+x^m\hat{A}\vec{e}_m.$$

Система образов дополнительных до базиса векторов

$$\left\{\hat{A}\vec{e}_{k+1}, \hat{A}\vec{e}_{k+2}, \ldots, \hat{A}\vec{e}_m\right\}$$

линейно независима. Действительно, предполагая линейную зависимость этой системы образов, приходим к равенству

$$\alpha_{k+1}\hat{A}\vec{e}_{k+1} + \alpha_{k+2}\hat{A}\vec{e}_{k+2} + \ldots + \alpha_m\hat{A}\vec{e}_m = \vec{0},$$

где не все коэффициенты равны нулю одновременно. Из последнего равенства в силу линейности оператора имеем

$$\hat{A}\left(\alpha_{k+1}\vec{e}_{k+1} + \alpha_{k+2}\vec{e}_{k+2} + \ldots + \alpha_m\vec{e}_m\right) = \vec{0},$$

то есть

$$\alpha_{k+1}\vec{e}_{k+1} + \alpha_{k+2}\vec{e}_{k+2} + \ldots + \alpha_m\vec{e}_m \in K\left(\hat{A}\right),$$

что приводит к противоречию. Следовательно, $\left\{\hat{A}\vec{e}_{k+1}, \hat{A}\vec{e}_{k+2}, \ldots, \hat{A}\vec{e}_m\right\}$ — базис в подпространстве $\hat{A}(X)$ и

$$r\left(\hat{A}\right) \equiv \dim\hat{A}(X) = m - k,$$

откуда имеем

$$r\left(\hat{A}\right) + d\left(\hat{A}\right) = m. \bullet\bullet$$

Конструкция линейного оператора. Установим ***конструкцию линейного оператора*** в общем случае оператора $\hat{A}: X^m \to Y^n$.

Пусть фиксированы базисы

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\right\} \subset X^m, \left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset Y^n.$$

Тогда векторы $\vec{x} \in X^m$ и $\vec{y} = \hat{A}\vec{x} \in Y^n$ можно представить в виде разложений

$$\vec{x} = x^1\vec{e}_1 + x^2\vec{e}_2 + \ldots + x^m\vec{e}_m = \sum_{j=1}^{m} x^j\vec{e}_j;$$

$$\vec{y} = y^1\vec{g}_1 + y^2\vec{g}_2 + \ldots + y^n\vec{g}_n = \sum_{i=1}^{n} y^i\vec{g}_i. \qquad (3.1.3)$$

Используя линейность оператора $\hat{A}$, получаем

$$\vec{y} = \hat{A}\vec{x} = \hat{A}\left(\sum_{j=1}^{m} x^j \vec{e}_j\right) = \sum_{j=1}^{m} x^j \hat{A}\vec{e}_j = \sum_{j=1}^{m} x^j \sum_{i=1}^{n} a_j^i \vec{g}_i =$$

$$= \sum_{i=1}^{n}\left(\sum_{j=1}^{m} x^j a_j^i\right)\vec{g}_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_j^i x^j\right)\vec{g}_i, \tag{3.1.4}$$

где векторы $\hat{A}\vec{e}_j$ разложены по базису пространства Y^n, то есть

$$\hat{A}\vec{e}_j = \sum_{i=1}^{n} a_j^i \vec{g}_i .$$

Сравнивая равенство (3.1.4) со вторым равенством (3.1.3), получаем

$$\left(\forall i = 1, 2, \ldots, n\right)\ y^i = \sum_{j=1}^{m} a_j^i x^j , \tag{3.1.5}$$

или в развёрнутом виде

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_m^1 x^m = y^1, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_m^2 x^m = y^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n x^1 + a_2^n x^2 + \ldots + a_m^n x^m = y^n. \end{cases} \tag{3.1.6}$$

Из соотношений (3.1.6) видно, что линейный оператор $\hat{A} : X^m \to Y^n$ полностью определяется таблицей

$$A_{(n\times m)} = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_m^n \end{pmatrix}. \tag{3.1.7}$$

Таблица (3.1.7) называется ***матрицей оператора*** $\hat{A}$. Интерпретируя соотношение (3.1.6) как равенство вектор-столбцов, перепишем его в следующем виде:

$$\begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_m^n \end{pmatrix}\begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^m \end{pmatrix} = \begin{pmatrix} y^1 \\ y^2 \\ \ldots \\ y^n \end{pmatrix}. \tag{3.1.8}$$

Соотношение (3.1.8) определяет действие ***умножения вектор-столбца на матрицу***.

Итак, доказана следующая теорема.

Теорема 3.1.5. *Каждый линейный оператор* $\hat{A}: X^m \to Y^n$ *при фиксированных базисах в пространствах* X^m *и* Y^n *порождает систему линейных соотношений вида (3.1.6) или, что то же самое соотношение (3.1.8) с матрицей (3.1.7), имеющей n строк и m столбцов, связывающих координаты образа* $\vec{y} \in Y^n$ *и прообраза* $\vec{x} \in X^m$.

Кроме обозначения (3.1.7) для матриц иногда используют следующие обозначения:

$$\begin{bmatrix} a_1^1 & a_2^1 & \dots & a_m^1 \\ a_1^2 & a_2^2 & \dots & a_m^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_m^n \end{bmatrix}, \begin{Vmatrix} a_1^1 & a_2^1 & \dots & a_m^1 \\ a_1^2 & a_2^2 & \dots & a_m^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_m^n \end{Vmatrix}.$$

Мы, однако, будем использовать исключительно обозначения вида (3.1.7).

В обозначении a_j^i ***элемента матрицы*** верхний индекс нумерует строки матрицы, а нижний индекс — столбцы. Причём число столбцов матрицы (3.1.7) равно размерности пространства прообразов X^m, а число строк равно размерности пространства образов X^n. Если матрица имеет n строк и m столбцов, то о ней говорят как о матрице размеров $n \times m$.

В случае, когда число строк совпадает с числом столбцов, матрица называется ***квадратной матрицей размером*** $n \times n$ и соответствует оператору $\hat{A}: X^n \to X^n$, действующему в одном и том же пространстве. Для квадратной матрицы выделяется ***главная диагональ***, которая образована элементами $a_j^j \left(j = 1, 2, \dots, n \right)$ с одинаковыми номерами. Главной диагонали соответствует побочная диагональ.

Таким образом, чтобы по координатам прообраза найти координаты образа, достаточно вычислить левые части выражений (3.1.6). Чтобы определить координаты прообраза по координатам образа, нужно решить СЛАУ (3.1.6) относительно неизвестных $x^1, x^2, \dots, x^m$.

Сформулируем ***правило нахождения матрицы*** линейного оператора $\hat{A}: X^m \to Y^n$, вытекающее из его конструкции:

элементы столбцов матрицы линейного оператора $\hat{A}: X^m \to Y^n$ ***являются координатами в разложении образов соответствующих базисных векторов пространства*** X^m ***по векторам базиса пространства образов*** Y^n.

При фиксированных в пространствах прообразов и образов базисах это правило позволяет найти и записать по столбцам матрицу линейного оператора, заданного некоторым высказыванием, действуя оператором на каждый базисный вектор пространства прообразов. Матрицу линейного оператора можно найти и путём действия оператора на произвольный вектор пространства прообразов. Однако этот способ несколько более громоздкий.

Пример 3.1.1. 1) Поставим каждому вектору $\vec{x} \in X^3$ в соответствие нулевой вектор того же пространства X^3.

Полученный оператор $\hat{O}: X^3 \to \left\{\vec{0}\right\}$ называется ***нулевым оператором***. Очевидно, матрица нулевого оператора

$$O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

2) Поставив в соответствие каждому вектору $\vec{x} \in X^3$ тот же самый вектор, получим ***единичный*** (***тождественный***) ***оператор*** $\hat{I}: X^3 \to X^3$, действие которого определяется правилом:

$$\hat{A}\vec{x} = 1 \cdot \vec{x} \equiv \vec{x}.$$

Используем сформулированное выше правило нахождения матрицы оператора. Для этого подействуем оператором $\hat{I}$ на базисные векторы $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ пространства X^3 и разложим образы базисных векторов по векторам того же базиса:

$$\hat{A}\vec{e}_1 = 1 \cdot \vec{e}_1 = 1 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3,$$

$$\hat{A}\vec{e}_2 = 1 \cdot \vec{e}_2 = 0 \cdot \vec{e}_1 + 1 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3,$$

$$\hat{A}\vec{e}_3 = \alpha \cdot \vec{e}_3 = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 1 \cdot \vec{e}_3.$$

Помещая коэффициенты в разложении образов базисных векторов — их координаты — на соответствующие по номерам места в столбцах матрицы оператора, получаем:

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3) Пусть $\hat{A}: X^3 \to X^3$. Определим оператор $\hat{B}: X^3 \to X^3$ высказыванием

$$\hat{B}\vec{x} \overset{def}{=} -\hat{A}\vec{x}.$$

Полученный оператор называется ***противоположным*** оператору $\hat{A}$**.**

Используем сформулированное выше правило нахождения матрицы оператора. Для этого подействуем оператором на базисные векторы $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ пространства X^3 и разложим образы базисных векторов по векторам того же базиса:

$$\hat{B}\vec{e}_1 = -\hat{A}\vec{e}_1 = -a_1^1 \cdot \vec{e}_1 - a_1^2 \cdot \vec{e}_2 - a_1^3 \cdot \vec{e}_3,$$

$$\hat{B}\vec{e}_2 = -\hat{A}\vec{e}_2 = -a_2^1 \cdot \vec{e}_1 - a_2^2 \cdot \vec{e}_2 - a_2^3 \cdot \vec{e}_3,$$

$$\hat{B}\vec{e}_3 = -\hat{A}\vec{e}_3 = -a_3^1 \cdot \vec{e}_1 - a_3^2 \cdot \vec{e}_2 - a_3^3 \cdot \vec{e}_3.$$

Помещая коэффициенты в разложении образов базисных векторов на соответствующие места в столбцах матрицы оператора, получаем:

$$-A = \begin{pmatrix} -a_1^1 & -a_2^1 & -a_3^1 \\ -a_1^2 & -a_2^2 & -a_3^2 \\ -a_1^3 & -a_2^3 & -a_3^3 \end{pmatrix}.$$

4) Пусть $\alpha \in R^1$. Определим оператор $\hat{A}: X^3 \to X^3$ высказыванием

$$\hat{A}\vec{x} = \alpha \cdot \vec{x}.$$

Полученный оператор называется ***скалярным***. Используем сформулированное выше правило нахождения матрицы оператора. Для этого подействуем оператором на базисные векторы $\left\{\vec{e}_1, \vec{e}_2, \vec{e}_3\right\}$ пространства X^3 и разложим образы базисных векторов по векторам того же базиса:

$$\hat{A}\vec{e}_1 = \alpha\cdot\vec{e}_1 = \alpha\left(1\cdot\vec{e}_1 + 0\cdot\vec{e}_2 + 0\cdot\vec{e}_3\right) = \alpha\cdot\vec{e}_1 + 0\cdot\vec{e}_2 + 0\cdot\vec{e}_3,$$

$$\hat{A}\vec{e}_2 = \alpha\cdot\vec{e}_2 = \alpha\left(0\cdot\vec{e}_1 + 1\cdot\vec{e}_2 + 0\cdot\vec{e}_3\right) = 0\cdot\vec{e}_1 + \alpha\cdot\vec{e}_2 + 0\cdot\vec{e}_3,$$

$$\hat{A}\vec{e}_3 = \alpha\cdot\vec{e}_3 = \alpha\left(0\cdot\vec{e}_1 + 0\cdot\vec{e}_2 + 1\cdot\vec{e}_3\right) = 0\cdot\vec{e}_1 + 0\cdot\vec{e}_2 + \alpha\cdot\vec{e}_3.$$

Здесь мы использовали правила умножения векторов на числа. Помещая коэффициенты в разложении образов базисных векторов на соответствующие места в столбцах матрицы оператора, получаем:

$$\mathrm{A} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}.$$

Полученная матрица также называется ***скалярной матрицей***. ⊗

Действия с линейными операторами и матрицами. Дадим теперь определения действий с (линейными) операторами и их матрицами.

Определение 3.1.4. *Операторы* $\hat{A}: X^m \to Y^n$ *и* $\hat{B}: X^m \to Y^n$ *считаются* ***равными*** *и пишут* $\hat{A} = \hat{B}$, *если*

$$\left(\forall \vec{x} \in R^m\right) \hat{A}\vec{x} = \hat{B}\vec{x}, \tag{3.1.9}$$

то есть ***при действии этих операторов на один и тот же вектор из пространства прообразов соответствующие образы в пространстве образов равны***. •

Определим равенство соответствующих матриц, используя правило нахождения матрицы линейного оператора. Для этого, зафиксировав в пространстве прообразов и в пространстве образов базисы

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\right\} \subset X^m, \left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset Y^n,$$

подействуем операторами $\hat{A}: X^m \to Y^n$ и $\hat{B}: X^m \to Y^n$ последовательно на базисные векторы пространства прообразов и разложим получившиеся образы по базису пространства образов:

$$\hat{A}\vec{e}_k = \sum_{j=1}^{n} a_k^j \vec{g}_j,\ \hat{B}\vec{e}_k = \sum_{j=1}^{n} b_k^j \vec{g}_j.$$

В силу того, что образы равны $\hat{A}\vec{e}_k = \hat{B}\vec{e}_k$, $k = 1, 2, \ldots, m$, получаем

$$\sum_{j=1}^{n} a_k^j \vec{g}_j = \sum_{j=1}^{n} b_k^j \vec{g}_j \Rightarrow \sum_{j=1}^{n} \left(a_k^j - b_k^j\right)\vec{g}_j = \vec{0},$$

или в развернутом виде

$$\left(a_k^1 - b_k^1\right)\vec{g}_1 + \left(a_k^2 - b_k^2\right)\vec{g}_2 + \ldots + \left(a_k^n - b_k^n\right)\vec{g}_n = \vec{0}.$$

Так как система базисных векторов пространства Y^n линейно независима, то по определению получаем

$$a_k^j = b_k^j \tag{3.1.10}$$

для любых $k = 1, 2, \ldots, m$ и $j = 1, 2, \ldots, n$. Итак, равенство матриц

$$\begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_m^n \end{pmatrix} = \begin{pmatrix} b_1^1 & b_2^1 & \ldots & b_m^1 \\ b_1^2 & b_2^2 & \ldots & b_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ b_1^n & b_2^n & \ldots & b_m^n \end{pmatrix}$$

означает, что для любых пар индексов $\{j, k\}$ выполнены равенства (3.1.10).

Установим теперь правило сложения операторов, определив их сумму.

Определение 3.1.5. *Оператор* $\hat{C} : X^m \to Y^n$ *называется* ***суммой оператора*** $\hat{A} : X^m \to Y^n$ ***и оператора*** $\hat{B} : X^m \to Y^n$ *и пишут* $\hat{C} = \hat{A} + \hat{B}$*, если*

$$\left(\forall \vec{x} \in X^m\right) \hat{C}\vec{x} = \left(\hat{A} + \hat{B}\right)\vec{x} \overset{def}{=} \hat{A}\vec{x} + \hat{B}\vec{x}, \tag{3.1.11}$$

то есть образ произвольного вектора $\vec{x} \in X^m$ *при действии суммы операторов равен сумме образов этого вектора при действии каждого из операторов.* •

Установим правило сложения соответствующих матриц. Для этого подействуем суммой операторов на произвольный базисный вектор пространства прообразов и результат разложим по базису пространства образов:

$$\hat{C}\vec{e}_k = \left(\hat{A} + \hat{B}\right)\vec{e}_k = \hat{A}\vec{e}_k + \hat{B}\vec{e}_k = \sum_{j=1}^{n} a_k^j \vec{g}_j + \sum_{j=1}^{n} b_k^j \vec{g}_j = \sum_{j=1}^{n}\left(a_k^j + b_k^j\right)\vec{g}_j .$$

С другой стороны, для оператора $\hat{C}$ имеем

$$\hat{C}\vec{e}_k = \sum_{j=1}^{n} c_k^j \vec{g}_j .$$

Здесь через c_k^j обозначены элементы матрицы суммы операторов. Таким образом, получаем следующее равенство:

$$\sum_{j=1}^{n} c_k^j \vec{g}_j = \sum_{j=1}^{n}\left(a_k^j + b_k^j\right)\vec{g}_j ,$$

из которого в силу линейной независимости системы базисных векторов пространства образов получаем, что

$$\left(\forall j = 1, 2, \ldots, n\right) \text{ и } \left(\forall k = 1, 2, \ldots, m\right),\ c_k^j = a_k^j + b_k^j. \tag{3.1.12}$$

Матрица C оператора $\hat{C} = \hat{A} + \hat{B}$ называется ***суммой матриц A и B***, что записывается в виде $C = A + B$.

Можно показать (рекомендуется сделать это в качестве упражнения), что ***сумма линейных операторов снова является линейным оператором***.

Определим теперь действие умножения оператора на число.

Определение 3.1.6. *Оператор* $\hat{C} : X^m \to Y^n$ *называется* ***произведением оператора*** $\hat{A} : X^m \to Y^n$ ***на число*** $\alpha \in R^1$ *и пишут* $\hat{C} = \alpha \cdot \hat{A}$*, если*

$$\left(\forall \vec{x} \in R^m\right) \hat{C}\vec{x} = \left(\alpha \cdot \hat{A}\right)\vec{x} = \alpha \cdot \left(\hat{A}\vec{x}\right), \tag{3.1.13}$$

то есть образ произвольного вектора $\vec{x} \in X^m$ *при действии произведения оператора на число равен произведению образа вектора при действии оператора* $\hat{A}$ *на данное число.* •

Правило получения матрицы произведения оператора на число устанавливается аналогично случаю суммы операторов и записывается в виде:

$$\left(\forall j = 1, 2, \ldots, n\right) \text{ и } \left(\forall k = 1, 2, \ldots, m\right) c_k^j = \alpha \cdot a_k^j, \tag{3.1.14}$$

где через c_k^j обозначены элементы матрицы произведения оператора $\hat{A}$ на число α. Матрица C оператора $\hat{C} = \alpha \cdot \hat{A}$ называется ***произведением матрицы A на число*** $\alpha \in R^1$, что записывается в виде $C = \alpha \cdot A$.

Нетрудно показать (рекомендуется сделать это в качестве упражнения), что ***произведение линейного оператора на число снова является линейным оператором***.

Пусть $\hat{A}: X^m \to Y^n$ и $\hat{B}: Y^n \to Z^p$ — два линейных оператора. Определим ***композицию*** этих операторов (иллюстрация приведена на рис. 3.1 на примере плоских диаграмм, на которых операторы изображены стрелками).

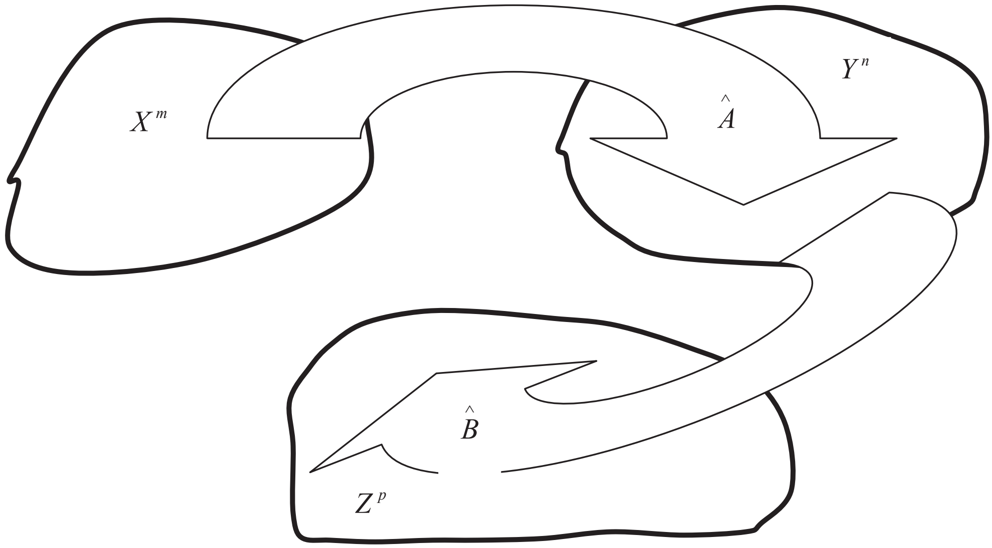


Рис. 3.1

Определение 3.1.7. *Композицией* $\hat{C} = \hat{B}\hat{A}$ *операторов* $\hat{A}: X^m \to Y^n$ *и* $\hat{B}: Y^n \to Z^p$ *называется оператор* $\hat{C}: X^m \to Z^p$*, действие которого производится по правилу*:

$$\hat{C}\vec{x} \stackrel{def}{=} \left(\hat{B}\hat{A}\right)\vec{x} = \hat{B}\left(\hat{A}\vec{x}\right). \bullet \tag{3.1.15}$$

Матрица композиции линейных операторов называется произведением матриц соответствующих операторов и записывается в виде $C = BA$. Получим форму для произведения матриц.

Для этого зафиксируем в пространствах X^m, Y^n и Z^p некоторые базисы

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\right\} \subset X^m,$$

$$\left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset Y^n,$$

$$\left\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_p\right\} \subset Z^p$$

и подействуем на произвольный вектор $\vec{e}_k$ базиса пространства прообразов X^m композицией операторов.

Цепочка преобразований выглядит так:

$$\hat{C}\vec{e}_k = \left(\hat{B}\hat{A}\right)\vec{e}_k = \hat{B}\left(\hat{A}\vec{e}_k\right) = \hat{B}\left(\sum_{j=1}^{n} a_k^j \vec{g}_j\right) = \sum_{j=1}^{n} a_k^j \hat{B}\vec{g}_j =$$

$$= \sum_{j=1}^{n} a_k^j \sum_{i=1}^{p} b_j^i \vec{q}_i = \sum_{i=1}^{p}\left(\sum_{j=1}^{n} b_j^i a_k^j\right)\vec{q}_i.$$

С другой стороны,

$$\hat{C}\vec{e}_k = \sum_{i=1}^{p} c_k^i \vec{q}_i.$$

Сравнивая правые части этих равенств, получаем:

$$c_k^i = \sum_{j=1}^{n} b_j^i a_k^j. \tag{3.1.16}$$

Можно показать (рекомендуется сделать это в качестве упражнения), что ***композиция линейных операторов снова является линейным оператором***.

Из формулы (3.1.16) следует правило, упрощающее запоминание выполнения умножения матриц:

элемент c_k^i произведения матриц A и B (записывается в порядке BA) получается путём вычисления скалярного произведения вектор-строки с номером i матрицы, стоящей в произведении слева, и вектор-столба с номером k матрицы, стоящей в произведении справа.

Из самого определения композиции операторов следует, что для выполнимости умножения матриц должно быть выполненным условие согласования размерностей: ***число столбцов матрицы, стоящей***

в произведении слева, должно равняться числу строк матрицы, стоящей в произведении справа.

Теперь легко определить степень оператора $\hat{A}: X \to X$.

Определение 3.1.8. ***Степенью*** $\left(\hat{A}\right)^p$ *линейного оператора* $\hat{A}: X \to X$ *называется оператор, который определяется следующими рекуррентными соотношениями:*

$$\left(\hat{A}\right)^0 \overset{def}{=} \hat{I},\ \left(\hat{A}\right)^p \overset{def}{=} \underbrace{\hat{A}\hat{A}\ldots\hat{A}}_{p}:$$

$$\left(\forall \vec{x} \in X\right)\left(\hat{A}\right)^0 \vec{x} = \hat{I}\vec{x},\ \left(\hat{A}\right)^p \vec{x} = \underbrace{\hat{A}\left(\hat{A}\left(\ldots\hat{A}\left(\hat{A}\vec{x}\right)\right)\right)}_{p}. \bullet$$

Векторное пространство линейных операторов. Рассмотрим множество Ω_{XY} операторов $\hat{A}: X \to Y$.

Теорема 3.1.6. *Множество операторов* Ω_{XY} *является векторным пространством*.

Д о к а з а т е л ь с т в о. Для доказательства утверждения теоремы нужно проверить выполнение всех аксиом векторного пространства.

1) Выше было показано, что операция сложения линейных операторов является алгебраической на множестве операторов Ω_{XY}.

2) Покажем, что операция сложения операторов коммутативна. Действительно:

$$\left(\hat{A} + \hat{B}\right)\vec{x} = \hat{A}\vec{x} + \hat{B}\vec{x} = \hat{B}\vec{x} + \hat{A}\vec{x} = \left(\hat{B} + \hat{A}\right)\vec{x}.$$

3) Покажем, что операция сложения операторов ассоциативна. Действительно,

$$\left(\left(\hat{A} + \hat{B}\right) + \hat{C}\right)\vec{x} = \left(\hat{A} + \hat{B}\right)\vec{x} + \hat{C}\vec{x} =$$

$$= \hat{A}\vec{x} + \hat{B}\vec{x} + \hat{C}\vec{x} = \hat{A}\vec{x} + \left(\hat{B}\vec{x} + \hat{C}\vec{x}\right) = \left(\hat{A} + \left(\hat{B} + \hat{C}\right)\right)\vec{x}.$$

4) В качестве нулевого элемента выбираем определённый выше нулевой оператор $\hat{O}$.

5) В качестве обратного (противоположного) элемента для оператора $\hat{A}$ выбираем определённый выше противоположный оператор $-\hat{A}$.

Из 1) — 5) следует, что ***множество*** Ω_{XY} ***является аддитивной абелевой группой по операции сложения операторов***.

6) Внешняя операция умножения оператора на число обладает следующими свойствами: $\left(\forall \hat{A}, \hat{B} \in \Omega_{XY}\right) \wedge (\forall \alpha, \beta \in P)$

$$1)\ 1 \cdot \hat{A} = \hat{A}; \quad 2)\ \alpha \cdot \left(\hat{A} + \hat{B}\right) = \alpha \cdot \hat{A} + \alpha \cdot \hat{B};$$

$$3)\ (\alpha + \beta) \cdot \hat{A} = \alpha \cdot \hat{A} + \beta \cdot \hat{A};\ 4)\ (\alpha \cdot \beta) \cdot \hat{A} = \alpha \cdot \left(\beta \cdot \hat{A}\right).$$

Справедливость, например свойства 3, вытекает из следующей цепочки равенств: $\left(\forall \vec{x} \in X\right)$

$$\left((\alpha + \beta) \cdot \hat{A}\right)\vec{x} = (\alpha + \beta) \cdot \hat{A}\vec{x} = \alpha \cdot \hat{A}\vec{x} + \beta \cdot \hat{A}\vec{x} =$$

$$= \left(\alpha \cdot \hat{A}\right)\vec{x} + \left(\beta \cdot \hat{A}\right)\vec{x} = \left(\alpha \cdot \hat{A} + \beta \cdot \hat{A}\right)\vec{x}.$$

Таким образом, на множестве Ω_{XY} выполняются все аксиомы векторного пространства, что и доказывает утверждение теоремы. ••

Кольцо операторов. Пусть K_{XX} множество операторов $\hat{A} : X \to X$.

Теорема 3.1.7. *Множество операторов K_{XX} является кольцом.*

Д о к а з а т е л ь с т в о. Выше было отмечено, что множество операторов Ω_{XY} и, в частности, K_{XX} является аддитивной абелевой группой по сложению. Было отмечено также, что операция композиции (произведения) операторов на множестве K_{XX} является алгебраической. Чтобы доказать, что K_{XX} является кольцом, нужно показать, что операции сложения и композиции операторов на множестве K_{XX} связаны дистрибутивными законами, а именно, что для любых операторов $\hat{A}, \hat{B}, \hat{C} \in \Omega_{XX}$ выполнены следующие законы, связывающие операции сложения и композиции:

$$1)\ \hat{A}(\hat{B} + \hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C},\ 2)\ (\hat{B} + \hat{C})\hat{A} = \hat{B}\hat{A} + \hat{C}\hat{A}.$$

Докажем, например, первый из законов. Сначала отметим, что операторы в левой и правой части этого равенства являются элементами

K_{XX}. Пусть $\vec{x}$ — произвольный вектор из X. Тогда, используя определения операций сложения и композиции операторов, имеем:

$$\hat{A}(\hat{B}+\hat{C})\vec{x} = \hat{A}\left((\hat{B}+\hat{C})\vec{x}\right) = \hat{A}\left(\hat{B}\vec{x}+\hat{C}\vec{x}\right) = \hat{A}\hat{B}\vec{x}+\hat{A}\hat{C}\vec{x},$$

$$\left(\hat{A}\hat{B}+\hat{A}\hat{C}\right)\hat{x} = \hat{A}\hat{B}\vec{x}+\hat{A}\hat{C}\vec{x}.$$

Так как правые части равны, то равны и левые, что и доказывает справедливость данного дистрибутивного закона. Справедливость второго дистрибутивного закона показывается аналогично. Итак, множество K_{XX} операторов $\hat{A}: X \to X$ является кольцом. ••

Отметим, что кольцо операторов K_{XX} является ***ассоциативным кольцом***.

Действительно, $\left(\forall \vec{x} \in X\right) \wedge \left(\forall \hat{A}, \hat{B}, \hat{C} \in K(X, X)\right)$

$$\left(\left(\hat{A}\hat{B}\right)\hat{C}\right)\vec{x} = \hat{A}\hat{B}\left(\hat{C}\vec{x}\right) = \hat{A}\left(\hat{B}\left(\hat{C}\vec{x}\right)\right) = \hat{A}\left(\hat{B}\hat{C}\vec{x}\right) = \left(\hat{A}\left(\hat{B}\hat{C}\right)\right)\vec{x},$$

откуда следует ассоциативность операции композиции операторов:

$$\left(\hat{A}\hat{B}\right)\hat{C} = \hat{A}\left(\hat{B}\hat{C}\right).$$

Отметим ещё одно полезное свойство операции композиции операторов на множестве K_{XX}

$$\alpha \cdot \left(\hat{B}\hat{A}\right) = \left(\alpha \cdot \hat{B}\right)\hat{A} = \hat{B}\left(\alpha \cdot \hat{A}\right),$$

которое легко проверить, используя свойство однородности линейного оператора.

Группа невырожденных операторов. Рассмотрим множество G_{XX} ***невырожденных операторов*** $\hat{A}: X \to X$ с определённой на нём операцией композиции операторов.

Теорема 3.1.8. *Множество операторов G_{XX} является мультипликативной группой с групповой операцией композиции операторов.*

Доказательство. Выше показано, что операция композиции операторов $\hat{A}: X \to X$ является ассоциативной. В качестве единичного элемента можно взять введённый выше единичный оператор. Так как операторы из множества G_{XX} невырожденные, то для них существует обратный оператор, который и принимается за обратный элемент. Таким образом, для элементов множества G_{XX} выполнены все аксиомы группы.••

Композиция операторов, вообще говоря, не является коммутативной операцией, что следует из свойств операции умножения матриц. Поэтому группа G_{XX} является некоммутативной. На множестве операторов G_{XX} можно построить и мультипликативные абелевы группы операторов, взяв в качестве их элементов степени операторов.

Алгебра операторов. Выше на множестве операторов $\hat{A}: X \to X$ введена внешняя операция — умножение операторов на числа из некоторого поля. Обозначая множество операторов с внешней операцией умножения на числа символом A_{XX}, сформулируем следующую теорему.

Теорема 3.1.9. *Множество операторов A_{XX} является некоммутативной алгеброй.*

Д о к а з а т е л ь с т в о. Выше уже показано, что множество всех операторов A_{XX} является ассоциативным кольцом. Требуется показать, что внешняя операция обладает следующими свойствами: $\left(\forall \hat{A}, \hat{B} \in A_{XX} \wedge \forall \alpha \in P\right)$

1) $(\alpha+\beta)\cdot \hat{A} = \alpha\cdot\hat{A} + \beta\cdot\hat{A}$; 2) $\alpha\cdot\left(\hat{A}+\hat{B}\right) = \alpha\cdot\hat{A}+\alpha\cdot\hat{B}$;

3) $1\cdot\hat{A}=\hat{A}$; 4) $\left(\alpha\cdot\hat{A}\right)\left(\beta\cdot\hat{B}\right) = (\alpha\beta)\cdot\left(\hat{A}\hat{B}\right)$.

Доказательство этих свойств не представляет труда. ••

3.2. Определитель как функция

Определители. На множестве квадратных матриц, представляющих в фиксированном базисе операторы (преобразования) $\hat{A}: X^n \to X^n$, вводится понятие ***функции***, которая носит название ***определитель***.

Определение 3.2.1. *Отображение D множества квадратных* $(n \times n)$ *матриц во множество действительных чисел называется* ***определителем (детерминантом)****, если его значение на каждой матрице из указанного множества находится по следующим правилам:*

1) *значением отображения D на матрице первого порядка* $\left(a_1^1\right)$ *является её единственный элемент* $D\left(a_1^1\right) \equiv \det\left(a_1^1\right) = a_1^1$;

2) *значением отображения D на матрице*

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix} \tag{3.2.1}$$

порядка $n > 1$ *является число, вычисляемое по формуле*

$$\det A = \sum_{k=1}^{n} (-1)^{1+k} a_k^1 M_k^1 =$$

$$= (-1)^{1+1} a_1^1 M_1^1 + (-1)^{1+2} a_2^1 M_2^1 + \dots + (-1)^{1+n} a_n^1 M_n^1, \tag{3.2.2}$$

где M_k^1 *— определитель матрицы порядка* $n-1$*, полученной из матрицы D вычёркиванием первой строки и столбца с номером* $k = \overline{1, n}$. •

Для краткости значение отображения D на множестве квадратных матриц $A = \left(a_j^i\right)_{j=1,2,\dots,n}^{i=1,2,\dots,n}$ также принято называть ***определителем*** — именно в этом смысле мы будем использовать данный термин в дальнейшем. Определитель M_k^1 матрицы порядка $n-1$, фигурирующий в определении 3.2.1, называется ***минором*** элемента a_k^1 матрицы A, а формула (3.2.2) называется формулой ***разложения определителя*** по элементам первой строки. Определитель обозначается так:

$$\det A = \det \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix} = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{vmatrix}.$$

Определители обладают рядом свойств, которые можно проверить непосредственно для определителей малого порядка, например порядка 3. По этой причине мы перечислим эти свойства, отсылая за доказательствами к списку литературы.

1. *Формула, аналогичная формуле (3.2.2), справедлива и для разложения определителя по элементам первого столбца, то есть*

$$\begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{vmatrix} = (-1)^{1+1} a_1^1 M_1^1 + (-1)^{2+1} a_1^2 M_1^2 + \dots + (-1)^{n+1} a_1^n M_1^n. \quad (3.2.3)$$

Здесь M_1^k — минор элемента a_1^k, расположенного в матрице на пересечении первого столбца и строки с номером $k = \overline{1, n}$.

2. *Формула, аналогичная формулам (3.2.2) и (3.2.3), справедлива и для разложения определителя по элементам любой строки или любого столбца определителя; например, для столбца с номером k получаем:*

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} a_k^i M_k^i =$$

$$= (-1)^{1+k} a_k^1 M_k^1 + (-1)^{2+k} a_k^2 M_k^2 + \dots + (-1)^{n+k} a_k^n M_k^n. \quad (3.2.4)$$

Здесь M_j^i — минор элемента матрицы a_j^i, расположенного на пересечении строки с номером i и столбца с номером j.

Поставим в соответствие матрице (3.2.1) матрицу

$$A^T \overset{def}{=} \begin{pmatrix} a_1^1 & a_1^2 & \dots & a_1^n \\ a_2^1 & a_2^2 & \dots & a_2^n \\ \dots & \dots & \dots & \dots \\ a_n^1 & a_n^2 & \dots & a_n^n \end{pmatrix},$$

у которой строки поменялись местами со столбцами. Такая матрица называется ***транспонированной матрицей*** по отношению к матрице (3.2.1). Легко видеть, что матрица A^T есть результат поворота на 180^o матрицы A вокруг главной диагонали.

3. Определитель матрицы A равен определителю транспонированной к ней матрицы, то есть

$$\det A = \det A^T.$$

4. *Если к элементам какой-либо строки (какого-либо столбца) определителя прибавить соответствующие элементы любой другой его стро-*

ки (любого другого столбца), умноженные на одно и то же число, то значение определителя не изменится.

5. *Если в определителе поменять местами две какие-либо строки (два каких-либо столбца), определитель изменит свой знак на противоположный.*

6. *Определитель, в котором имеются две одинаковые строки (два одинаковых столбца), равен нулю.*

7. *Если все элементы какой-либо строки (столбца) определителя равны нулю, то он равен нулю.*

8. *Система n векторов* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\right\} \subset R^n$ *линейно зависима в том и только в том случае, если определитель матрицы, составленной из вектор-столбцов (строк) координат векторов системы, равен нулю.*

9. *Определитель произведения двух матриц равен произведению их определителей.*

10. *Если все элементы j-го столбца определителя D представлены в виде линейной комбинации двух слагаемых*

$$a_j^i = \lambda b_j^i + \mu c_j^i \; (i = 1, 2, \ldots, n),$$

где λ *и* μ *— фиксированные числа, то определитель представляется в виде суммы*

$$D = \lambda D_1 + \mu D_2,$$

причём у каждого из определителей D_1 *и* D_2 *все столбцы, кроме j-го, такие же, как и у определителя D, а j-й столбец у определителя* D_1 *состоит из чисел* b_j^i*, а у определителя* D_2 *— из чисел* c_j^i.

Формулу разложения определителя по элементам строки с номером i можно записать в виде

$$\begin{vmatrix} a_1^1 & a_2^1 & \ldots & a_j^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_j^2 & \ldots & a_n^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_1^i & a_2^i & \ldots & a_j^i & \ldots & a_n^i \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_j^n & \ldots & a_n^n \end{vmatrix} =$$

$$= (-1)^{i+1} a_1^i M_1^i + (-1)^{i+2} a_2^i M_2^i + \ldots + (-1)^{i+n} a_n^i M_n^i =$$

$$= A_1^i a_1^i + A_2^i a_2^i + \ldots + A_n^i a_n^i,$$

если определить ***алгебраическое дополнение*** A_j^i элемента a_j^i, положив

$$A_j^i \overset{def}{=} (-1)^{i+j}\, M_j^i. \tag{3.2.5}$$

Аналогичную формулу можно написать и для разложения определителя по элементам любой строки или любого столбца. Используя определение (3.2.5), сформулируем ещё одно важное свойство определителей.

11. *Сумма произведений элементов какой-либо строки (какого-либо столбца) определителя на алгебраические дополнения какой-либо другой строки (какого-либо другого столбца) определителя равна нулю.*

Квадратная матрица (3.2.1) называется ***невырожденной***, если её определитель не равен нулю:

$$\det A = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{vmatrix} \neq 0.$$

Обратная матрица и формулы Крамера. Если записать формулу действия оператора на вектор в матрично-векторном виде

$$A|x\rangle = |y\rangle,$$

то в предположении $\det A \neq 0$ для ***обратного оператора*** $\hat{A}^{-1}$, удовлетворяющего условиям

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I}, \tag{3.2.6}$$

вводится понятие представляющей его ***обратной матрицы*** A^{-1}, для которой условия (3.2.6) имеют тот же вид:

$$A^{-1}A = AA^{-1} = I. \tag{3.2.7}$$

Здесь $\hat{I}$ и I — соответственно единичный оператор, и его матрица, имеющая вид

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \tag{3.2.8}$$

также называется ***единичной матрицей***.

Получим формулу для вычисления обратной матрицы. Для этого матрице оператора A поставим в соответствие транспонированную матрицу A^T

$$A^T = \begin{pmatrix} a_1^1 & a_1^2 & \dots & a_1^n \\ a_2^1 & a_2^2 & \dots & a_2^n \\ \dots & \dots & \dots & \dots \\ a_n^1 & a_n^2 & \dots & a_n^n \end{pmatrix},$$

повернув матрицу A на 180° вокруг главной диагонали в пространстве. Заменяя в транспонированной матрице элементы их алгебраическими дополнениями (3.2.5) элементов матрицы A, получим матрицу следующего вида:

$$adjA \overset{def}{=} \begin{pmatrix} A_1^1 & A_1^2 & \dots & A_1^n \\ A_2^1 & A_2^2 & \dots & A_2^n \\ \dots & \dots & \dots & \dots \\ A_n^1 & A_n^2 & \dots & A_n^n \end{pmatrix}. \qquad (3.2.9)$$

Матрица $adj\,A$ (читается «эджектив») называется ***союзной***, или ***присоединённой*** матрицей к матрице A. Из формулы разложения и десятого свойства определителей вытекает легко проверяемая формула

$$adj\,A \cdot A = \det A \cdot I,$$

из которой при условии $\det A \neq 0$ получаем

$$\frac{1}{\det A} adj\,A \cdot A = I.$$

Из последней формулы и определения обратной матрицы (3.2.7) следует, что ***при условии*** $\det A \neq 0$ для матрицы A ***существует обратная матрица***

$$A^{-1} = \frac{1}{\det A} adj\,A, \qquad (3.2.10)$$

где $adj\,A$ — ***матрица, союзная к матрице*** A***, которая вычисляется по формуле*** (3.2.9).

Из существования обратной матрицы теперь можно сделать вывод и о существовании обратного оператора.

Оператор $\hat{A}$, матрица которого удовлетворяет условию $\det A \neq 0$, называется ***невырожденным*** (матрица A при условии $\det A \neq 0$ также называется ***невырожденной***). Как будет видно из доказанной ниже леммы, это определение совпадает с введённым выше определением невырожденного оператора. Напомним, что для обратной матрицы, если она существует, выполняется следующее из условия (3.2.8) аналогичное условие (3.2.7).

Запишем систему линейных алгебраических уравнений (3.1.6) для случая квадратной матрицы

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_n^1 x^n = b^1, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_n^2 x^n = b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n x^1 + a_2^n x^2 + \ldots + a_n^n x^n = b^n \end{cases} \tag{3.2.11}$$

в матричном виде

$$A|x\rangle = |y\rangle,$$

где A — квадратная матрица (3.2.1), $|y\rangle$ — вектор-столбец правых частей СЛАУ. Пусть $\det A \neq 0$, тогда, умножая обе части системы уравнений слева на матрицу A^{-1}, нетрудно видеть, что ***формула для нахождения решения*** этой системы при помощи обратной матрицы имеет вид

$$|x\rangle = A^{-1}|y\rangle. \tag{3.2.12}$$

Действительно, умножая обе части системы уравнений, записанной в матричной форме, слева на матрицу A^{-1}, получаем цепочку

$$A^{-1}A|x\rangle = A^{-1}|y\rangle \Rightarrow \left(A^{-1}A\right)|x\rangle = A^{-1}|y\rangle \Rightarrow I|x\rangle = A^{-1}|y\rangle,$$

что эквивалентно формуле (3.2.12).

Формулу (3.2.12) можно записать в развёрнутом виде

$$\begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^n \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} A_1^1 & A_1^2 & \ldots & A_1^n \\ A_2^1 & A_2^2 & \ldots & A_2^n \\ \ldots & \ldots & \ldots & \ldots \\ A_n^1 & A_n^2 & \ldots & A_n^n \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ \ldots \\ y^n \end{pmatrix}. \tag{3.2.13}$$

Производя умножение вектор-столбца в правой части (3.2.13) на матрицу и используя запись формулы разложения определителя через алгебраические дополнения, перепишем формулу решения (3.2.13) СЛАУ (3.2.11) в виде ***формул Крамера***, а именно для неизвестного с номером $k = 1, 2, \ldots, n$ имеем:

$$x^k = \frac{1}{\det A}\begin{vmatrix} a_1^1 & \dots & a_{k-1}^1 & b^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & \dots & a_{k-1}^2 & b^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^n & \dots & a_{k-1}^n & b^n & a_{k+1}^n & \dots & a_n^n \end{vmatrix}. \qquad (3.2.14)$$

В заключение отметим, что ***операция композиции на множестве всех невырожденных операторов $\hat{A}: X^n \to X^n$ (квадратных матриц) является некоммутативной, то есть $\hat{B}\hat{A} \neq \hat{A}\hat{B}$.***

Критерий невырожденности линейного оператора. Столбцы (строки) $(n \times n)$ матрицы оператора $\hat{A}: X^n \to X^n$ и её определителя

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix}, \det A = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{vmatrix} \qquad (3.2.15)$$

можно рассматривать как вектор-столбцы координатного пространства R^n, а именно:

$$|a_1\rangle = \begin{pmatrix} a_1^1 \\ a_1^2 \\ \dots \\ a_1^n \end{pmatrix}, |a_2\rangle = \begin{pmatrix} a_2^1 \\ a_2^2 \\ \dots \\ a_2^n \end{pmatrix}, \dots, |a_n\rangle = \begin{pmatrix} a_n^1 \\ a_n^2 \\ \dots \\ a_n^n \end{pmatrix}.$$

Следовательно, для столбцов (строк) матрицы A и её определителя справедливы все теоремы о линейной зависимости и независимости систем векторов. Установим критерий невырожденности линейного оператора $\hat{A}: X^n \to X^n$.

Лемма 3.2.1. *Система столбцов квадратной матрицы линейно зависима в том и только в том случае, если её определитель равен нулю.*

Доказательство. Пусть система столбцов матрицы A линейно зависима. Тогда, согласно третьему свойству определителей, некоторый столбец определителя (3.2.15) можно представить в виде линейной комбинации остальных столбцов. По десятому свойству определитель матрицы (3.2.1) в этом случае распадётся на сумму $n-1$ определителей, каждый из которых будет иметь по два равных столбца, но тогда $\det A = 0$.

Покажем теперь, что если $\det A = 0$, то система его столбцов линейно зависима. Действительно, потребуем выполнения равенства

$$\alpha_1 \begin{pmatrix} a_1^1 \\ a_1^2 \\ \dots \\ a_1^n \end{pmatrix} + \alpha_2 \begin{pmatrix} a_2^1 \\ a_2^2 \\ \dots \\ a_2^n \end{pmatrix} + \dots + \alpha_n \begin{pmatrix} a_n^1 \\ a_n^2 \\ \dots \\ a_n^n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}.$$

Из этого равенства вытекает СЛАУ относительно введённых неизвестных — коэффициентов $\alpha_1, \alpha_2, \dots, \alpha_n$, имеющая вид

$$\begin{cases} a_1^1\alpha_1 + a_2^1\alpha_2 + \dots + a_n^1\alpha_n = 0, \\ a_1^2\alpha_1 + a_2^2\alpha_2 + \dots + a_n^2\alpha_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \\ a_1^n\alpha_1 + a_2^n\alpha_2 + \dots + a_n^n\alpha_n = 0. \end{cases} \tag{3.2.16}$$

Полученная СЛАУ (3.2.16) называется ***однородной***. Очевидно, что у СЛАУ (3.2.16) всегда существует решение $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, которое называется ***тривиальным решением***. Нас интересует наличие у СЛАУ ***нетривиальных решений***, то есть ***ненулевых*** решений.

Умножая обе части уравнений СЛАУ (3.2.16) соответственно на алгебраические дополнения $A_1^1, A_1^2, \dots, A_1^n$ элементов первого столбца определителя СЛАУ и складывая результаты, получим:

$$\left(a_1^1 A_1^1 + a_1^2 A_1^2 + \dots + a_1^n A_1^n\right)\cdot\alpha_1 + \left(a_2^1 A_1^1 + a_2^2 A_1^2 + \dots + a_2^n A_1^n\right)\cdot\alpha_2 +$$
$$+\left(a_n^1 A_1^1 + a_n^2 A_1^2 + \dots + a_n^n A_1^n\right)\cdot\alpha_n = 0.$$

Нетрудно видеть, что множитель при α_1 равен $\det A$. Множители при остальных неизвестных представляют собой суммы произведений алгебраических дополнений элементов первого столбца на элементы второго, третьего, ..., n-го столбцов. Но по свойству 7 определителей такие суммы равны нулю. Поэтому получаем, что произведение

$$\det A \cdot \alpha_1 = 0.$$

Так как по предположению $\det A = 0$, то α_1 может принимать любые конечные ненулевые значения.

Аналогично умножая уравнения системы (3.2.16) последовательно на алгебраические дополнения элементов второго, третьего, ..., n-го столбцов определителя и складывая результаты, можно показать, что оставшиеся коэффициенты $\alpha_1, \alpha_2, \ldots, \alpha_n$ могут принимать любые конечные ненулевые значения.

Таким образом, можно подобрать такие неравные одновременно нулю числа $\alpha_1, \alpha_2, \ldots, \alpha_n$, что требуемое равенство выполняется. Последнее и означает, что система вектор-столбцов линейно зависима. ••

Логическим противопоставлением этой лемме служит следующее утверждение.

Лемма 3.2.2. *Система столбцов квадратной матрицы линейно независима в том и только в том случае, если её определитель неравен нулю.*

Утверждение леммы 3.2.2 даёт критерий невырожденности матрицы и линейного оператора, представителем которого она является.

3.3. Преобразование координат вектора и матрицы линейного оператора при изменении базиса

При изменении базиса

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \to \left\{\vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'}\right\}$$

пространства очевидным образом изменяются координаты (произвольного) вектора, а также матрица действующего в пространстве оператора $\hat{T} : X^n \to X^n$. Рассмотрим последовательно преобразование координат произвольного вектора и матрицы оператора при переходе от одного базиса пространства X^n к другому базису.

Преобразование базисных векторов. Пусть в векторном пространстве X^n заданы два базиса, «***старый***» базис

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$$

и «***новый***» базис

$$\left\{\vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'}\right\} \subset X^n.$$

Заметим, что мы рассматриваем только преобразования «***типа поворота***», так как в абстрактном векторном пространстве понятие начала репера введено быть не может и, следовательно, о параллельном переносе речь не идёт. Штрих в обозначении вектора базиса показывает, что вектор принадлежат новому базису, но при этом $n = n'$.

Разложим векторы нового базиса по векторам старого базиса $\left(i' = \overline{1, n}\right)$:

$$\vec{a}_{i'} = A_{i'}^{1}\vec{a}_{1} + A_{i'}^{2}\vec{a}_{2} + \ldots + A_{i'}^{n}\vec{a}_{n} = \sum_{j=1}^{n} A_{i'}^{j}\vec{a}_{j}. \qquad (3.3.1)$$

Здесь коэффициенты $A_{i'}^{j}$ являются элементами матрицы преобразования от векторов старого базиса к векторам нового базиса, то есть

$$A : \left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \to \left\{\vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'}\right\},$$

или в матричном виде

$$\begin{pmatrix} \vec{a}_{1'} \\ \vec{a}_{2'} \\ \ldots \\ \vec{a}_{n'} \end{pmatrix} = \begin{pmatrix} A_{1'}^{1} & A_{1'}^{2} & \ldots & A_{1'}^{n} \\ A_{2'}^{1} & A_{2'}^{2} & \ldots & A_{2'}^{n} \\ \ldots & \ldots & \ldots & \ldots \\ A_{n'}^{1} & A_{n'}^{2} & \ldots & A_{n'}^{n} \end{pmatrix} \begin{pmatrix} \vec{a}_{1} \\ \vec{a}_{2} \\ \ldots \\ \vec{a}_{n} \end{pmatrix}. \qquad (3.3.2)$$

Матрица A в соотношении (3.3.2) носит название «***матрица перехода от старого базиса к новому базису***».

Так как векторы нового базиса в совокупности образуют линейно независимую систему, то и строки матрицы (3.3.2) образуют линейно независимую систему как векторы пространства R^n. Поэтому выполняется следующее условие:

$$\det A = \begin{vmatrix} A_{1'}^{1} & A_{1'}^{2} & \ldots & A_{1'}^{n} \\ A_{2'}^{1} & A_{2'}^{2} & \ldots & A_{2'}^{n} \\ \ldots & \ldots & \ldots & \ldots \\ A_{n'}^{1} & A_{n'}^{2} & \ldots & A_{n'}^{n} \end{vmatrix} \neq 0.$$

Это единственное условие, которое накладывается на преобразование (3.3.1). Удовлетворяющие данному условию преобразования векторов базиса называются ***невырожденными***.

Разложим теперь векторы старого базиса по векторам нового базиса:

$$\vec{a}_{j} = A_{j}^{1'}\vec{a}_{1'} + A_{j}^{2'}\vec{a}_{2'} + \ldots + A_{j}^{n'}\vec{a}_{n'} = \sum_{k'=1}^{n'} A_{j}^{k'}\vec{a}_{k'}. \qquad (3.3.3)$$

Здесь коэффициенты $A_j^{k'}$ являются элементами матрицы преобразования от векторов нового базиса к векторам старого базиса, то есть

$$A':\left\{\vec{a}_{1'},\vec{a}_{2'},\ldots,\vec{a}_{n'}\right\}\to\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n\right\},$$

или в матричном виде

$$\begin{pmatrix}\vec{a}_1\\ \vec{a}_2\\ \ldots\\ \vec{a}_n\end{pmatrix}=\begin{pmatrix}A_1^{1'} & A_1^{2'} & \ldots & A_1^{n'}\\ A_2^{1'} & A_2^{2'} & \ldots & A_2^{n'}\\ \ldots & \ldots & \ldots & \ldots\\ A_n^{1'} & A_n^{2'} & \ldots & A_n^{n'}\end{pmatrix}\begin{pmatrix}\vec{a}_{1'}\\ \vec{a}_{2'}\\ \ldots\\ \vec{a}_{n'}\end{pmatrix}. \tag{3.3.4}$$

Матрица A' в соотношении (3.3.4) носит название «***матрица перехода от нового базиса к старому базису***».

Так как векторы старого базиса в совокупности образуют линейно независимую систему, то определитель этой матрицы неравен нулю, то есть

$$\det A'=\begin{vmatrix}A_1^{1'} & A_1^{2'} & \ldots & A_1^{n'}\\ A_2^{1'} & A_2^{2'} & \ldots & A_2^{n'}\\ \ldots & \ldots & \ldots & \ldots\\ A_n^{1'} & A_n^{2'} & \ldots & A_n^{n'}\end{vmatrix}\neq 0.$$

Лемма 3.3.1. *Матрицы перехода от старого к новому и от нового к старому базисам взаимно обратные.*

Доказательство. Для доказательства подставим $\vec{a}_j$ из формул (3.3.3) в формулы (3.3.1):

$$\vec{a}_{i'}=\sum_{j=1}^{n}A_{i'}^{j}\vec{a}_j=\sum_{j=1}^{n}A_{i'}^{j}\sum_{k'=1}^{n'}A_j^{k'}\vec{a}_{k'}=\sum_{k'=1}^{n'}\left(\sum_{j=1}^{n}A_j^{k'}A_{i'}^{j}\right)\vec{a}_{k'}. \tag{3.3.5}$$

Из формул (3.3.5) получаем

$$\sum_{j=1}^{n}A_j^{k'}A_{i'}^{j}=\delta_{i'}^{k'}=\begin{cases}1, k'=i';\\ 0, k'\neq i'.\end{cases} \tag{3.3.6}$$

Здесь $\delta_{i'}^{k'}$ – ***символ Кронекера***. Равенство (3.3.6) означает, что $A'=A^{-1}$.

Аналогично подставляя $\vec{a}_{k'}$ из формул (3.3.1) в формулы (3.3.3), получаем:

$$\vec{a}_j = \sum_{k'=1}^{n'} A_j^{k'} \vec{a}_{k'} = \sum_{k'=1}^{n'} A_j^{k'} \sum_{i=1}^{n} A_{k'}^{i} \vec{a}_i = \sum_{i=1}^{n} \left(\sum_{k'=1}^{n'} A_{k'}^{i} A_j^{k'} \right) \vec{a}_i . \qquad (3.3.7)$$

Из формул (3.3.7) получаем

$$\sum_{k'=1}^{n'} A_{k'}^{i} A_j^{k'} = \delta_j^i = \begin{cases} 1, i = j; \\ 0, i \neq j. \end{cases} \qquad (3.3.8)$$

Равенство (3.3.8) означает, что $A = (A')^{-1}$. Утверждение леммы доказано. ••

Преобразование координат вектора. Естественно, что при изменении базиса координаты произвольного вектора также изменяются, но сам вектор остаётся неизменным как геометрический или физический объект, то есть ***вектор инвариантен относительно изменения базиса***. Воспользуемся инвариантностью вектора для получения формул преобразования его координат при замене ***старого*** базиса ***новым*** базисом.

Итак, пусть старый базис $\left\{ \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n \right\} \subset X^n$ заменяется новым базисом $\left\{ \vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'} \right\} \subset X^n$. Разложим некоторый произвольный вектор $\vec{x}$ по векторам старого и нового базисов:

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{a}_i ; \qquad (3.3.9)$$

$$\vec{x} = \sum_{j'=1}^{n'} x^{j'} \vec{a}_{j'} . \qquad (3.3.10)$$

В силу инвариантности вектора относительно изменения базиса из разложений (3.3.9) и (3.3.10) имеем

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{a}_i = \sum_{j'=1}^{n'} x^{j'} \vec{a}_{j'} . \qquad (3.3.11)$$

Выразим координаты вектора $\vec{x}$ в новом базисе через его координаты в старом базисе. Для этого в левой части разложения (3.3.11) выразим векторы старого базиса через векторы нового базиса по формулам (3.3.3):

$$\sum_{i=1}^{n} x^i \vec{a}_i = \sum_{i=1}^{n} x^i \left(\sum_{j'=1}^{n'} A_i^{j'} \vec{a}_{j'} \right) = \sum_{j'=1}^{n'} \left(\sum_{i=1}^{n} A_i^{j'} x^i \right) \vec{a}_{j'} .$$

Подстановка в левую часть разложения (3.3.11) даёт

$$\sum_{j'=1}^{n'}\left(\sum_{i=1}^{n} A_i^{j'} x^i\right)\vec{a}_{j'} = \sum_{j'=1}^{n'} x^{j'}\,\vec{a}_{j'},$$

откуда получаем формулу преобразования координат вектора при замене ***старого*** базиса ***новым*** базисом:

$$x^{j'} = \sum_{i=1}^{n} A_i^{j'} x^i. \qquad (3.3.12)$$

В матричном виде

$$\begin{pmatrix} x^{1'} \\ x^{2'} \\ \dots \\ x^{n'} \end{pmatrix} = \begin{pmatrix} A_1^{1'} & A_2^{1'} & \dots & A_n^{1'} \\ A_1^{2'} & A_2^{2'} & \dots & A_n^{2'} \\ \dots & \dots & \dots & \dots \\ A_1^{n'} & A_2^{n'} & \dots & A_n^{n'} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix}. \qquad (3.3.13)$$

Из формул (3.3.13), (3.3.4) и (3.3.2) видим, что матрица $\left(A_i^{j'}\right)_{i=1,2,\dots,n}^{j'=1,2,\dots,n'}$ является транспонированной по отношению к матрице (3.3.4), которая является обратной для матрицы (3.3.2). Получили следующий результат.

Лемма 3.3.2. *При переходе в пространстве X^n от старого базиса к новому базису* ***старые координаты*** *произвольного вектора $\vec{x} \in X^n$ преобразуются в* ***новые координаты*** *при помощи* ***транспонированной обратной матрицы перехода от старого базиса к новому базису*** *или, что то же самое, при помощи* ***транспонированной матрицы перехода от нового базиса к старому базису****.*

Теперь формулу (3.3.13) можно записать в символическом виде

$$|x'\rangle = \left(A^{-1}\right)^T |x\rangle. \qquad (3.3.14)$$

Если теперь в правой части разложения (3.3.11) выразить векторы нового базиса через векторы старого базиса по формулам (3.3.1), то получаем

$$\sum_{i=1}^{n} x^i\,\vec{a}_i = \sum_{j'=1}^{n'} x^{j'}\,\vec{a}_{j'} = \sum_{j'=1}^{n'} x^{j'} \sum_{i=1}^{n} A_{j'}^{i}\,\vec{a}_i = \sum_{i=1}^{n}\left(\sum_{j'=1}^{n'} A_{j'}^{i} x^{j'}\right)\vec{a}_i,$$

откуда следует, что

$$x^i = \sum_{j'=1}^{n'} A_{j'}^{i} x^{j'}. \qquad (3.3.15)$$

В матричном виде

$$\begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} = \begin{pmatrix} A^1_{1'} & A^1_{2'} & \dots & A^1_{n'} \\ A^2_{1'} & A^2_{2'} & \dots & A^2_{n'} \\ \dots & \dots & \dots & \dots \\ A^n_{1'} & A^n_{2'} & \dots & A^n_{n'} \end{pmatrix} \begin{pmatrix} x^{1\,'} \\ x^{2\,'} \\ \dots \\ x^{n\,'} \end{pmatrix}. \qquad (3.3.16)$$

В формулах (3.3.16) матрица преобразования является транспонированной обратной к матрице перехода от нового базиса к старому, то есть транспонированной матрицей перехода от старого базиса к новому. Имеем такой результат.

Лемма 3.3.3. *При переходе в пространстве X^n от нового базиса к старому базису* ***новые координаты*** *произвольного вектора $\vec{x} \in X^n$ преобразуются в* ***старые координаты*** *при помощи* ***транспонированной обратной матрицы перехода от нового базиса к старому базису*** *или, что то же самое, при помощи* ***транспонированной матрицы перехода от старого базиса к новому базису****.*

Формулу (3.3.16) можно записать в символическом виде

$$|x\rangle = A^T\,|x'\rangle. \qquad (3.3.17)$$

Преобразование матрицы оператора. Как уже было сказано выше, для простоты рассматривается случай линейного оператора $\hat{T} : X^n \to X^n$, то есть преобразования. Зафиксируем в пространстве X^n старый и новый базисы:

$$\left\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\right\} \subset X^n; \left\{\vec{a}_{1'}, \vec{a}_{2'}, \dots, \vec{a}_{n'}\right\} \subset X^n.$$

Тогда представителем оператора является $n \times n$ матрица T, такая, что в старом базисе действие оператора на произвольный вектор $\vec{x} \in X^n$ в координатной форме запишется так

$$y^j = \sum_{i=1}^{n} T_i^j x^i, \qquad (3.3.18)$$

или в символическом виде

$$|y\rangle = T\,|x\rangle. \qquad (3.3.19)$$

Здесь учтено, что

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{a}_i, \; \vec{y} = \sum_{j=1}^{n} y^j \vec{a}_j.$$

Аналогичная формула в новом базисе имеет вид

$$y^{q'} = \sum_{p'=1}^{n'} T_{p'}^{q'} x^{p'}, \tag{3.3.20}$$

или в символическом виде

$$|y'\rangle = T'|x'\rangle. \tag{3.3.21}$$

В формуле (3.3.18) перейдём к новому базису

$$y^j = \sum_{i=1}^{n} T_i^j x^i \Rightarrow \sum_{k'=1}^{n'} A_{k'}^j y^{k'} = \sum_{i=1}^{n} T_i^j \sum_{p'=1}^{n'} A_{p'}^i x^{p'},$$

то есть

$$\sum_{k'=1}^{n'} A_{k'}^j y^{k'} = \sum_{p'=1}^{n'} \left(\sum_{i=1}^{n} T_i^j A_{p'}^i \right) x^{p'}. \tag{3.3.22}$$

Умножая выражения (3.3.22) слева на обратную матрицу $A_j^{q'}$, получим:

$$\sum_{j=1}^{n} A_j^{q'} \sum_{k'=1}^{n'} A_{k'}^j y^{k'} = \sum_{j=1}^{n} A_j^{q'} \sum_{p'=1}^{n'} \left(\sum_{i=1}^{n} T_i^j A_{p'}^i \right) x^{p'}.$$

Меняя порядок суммирования и учитывая, что

$$\sum_{j=1}^{n} A_j^{q'} A_{k'}^j = \delta_{k'}^{q'}$$

(матрицы взаимно обратные), имеем:

$$y^{q'} = \sum_{p'=1}^{n'} \left(\sum_{j=1}^{n} \sum_{i=1}^{m} A_j^{q'} T_i^j A_{p'}^i \right) x^{p'}. \tag{3.3.23}$$

Сравнивая формулы (3.3.23) и (3.3.20), приходим к формуле вычисления элементов матрицы оператора $\hat{T}$ в новом базисе:

$$T_{p'}^{q'} = \sum_{j=1}^{n} \sum_{i=1}^{n} A_j^{q'} T_i^j A_{p'}^i.$$

Теорема 3.3.1. *Матрица оператора* $\hat{T}: X^n \to X^n$ *при замене старого базиса* $\left\{ \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n \right\} \subset X^n$ *на новый базис* $\left\{ \vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'} \right\} \subset X^n$ *преобразуется по формуле*

$$T_{p'}^{q'} = \sum_{j=1}^{n} \sum_{i=1}^{n} A_j^{q'} T_i^j A_{p'}^i, \tag{3.3.24}$$

где $A_j^{q'}$ *— транспонированная обратная матрица перехода от старого базиса к новому базису (транспонированная матрица перехода от нового базиса к старому базису), а* $A_{p'}^{i}$ *— транспонированная матрица перехода от старого базиса к новому базису.*

Вывод формулы преобразования матрицы оператора в матричных обозначениях. Вывод формулы (3.3.24) можно провести, используя матричную форму записи, что несколько сокращает и, вероятно, упрощает усвоение выкладок.

Запишем формулу (3.3.19) в новом (штрихованном) базисе:

$$A^T\left|y'\right\rangle = TA^T\left|x'\right\rangle.$$

Умножая обе части последнего равенства слева на $\left(A^T\right)^{-1}$, получаем:

$$\left(A^T\right)^{-1}A^T\left|y'\right\rangle = \left(A^T\right)^{-1}TA^T\left|x'\right\rangle.$$

Так как

$$\left(A^T\right)^{-1}A^T = I,$$

приходим к формуле

$$\left|y'\right\rangle = \left(A^T\right)^{-1}TA^T\left|x'\right\rangle. \tag{3.3.25}$$

Сравнивая формулы (3.3.25) и (3.3.21), получаем

$$T' = \left(A^{-1}\right)^T TA^T. \tag{3.3.26}$$

В развёрнутом матричном виде формула (3.3.25) записывается так:

$$\begin{pmatrix} T_{1'}^{1'} & T_{2'}^{1'} & \dots & T_{n'}^{1'} \\ T_{1'}^{2'} & T_{2'}^{2'} & \dots & T_{n'}^{2'} \\ \dots & \dots & \dots & \dots \\ T_{1'}^{n'} & T_{2'}^{n'} & \dots & T_{n'}^{n'} \end{pmatrix} =$$

$$= \begin{pmatrix} A_1^{1'} & A_2^{1'} & \dots & A_n^{1'} \\ A_1^{2'} & A_2^{2'} & \dots & A_n^{2'} \\ \dots & \dots & \dots & \dots \\ A_1^{n'} & A_2^{n'} & \dots & A_n^{n'} \end{pmatrix} \begin{pmatrix} T_1^1 & T_2^1 & \dots & T_n^1 \\ T_1^2 & T_2^2 & \dots & T_n^2 \\ \dots & \dots & \dots & \dots \\ T_1^n & T_2^n & \dots & T_n^n \end{pmatrix} \begin{pmatrix} A_{1'}^1 & A_{2'}^1 & \dots & A_{n'}^1 \\ A_{1'}^2 & A_{2'}^2 & \dots & A_{n'}^2 \\ \dots & \dots & \dots & \dots \\ A_{1'}^n & A_{2'}^n & \dots & A_{n'}^n \end{pmatrix}.$$

3.4. Ранг матрицы и ранг оператора. Совместность СЛАУ общего вида

Выше получены ***конструктивные процедуры*** решения СЛАУ, то есть даны ***прямые алгоритмы*** отыскания решения. Часто знать решение СЛАУ не требуется, достаточно выяснить вопрос о её совместности. В этом параграфе рассмотрено понятие «***ранг матрицы***» и некоторые теоремы о совместности СЛАУ общего вида, позволяющие выяснить вопрос о совместности СЛАУ, не находя её решения, а также найти её решение.

Понятие ранга матрицы. Рассмотрим матрицу

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_m^n \end{pmatrix}, \qquad (3.4.1)$$

где $n \neq m$. Предположим, что число $k \leq \min(n, m)$.

Определение 3.4.1. *Выберем в матрице (3.4.1) произвольно k строк и k столбцов. Определитель k-го порядка, образованный элементами матрицы A, стоящими на пересечении выбранных строк и столбцов, называется её* ***минором порядка k****. Минор, расположенный в первых k строках и в первых k столбцах, называется* ***главным*** *или* ***угловым минором матрицы A****.* •

Определение 3.4.2. *Если некоторая прямоугольная матрица A имеет не только нулевые элементы, то наивысший порядок r отличных от нуля миноров этой матрицы называется* ***рангом матрицы A****. Любой отличный от нуля минор порядка r матрицы A называется* ***базисным минором****, а строки и столбцы матрицы, в которых он расположен, называются* ***базисными.*** •

Ранг матрицы обозначается $r(A)$. По определению полагают, что ***ранг нулевой матрицы равен нулю***.

Пусть в пространстве X^n зафиксирован базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ и задана система из m векторов

$$\begin{aligned}
&\vec{x}_1 = x_1^1\vec{e}_1 + x_1^2\vec{e}_2 + \ldots + x_1^n\vec{e}_n,\\
&\vec{x}_2 = x_2^1\vec{e}_1 + x_2^2\vec{e}_2 + \ldots + x_2^n\vec{e}_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\\
&\vec{x}_m = x_m^1\vec{e}_1 + x_m^2\vec{e}_2 + \ldots + x_m^n\vec{e}_n.
\end{aligned} \tag{3.4.2}$$

Запишем матрицу, столбцы которой составлены из координат векторов (3.4.2):

$$A = \begin{pmatrix} x_1^1 & x_2^1 & \ldots & x_m^1 \\ x_1^2 & x_2^2 & \ldots & x_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ x_1^n & x_2^n & \ldots & x_m^n \end{pmatrix}. \tag{3.4.3}$$

Определение 3.4.3. *Пусть система (3.4.2) векторов* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ *линейно зависима и не все векторы системы нулевые. Если в этой системе можно выделить некоторую линейно независимую подсистему векторов, то эта подсистема называется* ***базой исходной системы***. •

Теорема о базисном миноре. Связь понятия ранга оператора и ранга его матрицы устанавливается при помощи следующей теоремы.

Теорема 3.4.1 (о базисном миноре). *Любые базисные столбцы (строки) прямоугольной матрицы A образуют базу её столбцов (строк).*

Теорема 3.4.1 позволяет свести отыскание базы системы векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ к отысканию базисного минора матрицы (3.4.3).

Доказательство. Для доказательства теоремы нужно показать, что базисные столбцы прямоугольной матрицы *A* образуют линейно независимую систему и любой столбец матрицы линейно выражается через векторы этой системы.

Если бы базисные столбцы матрицы *A* составляли линейно зависимую систему, то один из них линейно выражался бы через остальные базисные столбцы. Но тогда по свойствам определителя базисный минор матрицы можно представить в виде суммы определителей с двумя равными столбцами, следовательно, и сам базисный минор равен нулю. А это противоречит условию теоремы.

Добавим теперь к базисным столбцам любой другой столбец матрицы *A*. Тогда по определению базисного минора все миноры порядка

$r+1$, расположенные на этих столбцах, будут равны нулю. Предположим, что данные столбцы (базисные плюс добавленный столбец) образуют линейно независимую систему. Дополнив эту систему до базиса, получим квадратную матрицу, определитель которой не должен равняться нулю. Но, раскладывая полученный определитель последовательно по элементам дополнительных (до базиса) столбцов, приходим к сумме произведений элементов дополнительных столбцов на миноры, расположенные в исходных $r+1$ столбцах, которые равны нулю. Таким образом, полученный определитель (который не должен равняться нулю) равен нулю. Получили противоречие. Следовательно, система столбцов, состоящая из базисных столбцов плюс любой другой столбец матрицы A, линейно зависима. Поэтому через базисные столбцы можно выразить линейно любой другой столбец матрицы. ••

Следствие из теоремы 3.4.1. *Пусть* $\hat{A}: X^m \to Y^n$ *— некоторый линейный оператор и*

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \subset X^m, \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right\} \subset X^n$$

— фиксированные базисы в пространствах X^m и Y^n. Тогда ранг оператора равен рангу его матрицы

$$r\left(\hat{A}\right) = r(A).$$

Д о к а з а т е л ь с т в о. Для доказательства достаточно вспомнить, что согласно конструкции линейного оператора столбцы его матрицы составлены из координат образов

$$\left\{\hat{A}\vec{a}_1, \hat{A}\vec{a}_2, \ldots, \hat{A}\vec{a}_m\right\} \subset Y^n$$

векторов базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \subset X^m$, причём база столбцов матрицы равна рангу оператора, а по теореме о базисном миноре — рангу его матрицы. ••

Один из методов расчёта ранга матрицы основан на использовании так называемых ***элементарных операций над столбцами (строками) матрицы***, не изменяющих её ранга и приводящих матрицу к ступенчатому виду.

Определение 3.4.4. ***Элементарными операциями*** *над столбцами (строками) прямоугольной матрицы А называются следующие операции:*

1) *перестановка столбцов (строк) матрицы;*

2) *отбрасывание ненулевого общего множителя элементов данного столбца (строки);*

3) *прибавление к данному столбцу (строке) матрицы другого её столбца с произвольным множителем;*

4) *вычёркивание столбца (строки), состоящего (состоящей) из одних нулей;*

5) *вычёркивание столбца (строки), являющегося (являющейся) линейной комбинацией других столбцов (строк).* •

Справедлива следующая теорема.

Теорема 3.4.2. *Элементарные операции над столбцами (строками) некоторой матрицы A не изменяют её ранга.*

Д о к а з а т е л ь с т в о. Доказательство этой теоремы основано на том, что элементарные операции над столбцами матрицы не изменяют размерности линейной оболочки последних. Рассмотрим, например, первую элементарную операцию — перестановку столбцов матрицы. Размерность линейной оболочки столбцов матрицы, которая равна числу базисных столбцов, то есть её рангу, не зависит от порядка их следования. Поэтому и ранг матрицы от порядка следования её столбцов не зависит. Остальные утверждения теоремы доказываются аналогично. ••

Продемонстрируем применение этой теоремы для нахождения ранга матрицы на примере.

Пример 3.4.1. Пусть требуется найти ранг матрицы

$$A = \begin{pmatrix} 0 & -1 & 3 & 0 & 2 \\ 2 & -4 & 1 & 5 & 3 \\ -4 & 5 & 7 & -10 & 0 \\ -2 & 1 & 8 & -5 & 3 \end{pmatrix}.$$

Р е ш е н и е. 1. Чтобы элемент матрицы $a_1^1 \neq 0$, меняем местами первую и вторую строки матрицы:

$$\begin{pmatrix} 0 & -1 & 3 & 0 & 2 \\ 2 & -4 & 1 & 5 & 3 \\ -4 & 5 & 7 & -10 & 0 \\ -2 & 1 & 8 & -5 & 3 \end{pmatrix} \to \begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \\ -4 & 5 & 7 & -10 & 0 \\ -2 & 1 & 8 & -5 & 3 \end{pmatrix}.$$

2. Прибавим первую строку к четвёртой и, умножая её на 2, к третьей:

$$\begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \\ -4 & 5 & 7 & -10 & 0 \\ -2 & 1 & 8 & -5 & 3 \end{pmatrix} \to \begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \\ 0 & -3 & 9 & 0 & 6 \\ 0 & -3 & 9 & 0 & 6 \end{pmatrix}.$$

3. Вычтем из четвёртой строки третью, а из третьей — умноженную на три, вторую:

$$\begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \\ 0 & -3 & 9 & 0 & 6 \\ 0 & -3 & 9 & 0 & 6 \end{pmatrix} \to \begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

4. Вычёркиваем строки, состоящие из одних нулей:

$$\begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & -4 & 1 & 5 & 3 \\ 0 & -1 & 3 & 0 & 2 \end{pmatrix}.$$

5. Полученная матрица имеет ***ступенчатый вид***. Очевидно, что «угловой» минор второго порядка, расположенный в первом и втором столбцах, отличен от нуля. Поэтому ранг данной матрицы равен двум. ⊗

Критерии совместности СЛАУ. Рассмотрим применение понятия ранга матрицы к выяснению условий совместности СЛАУ.

Определение 3.4.5. *Система линейных алгебраических уравнений вида*

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_n^1 x^n = 0, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_n^2 x^n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n x^1 + a_2^n x^2 + \ldots + a_n^n x^n = 0 \end{cases} \qquad (3.4.4)$$

называется ***однородной***. •

Очевидно, что у СЛАУ (3.4.4) всегда существует хотя бы одно решение, например, нулевое. Если кроме нулевого решения у СЛАУ (3.4.4) других решений нет, то такая система называется ***тривиально совместной***. Справедлива следующая теорема.

Теорема 3.4.3. *Если ранг матрицы A однородной СЛАУ (3.4.4) равен числу неизвестных n, то СЛАУ тривиально совместна. СЛАУ (3.4.4) нетривиально совместна в том и только в том случае, если ранг матрицы A меньше числа неизвестных.*

Доказательство. Для доказательства представим СЛАУ в виде

$$x^1\begin{pmatrix} a_1^1 \\ a_1^2 \\ \ldots \\ a_1^n \end{pmatrix} + x^2\begin{pmatrix} a_2^1 \\ a_2^2 \\ \ldots \\ a_2^n \end{pmatrix} + \ldots + x^n\begin{pmatrix} a_n^1 \\ a_n^2 \\ \ldots \\ a_n^n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix}.$$

Из этого представления следует, что существование ненулевого решения у СЛАУ (3.4.4) равносильно линейной зависимости столбцов её матрицы,

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_n^n \end{pmatrix},$$

как вектор-столбцов пространства R^n. По следствию из теоремы 3.4.1 эта линейная зависимость имеет место в том и только в том случае, когда ранг матрицы A меньше числа неизвестных m. ••

Таким образом, необходимым и достаточным условием нетривиальной совместности СЛАУ (3.4.4) является ***равенство нулю определителя её матрицы***, то есть условие $\det A = 0$.

Так как система уравнений (3.4.4) вытекает из условия

$$\hat{A}\vec{x} = \vec{0},$$

то её решением являются, очевидно, координаты векторов, составляющих ядро $k\left(\hat{A}\right)$ линейного оператора $\hat{A}: X^n \to X^n$, имеющего своим представителем в фиксированных базисах матрицу, равную матрице A однородной СЛАУ (3.4.4). База вектор-столбцов матрицы A в этом случае называется ***фундаментальной системой решений*** однородной СЛАУ (3.4.4).

Определение 3.4.6. *Пусть дана СЛАУ общего вида:*

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_m^1 x^m = b^1, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_m^2 x^m = b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n x^1 + a_2^n x^2 + \ldots + a_m^n x^m = b^n. \end{cases} \tag{3.4.5}$$

Расширенной матрицей СЛАУ *(3.4.5) называется матрица*

$$B = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 & b^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 & b^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_m^n & b^n \end{pmatrix}. \bullet \tag{3.4.6}$$

Теорема 3.4.4. *СЛАУ общего вида (3.4.5) совместна в том и только в том случае, если ранг расширенной матрицы (3.4.6) этой СЛАУ равен рангу её основной матрицы.*

Д о к а з а т е л ь с т в о. Пусть СЛАУ (3.4.5) совместна. Тогда, если координаты вектор-столбца

$$|c\rangle = \begin{pmatrix} c^1 \\ c^2 \\ \ldots \\ c^m \end{pmatrix}$$

являются её решением, справедливы тождества:

$$\begin{cases} a_1^1 c^1 + a_2^1 c^2 + \ldots + a_m^1 c^m = b^1, \\ a_1^2 c^1 + a_2^2 c^2 + \ldots + a_m^2 c^m = b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n c^1 + a_2^n c^2 + \ldots + a_m^n c^m = b^n, \end{cases} \Rightarrow$$

$$\Rightarrow c^1 \begin{pmatrix} a_1^1 \\ a_1^2 \\ \ldots \\ a_1^n \end{pmatrix} + c^2 \begin{pmatrix} a_2^1 \\ a_2^2 \\ \ldots \\ a_2^n \end{pmatrix} + \ldots + c^m \begin{pmatrix} a_m^1 \\ a_m^2 \\ \ldots \\ a_m^n \end{pmatrix} = \begin{pmatrix} b^1 \\ b^2 \\ \ldots \\ b^n \end{pmatrix}.$$

Таким образом, последний столбец расширенной матрицы (3.4.6) СЛАУ (3.4.5) является линейной комбинацией остальных столбцов. По теореме 3.4.2 этот столбец можно вычеркнуть из матрицы без изменения её ранга. Но при вычёркивании последнего столбца в матрице (3.4.6) мы как раз и получаем основную матрицу СЛАУ.

Пусть теперь A и B имеют одинаковый ранг. Если r — ранг матрицы A (следовательно, и B), то r базисных столбцов матрицы A будут базисными столбцами и матрицы B. Тогда последний столбец матрицы B является линейной комбинацией базисных столбцов с некоторыми коэффициентами, например, $c^1, c^2, \ldots, c^r$, то есть

$$c^1\begin{pmatrix} a_1^1 \\ a_1^2 \\ \ldots \\ a_1^n \end{pmatrix} + c^2\begin{pmatrix} a_2^1 \\ a_2^2 \\ \ldots \\ a_2^n \end{pmatrix} + \ldots + c^r\begin{pmatrix} a_r^1 \\ a_r^2 \\ \ldots \\ a_r^n \end{pmatrix} + 0\begin{pmatrix} a_{r+1}^1 \\ a_{r+1}^2 \\ \ldots \\ a_{r+1}^n \end{pmatrix} + \ldots + 0\begin{pmatrix} a_m^1 \\ a_m^2 \\ \ldots \\ a_m^n \end{pmatrix} = \begin{pmatrix} b^1 \\ b^2 \\ \ldots \\ b^n \end{pmatrix}.$$

Откуда имеем

$$\begin{cases} a_1^1c^1 + a_2^1c^2 + \ldots + a_r^1c^r + a_{r+1}^1 \cdot 0 + \ldots + a_m^1 \cdot 0 = b^1, \\ a_1^2c^1 + a_2^2c^2 + \ldots + a_r^2c^r + a_{r+1}^2 \cdot 0 + \ldots + a_m^2 \cdot 0 = b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^nc^1 + a_2^nc^2 + \ldots + a_r^nc^r + a_{r+1}^n \cdot 0 + \ldots + a_m^n \cdot 0 = b^n. \end{cases}$$

Таким образом, СЛАУ (3.4.5) удовлетворяется значениями неизвестных

$$x^1 = c^1,\ x^2 = c^2, \ldots, x^r = c^r,\ x^{r+1} = 0, \ldots, x^m = 0$$

и, следовательно, совместна. ••

Решение СЛАУ общего вида. Пусть ранг основной матрицы СЛАУ (3.4.5) равен рангу её расширенной матрицы, то есть $r(A) = r(B)$. Не уменьшая общности рассуждений, предположим, что отличный от нуля минор матрицы порядка r расположен в левом верхнем углу матрицы A. Тогда первые r строк матрицы A по теореме о базисном миноре будут линейно независимыми. Следовательно, первые r уравнений системы также будут линейно независимыми в том смысле, что остальные $n - r$ уравнений будут их линейными комбинациями, то есть ***следствиями*** первых r уравнений СЛАУ. Все решения последних $n - r$ уравнений будут содержаться во множестве решений первых r уравнений. Поэтому достаточно решить только первые r уравнений.

При решении СЛАУ вида (3.4.5) возможны два случая.

1. *Ранг основной матрицы СЛАУ (3.4.5) равен числу неизвестных*. Тогда достаточно решить систему из первых $r = m$ уравнений, например, по формулам Крамера: ***СЛАУ совместна и определённа (имеет единственное решение)***.

2. *Ранг основной матрицы СЛАУ меньше числа неизвестных* $r < m$. В этом случае в левой части первых r уравнений СЛАУ оставляем первые r неизвестных (***базисные неизвестные***), а оставшиеся $m - r$ неизвестных переносим в правые части и объявляем ***свободными неизвестными***, которым можно придавать любые значения, то есть записываем СЛАУ в виде

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_r^1 x^r = b^1 - a_{r+1}^1 x^{r+1} - \ldots - a_m^1 x^m, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_r^2 x^r = b^2 - a_{r+1}^2 x^{r+1} - \ldots - a_m^2 x^m, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^r x^1 + a_2^r x^2 + \ldots + a_r^r x^r = b^r - a_{r+1}^r x^{r+1} - \ldots - a_m^r x^m. \end{cases} \tag{3.4.7}$$

Решая систему уравнений (3.4.7) по формулам Крамера, получаем ***общее решение СЛАУ (3.4.5)***: ***СЛАУ совместна и неопределённа***.

Пример 3.4.2. Решить СЛАУ:

$$\begin{cases} x^1 + 2x^2 + 3x^3 - x^4 = 0, \\ x^1 - x^2 + x^3 + 2x^4 = 4, \\ x^1 + 5x^2 + 5x^3 - 4x^4 = -4, \\ x^1 + 8x^2 + 7x^3 - 7x^4 = -8. \end{cases}$$

Решение. Применяя элементарные преобразования, сводим вычисление ранга расширенной и основной матриц к нахождению ранга матрицы

$$\begin{pmatrix} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \end{pmatrix}.$$

Нетрудно видеть, что $r(A) = r(B) = 2$, и, следовательно, СЛАУ совместна и неопределённа. Главный минор

$$\begin{vmatrix} 1 & 2 \\ 0 & -3 \end{vmatrix} = -3 \neq 0.$$

Принимаем третье и четвёртое неизвестные за свободные неизвестные, полагая $x^3 = a$, $x^4 = b$. Решая первые два уравнения по формулам Крамера, получаем общее решение СЛАУ в виде

$$x^1 = \frac{8}{3} - \frac{5}{3}a - b, x^2 = -\frac{4}{3} - \frac{2}{3}a + b. \otimes$$

3.5. Плоскость и прямая линия в n-мерном аффинном пространстве *)

Определение плоскости в аффинном пространстве. Выше мы ввели определение плоскости в пространстве R^3. Здесь приведён вывод уравнений плоскости и прямой линии в многомерном аффинном пространстве A^n.

Определение 3.5.1. *Пусть A^n — аффинное пространство n измерений. Выбирая некоторую точку $O^* \in A^n$ и линейно независимую систему из m векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$, построим множество всех точек, для которых радиус-вектор $\overrightarrow{O^*M}$ допускает разложение по системе векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ вида*

$$\overrightarrow{O^*M} = t^1 \vec{a}_1 + t^2 \vec{a}_2 + \ldots + t^m \vec{a}_m, \tag{3.5.1}$$

*где коэффициенты разложения $t^1, t^2, \ldots, t^m$ принимают всевозможные числовые значения. Полученное множество точек называется **m-мерной плоскостью** H^m **в n-мерном аффинном пространстве** A^n.* •

Как следует из определения, точка M пробегает всю плоскость H^m и на этом основании называется ***текущей точкой*** плоскости. Векторы

$$\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m \subset H^m$$

в разложении (3.5.1) радиус-вектора $\overrightarrow{O^*M}$ произвольной точки M плоскости H^m называются ***направляющими векторами*** этой плоскости. Отметим, что описанное в определении построение является, в понятном смысле слова, ***сужением процедуры построения аффинной системы координат*** в аффинном пространстве на случай m-мерной плоскости H^m (дальше будем указывать размерность плоскости при помощи верхнего индекса), то есть $\left\{O^*, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ является ***локальным репером*** в H^m, а числа $t^1, t^2, \ldots, t^m$ из разложения (3.5.1) — ***локальными аффинными координатами точек*** в этом репере.

Параметрические и неявные уравнения *m*-мерной плоскости в *n*-мерном аффинном пространстве. Из разложения (3.5.1) следуют уравнения m-мерной плоскости в n-мерном аффинном пространстве A^n, а именно, справедлива следующая важная теорема, на которой основана, по существу, часть аналитической геометрии, посвящённая теории прямых линий и плоскостей, в многомерном аффинном пространстве.

Теорема 3.5.1. *Текущие координаты x^i произвольной точки M данной m-мерной плоскости $H^m \subset A^n$ выражаются линейными функциями m независимых параметров $t^1, t^2, \ldots, t^m$ вида*

$$x^i = x_0^i + a_1^i t^1 + a_2^i t^2 + \ldots + a_m^i t^m \; (i = 1, 2, \ldots, n), \tag{3.5.2}$$

причём ранг матрицы коэффициентов равен m и, следовательно, система столбцов матрицы линейно независима.

Доказательство. Пусть в пространстве A^n введён некоторый репер $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ и M — любая точка плоскости $H^m \subset A^n$. Тогда радиус-вектор точки M имеет вид:

$$\vec{OM} = \vec{OO^*} + \vec{O^*M} = \vec{OO^*} + t^1 \vec{a}_1 + t^2 \vec{a}_2 + \ldots + t^m \vec{a}_m. \tag{3.5.3}$$

Здесь $\vec{OO^*}$ и $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ — фиксированные векторы, а $t^1, t^2, \ldots, t^m$ — независимые переменные (координаты на плоскости H^m). Поэтому выражение (3.5.3) является ***параметрическим уравнением плоскости H^m в векторной форме***.

Раскладывая все векторы из уравнения (3.5.3) по векторам репера $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$

$$\begin{aligned}
&\vec{OM} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^n \vec{e}_n, \\
&\vec{OO^*} = x_0^1 \vec{e}_1 + x_0^2 \vec{e}_2 + \ldots + x_0^n \vec{e}_n, \\
&\vec{a}_1 = a_1^1 \vec{e}_1 + a_1^2 \vec{e}_2 + \ldots + a_1^n \vec{e}_n, \\
&\vec{a}_2 = a_2^1 \vec{e}_1 + a_2^2 \vec{e}_2 + \ldots + a_2^n \vec{e}_n, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\
&\vec{a}_m = a_m^1 \vec{e}_1 + a_m^2 \vec{e}_2 + \ldots + a_m^n \vec{e}_n,
\end{aligned} \tag{3.5.4}$$

подставляя разложения (3.5.4) в (3.5.3) и приравнивая коэффициенты при соответствующих векторах базиса, получим:

$$\begin{cases} x^1 = x_0^1 + a_1^1 t^1 + a_2^1 t^2 + \ldots + a_m^1 t^m, \\ x^2 = x_0^2 + a_1^2 t^1 + a_2^2 t^2 + \ldots + a_m^2 t^m, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ x^n = x_0^n + a_1^n t^1 + a_2^n t^2 + \ldots + a_m^n t^m. \end{cases} \tag{3.5.5}$$

Уравнения (3.5.5) являются ***параметрическими уравнениями плоскости*** H^m в координатной форме, и теорема доказана. Ранг матрицы коэффициентов системы уравнений (3.5.5) максимален и равен m, так как система векторов $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ линейно независима.••

Справедливо и обратное утверждение.

Теорема 3.5.2. *Уравнения (3.5.2) при условии линейной независимости системы столбцов матрицы коэффициентов, то есть при условии максимального ранга* $r = m$ *этой матрицы, всегда определяют в фиксированном репере*

$$\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset A^n$$

некоторую m*-мерную плоскость.*

Из параметрических уравнений плоскости (3.5.2) или (3.5.5) можно получить ***неявные уравнения плоскости*** H^m.

Теорема 3.5.3. *Плоскость* $H^m \subset A^n$ *может быть задана* $n - m$ *независимыми линейными уравнениями между текущими координатами* $x^1, x^2, \ldots, x^n$.

Д о к а з а т е л ь с т в о. Для доказательства нужно из m уравнений (3.5.5) выразить m параметров $t^1, t^2, \ldots, t^m$ и исключить их из оставшихся $n - m$ уравнений. Последнее всегда выполнимо, так как направляющие векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ плоскости H^m образуют линейно независимую систему и, следовательно, ранг матрицы коэффициентов разложения (3.5.3) равен m, то есть максимален. Предположим, что базисный минор располагается в первых m строках матрицы A. Выражая параметры $t^1, t^2, \ldots, t^m$ через координаты $x^1, x^2, \ldots, x^m$ из первых m уравнений и подставляя их в оставшиеся уравнения, получаем следующую систему $n - m$ уравнений:

$$\begin{cases} x^{m+1} = b_1^{m+1} x^1 + b_2^{m+1} x^2 + \ldots + b_m^{m+1} x^m + b^{m+1}, \\ x^{m+2} = b_1^{m+2} x^1 + b_2^{m+2} x^2 + \ldots + b_m^{m+2} x^m + b^{m+2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^n = b_1^n x^1 + b_2^n x^2 + \ldots + b_m^n x^m + b^n. \end{cases} \tag{3.5.6}$$

Здесь коэффициенты b_q^p выражаются некоторым образом через коэффициенты a_j^i (способ выражения нас в данный момент не интересует). ••

Теорема 3.5.1 обобщает теорию плоскостей, изложенную выше для арифметического пространства R^3, на многомерные аффинные пространства и вместе с теоремами 3.5.2, 3.5.3 фактически подводит некоторый итог в исследовании объектов, которые мы назвали плоскостями. В собственно евклидовом пространстве, например в пространстве R^n все построения этого параграфа сохраняют силу, но в качестве векторов «глобального» и «локального» реперов

$$\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset H^m, \left\{O^*, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \subset H^m$$

выбираются канонические реперы в пространстве R^n и на плоскости H^m.

Частные случаи задания *m*-мерной плоскости в *n*-мерном аффинном пространстве. Рассмотрим некоторые частные случаи задания m-мерной плоскости в n-мерном аффинном пространстве.

Пример 3.5.1. 1. Пусть $m = 0$. Тогда вместо уравнений (3.5.5) получаем:

$$x^i = x_0^i \; (i = 1, 2, \ldots, n).$$

Имеем ***точку n-мерного пространства***.

2. Пусть теперь $m = 1$. Тогда вместо уравнений (3.5.5) имеем:

$$x^i = x_0^i + a_1^i t_1 \left(i = 1, 2, \ldots, n\right). \tag{3.5.7}$$

Видим, что текущие координаты являются линейными функциями одного параметра — ***плоскость одномерна****, то есть является* ***прямой линией*** *в n-мерном аффинном пространстве.*

Выпишем неявные уравнения прямой линии в n-мерном пространстве. Из первого уравнения (3.5.7) ($i = 1$) при условии $a_1^1 \neq 0$ получаем:

$$t_1 = \frac{(x^1 - x_0^1)}{a_1^1}.$$

Подстановка во второе уравнение даёт

$$x^2 = b_1^2 x^1 + b^2,$$

где $b_1^2 \equiv \dfrac{a_1^2}{a_1^1}$, $b^2 \equiv x_0^2 - \dfrac{a_1^2}{a_1^1}x_0^1$. И вообще $(\forall i = 2, 3, \ldots, n)$

$$x^i = b_1^i x^1 + b^i,$$

где $b_1^i \equiv \dfrac{a_1^i}{a_1^1}$, $b^i \equiv x_0^i - \dfrac{a_1^i}{a_1^1}x_0^i$.

Таким образом, имеем:

$$\begin{cases} x^2 = b_1^2 x^1 + b^2, \\ x^3 = b_1^3 x^1 + b^3, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x^n = b_1^n x^1 + b^n. \end{cases} \tag{3.5.8}$$

В случае $n = 3$

$$\begin{cases} x^1 = x_0^1 + t_1 a_1^1, \\ x^2 = x_0^2 + t_1 a_1^2, \\ x^3 = x_0^3 + t_1 a_1^3 \end{cases} \tag{3.5.9}$$

и неявные уравнения принимают вид:

$$\begin{cases} x^2 = b_1^2 x^1 + b^2, \\ x^3 = b_1^3 x^1 + b^3. \end{cases} \tag{3.5.10}$$

Уравнения (3.5.9) — параметрические уравнения прямой линии в трёхмерном аффинном пространстве A^3, а уравнения (3.5.10) — уравнения прямой линии в неявной форме в трёхмерном аффинном пространстве A^3.

3. Пусть $m = 2$. Тогда вместо уравнений (3.5.5) имеем:

$$x^i = x_0^i + a_1^i t_1 + a_2^i t_2 \; (i = 1, 2, \ldots, n). \tag{3.5.11}$$

Текущие координаты являются линейными функциями двух параметров. Получили двумерную плоскость в n-мерном аффинном пространстве A^n. Для представления уравнений в неявном виде нужно выразить из первых двух уравнений параметры t_1 и t_2, а затем подставить их в оставшиеся уравнения. В результате (решая первые два уравнения из системы (3.5.11) методом Крамера) получим для $i = 3, 4, \ldots, n$:

$$x^i = b_1^i x^1 + b_2^i x^2 + b^i. \tag{3.5.12}$$

Здесь коэффициенты b_q^p выражаются через коэффициенты a_i^j так:

$$b_1^i \equiv \frac{1}{\det A}\left(a_1^i a_2^2 - a_2^i a_1^2\right);\ b_2^i \equiv \frac{1}{\det A}\left(a_2^i a_1^1 - a_1^i a_2^1\right);$$

$$b^i \equiv x_0^i + \frac{1}{\det A}\left[x_0^1(a_2^i a_1^2 - a_1^i a_2^2) + x_0^2(a_1^i a_2^1 - a_2^i a_1^1)\right].$$

Уравнения (3.5.12) — неявная форма уравнений двумерной плоскости в n-мерном аффинном пространстве A^n. ⊗

Определение 3.5.2. *Плоскость H^m размерности $m = n-1$ в n-мерном аффинном пространстве A^n называется **гиперплоскостью**.* •

В случае гиперплоскости в системе (3.5.6) остаётся лишь одно уравнение. Поэтому ***гиперплоскость может быть задана одним (неявным) линейным уравнением относительно текущих координат***.

Пример 3.5.2. Рассмотрим снова изученный выше случай $n = 3$. Гиперплоскость имеет размерность $m = 2$ и определяется системой уравнений вида:

$$\begin{cases} x^1 = x_0^1 + t_1 a_1^1 + t_2 a_2^1, \\ x^2 = x_0^2 + t_1 a_1^2 + t_2 a_2^2, \\ x^3 = x_0^3 + t_1 a_1^3 + t_2 a_2^3. \end{cases}$$

Находя t_1 и t_2 из первых двух уравнений по формулам Крамера и подставляя результат в третье уравнение, получаем неявное уравнение гиперплоскости

$$x^3 = b_1^3 x^1 + b_2^3 x^2 + b^3,$$

где

$$b_1^3 \equiv -\frac{1}{\det A}\cdot\det\begin{pmatrix} a_1^2 & a_2^2 \\ a_1^3 & a_2^3 \end{pmatrix},\ b_2^3 \equiv -\frac{1}{\det A}\cdot\det\begin{pmatrix} a_1^1 & a_2^1 \\ a_1^3 & a_2^3 \end{pmatrix},$$

$$b^3 \equiv x_0^3 + \frac{1}{\det A}\left[x_0^1\cdot\det\begin{pmatrix} a_1^2 & a_2^2 \\ a_1^3 & a_2^3 \end{pmatrix} + x_0^2\cdot\det\begin{pmatrix} a_1^1 & a_2^1 \\ a_1^3 & a_2^3 \end{pmatrix}\right]. \otimes$$

О плоскости H^m в n-мерном аффинном пространстве A^n, заданной векторным уравнением (3.5.3) или координатными уравнениями (3.5.5), говорят, что она ***получена параллельным сдвигом*** проходящей через начало системы координат плоскости, определённой векторным уравнением

$$\vec{OM} = t_1\vec{a}_1 + t_2\vec{a}_2 + \ldots + t_m\vec{a}_m$$

или координатными уравнениями

$$x^i = x_0^i + a_1^i t_1 + a_2^i t_2 + \ldots + a_m^i t_m \ (i = 1, 2, \ldots, n),$$

на вектор $\overrightarrow{OO^*} = x_0^1 \vec{e}_1 + x_0^2 \vec{e}_2 + \ldots + x_0^3 \vec{e}_n$ — ***вектор сдвига***.

Тривиальным является случай задания плоскости при $m = n$, когда плоскость H^m совпадает с самим пространством A^n (говорят, что плоскость заполняет пространство).

Система всех плоскостей m измерений в n-мерном аффинном пространстве, проходящих через одну точку O, называется ***связкой плоскостей***. В случае связки плоскостей для задания конкретной плоскости нужно знать её направляющие векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$.

Глава 4. Некоторые специальные виды линейных операторов

4.1. Собственные подпространства и характеристический многочлен линейного оператора

Собственные векторы линейного оператора и их свойства. Пусть линейный оператор $\hat{T}$ действует в пространстве X^n, то есть является преобразованием

$$\hat{T} : X^n \to \hat{T}\left(X^n\right).$$

Это означает, что каждому вектору $\vec{x} \in X^n$ ставится в соответствие некоторый вектор $\vec{y} = \hat{T}\vec{x} \in X^n$. Может реализоваться ситуация, когда образ $\vec{y}$ будет ***коллинеарен*** прообразу $\vec{x}$.

Определение 4.1.1. *Пусть* $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ *— линейный оператор. Тогда, если для вектора* $\vec{x} \in X^n \left(\vec{x} \neq \vec{0}\right)$ *и некоторого числа* μ *выполняется соотношение*

$$\hat{T}\vec{x} = \mu \cdot \vec{x}, \tag{4.1.1}$$

то вектор $\vec{x}$ *называется* ***собственным вектором*** *оператора* $\hat{T}$*, соответствующим* ***собственному значению*** μ. •

Отметим, что нулевой вектор по определению ***не является собственным вектором никакого линейного оператора***.

Лемма 4.1.1. *Пусть* $\hat{T} : X^n \to \hat{T}\left(X^n\right)$*. Тогда, если* $\vec{x}_1, \vec{x}_2 \in X^n$ *— два собственных вектора оператора* $\hat{T}$*, соответствующих одному и тому же* *собственному значению* μ*, а* $\alpha_1 \neq 0$*,* $\alpha_2 \neq 0$ *— некоторые числа, то линейная комбинация* $\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2$ *тоже будет собственным вектором, соответствующим тому же собственному значению* μ*.*

Д о к а з а т е л ь с т в о. Действительно,

$$\hat{T}\left(\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2\right) = \alpha_1 \cdot \mu \vec{x}_1 + \alpha_2 \cdot \mu \vec{x}_2 = \mu\left(\alpha_1 \cdot \vec{x}_1 + \alpha_2 \cdot \vec{x}_2\right). \bullet\bullet$$

Итак, линейные комбинации собственных векторов оператора, соответствующих одному и тому же собственному значению, также являются собственными векторами. Ввиду того, что множество собственных векторов не содержит нулевой вектор, оно не является подпространством.

Определение 4.1.2. *Подпространство* $X_\mu \subset X^n$*, полученное добавлением вектора* $\vec{0}$ *к множеству всех собственных векторов оператора*

$$\hat{T} : X^n \to \hat{T}\left(X^n\right),$$

соответствующих одному и тому же собственному значению μ*, называется* ***собственным подпространством*** *этого оператора.* •

Из определений 4.1.1 и 4.1.2 следует, что ***собственное подпространство линейного оператора*** $\hat{T}$***, соответствующее собственному значению*** μ***, совпадает с ядром оператора*** $\hat{T} - \mu \cdot \hat{I}$, действительно:

$$\hat{T}\vec{x} = \mu \cdot \vec{x} \Rightarrow \hat{T}\vec{x} - \mu \cdot \vec{x} = \vec{0} \Rightarrow \left(\hat{T} - \mu \cdot \hat{I}\right)\vec{x} = \vec{0}.$$

Теперь по определению ядра оператора видим, что вектор $\vec{x}$ является элементом ядра оператора $\hat{T} - \mu \cdot \hat{I}$.

Теорема 4.1.1. *Система* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ *собственных векторов линейного оператора* $\hat{T} : X^n \to \hat{T}\left(X^n\right)$*, соответствующих попарно различным собственным значениям* $\mu_1, \mu_2, \ldots, \mu_m$*, линейно независима.*

Д о к а з а т е л ь с т в о. Докажем теорему по индукции. Собственные векторы по определению ненулевые. Следовательно, теорема верна для системы из одного собственного вектора, то есть при $m = 1$.

Предположим, что система

$$\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{m-1}\right\},$$

состоящая из $m-1$ собственных векторов, соответствующих различным собственным значениям $\mu_1, \mu_2, \ldots, \mu_{m-1}$, линейно независима, но система из m собственных векторов

$$\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{m-1}, \vec{x}_m\right\},$$

соответствующих различным собственным значениям $\mu_1, \mu_2, \ldots, \mu_m$, уже линейно зависима. Тогда, по определению линейной зависимости, найдутся такие не равные одновременно нулю числа $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}, \alpha_m$, что значением линейной комбинации векторов системы с коэффициентами $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}, \alpha_m$, среди которых не все равны нулю одновременно, будет нуль-вектор:

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \ldots + \alpha_{m-1} \vec{x}_{m-1} + \alpha_m \vec{x}_m = \vec{0}. \qquad (4.1.2)$$

Действуя оператором $\hat{T}$ на обе части равенства (4.1.2), получаем:

$$\alpha_1\mu_1 \vec{x}_1 + \ldots + \alpha_{m-1}\mu_{m-1} \vec{x}_{m-1} + \alpha_m\mu_m \vec{x}_m = \vec{0}. \qquad (4.1.3)$$

Умножим теперь обе части равенства (4.1.2) на собственное значение μ_m и найдём разности левых и правых частей полученного равенства и равенства (4.1.3), в результате получим:

$$\alpha_1(\mu_1 - \mu_m)\vec{x}_1 + \alpha_2(\mu_2 - \mu_m)\vec{x}_2 + \ldots + \alpha_{m-1}(\mu_{m-1} - \mu_m)\vec{x}_{m-1} = \vec{0}.$$

По предположению индукции система векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{m-1}\right\}$ линейно независима, следовательно,

$$(\forall k = 1, 2, \ldots, m-1)\ \alpha_k(\mu_k - \mu_m) = 0.$$

В силу предположенной линейной зависимости системы векторов

$$\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{m-1}, \vec{x}_m\right\}$$

можно положить, например, что $\alpha_1 \neq 0$. Но тогда, так как μ_k есть элемент числового поля, должно выполняться условие $\mu_1 = \mu_m$, которое противоречит условию теоремы. Следовательно, система векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ линейно независима. ••

Следствие из теоремы 4.1.1. *Любой из линейных операторов*

$$\hat{T} : X^n \to \hat{T}\left(X^n\right),$$

действующих в n-мерном пространстве X^n, не может иметь более чем n попарно различных собственных значений.

В этом случае по теореме 4.1.1 оператор имеет не более чем n собственных векторов, образующих линейно независимую систему. Поэтому мы можем составить ***базис пространства X^n целиком из собственных векторов оператора***.

Определение 4.1.3. *Линейный оператор $\hat{T}$, действующий в n-мерном пространстве X^n и имеющий n попарно различных собственных значений, называется **оператором простой структуры**.* •

Лемма 4.1.2. *Оператор $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ является оператором простой структуры в том и только в том случае, если в базисе собственных векторов он имеет диагональную матрицу.*

Доказательство. *Докажем необходимость.* Пусть $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ — линейно независимая система собственных векторов оператора простой структуры $\hat{T}$. Выберем их за базис пространства X^n, тогда

$$\hat{T}\vec{a}_1 = \mu_1 \vec{a}_1 = \mu_1 \vec{a}_1 + 0\vec{a}_2 + \ldots + 0\vec{a}_n,$$

$$\hat{T}\vec{a}_2 = \mu_2 \vec{a}_2 = 0\vec{a}_1 + \mu_2 \vec{a}_2 + \ldots + 0\vec{a}_n,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\hat{T}\vec{a}_n = \mu_n \vec{a}_n = 0\vec{a}_1 + 0\vec{a}_2 + \ldots + \mu_n \vec{a}_n.$$

Записывая матрицу оператора по столбцам, получаем:

$$T = \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu_n \end{pmatrix}. \tag{4.1.4}$$

Докажем достаточность. Пусть в некотором базисе

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$$

матрица линейного оператора $\hat{T}: X^n \to \hat{T}\left(X^n\right)$ имеет диагональный вид (4.1.4), причём диагональные элементы матрицы попарно различны. Тогда, вспоминая, что столбцы матрицы оператора составлены из координат разложений образов базисных векторов по базису пространства X^n, получаем

$$\hat{T}\vec{a}_1 = \mu_1 \cdot \vec{a}_1 + 0 \cdot \vec{a}_2 + \ldots + 0 \cdot \vec{a}_n = \mu_1 \vec{a}_1,$$

$$\hat{T}\vec{a}_2 = 0 \cdot \vec{a}_1 + \mu_2 \cdot \vec{a}_2 + \ldots + 0 \cdot \vec{a}_n = \mu_2 \vec{a}_2,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\hat{T}\vec{a}_n = 0_1 \cdot \vec{a}_1 + 0 \cdot \vec{a}_2 + \ldots + \mu_n \cdot \vec{a}_n = \mu_n \vec{a}_n. \bullet\bullet$$

Очевидно, что в базисе из собственных векторов действие оператора простой структуры сводится к ***растяжению координат вектора***.

Пример 4.1.1. Пусть в базисе собственных векторов $\left\{\vec{a}_1, \vec{a}_2, \vec{a}_3\right\}$ задана матрица оператора $\hat{T}$ и вектор

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \vec{x} = 3\vec{a}_1 + 11\vec{a}_2 + 7\vec{a}_3.$$

Найти образ вектора $\vec{x}$ при действии оператора $\hat{T}$.

Р е ш е н и е. Образ имеет координаты:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 3 \\ 11 \\ 7 \end{pmatrix} = \begin{pmatrix} 3 \\ 22 \\ 21 \end{pmatrix}.$$

Очевидно, что вторая и третья координаты вектора «растянуты», соответственно, в 2 и 3 раза. $\otimes$

Характеристический многочлен линейного оператора и его свойства. Запишем определение собственного вектора оператора (4.1.1) в матричной форме, зафиксировав некоторый базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ пространства X^n:

$$\begin{pmatrix} t_1^1 & t_2^1 & \ldots & t_n^1 \\ t_1^2 & t_2^2 & \ldots & t_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ t_1^n & t_2^n & \ldots & t_n^n \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^n \end{pmatrix} = \mu \cdot \begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^n \end{pmatrix}.$$

После умножения вектор-столбца на матрицу получаем СЛАУ

$$\begin{cases} t_1^1 x^1 + t_2^1 x^2 + \ldots + t_n^1 x^n = \mu \cdot x^1, \\ t_1^2 x^1 + t_2^2 x^2 + \ldots + t_n^2 x^n = \mu \cdot x^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ t_1^n x^1 + t_2^n x^2 + \ldots + t_n^n x^n = \mu \cdot x^n. \end{cases} \tag{4.1.5}$$

СЛАУ (4.1.5) перепишем в виде

$$\begin{cases} \left(t_1^1 - \mu\right) x^1 + t_2^1 x^2 + \ldots + t_n^1 x^n = 0, \\ t_1^2 x^1 + \left(t_2^2 - \mu\right) x^2 + \ldots + t_n^2 x^n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ t_1^n x^1 + t_2^n x^2 + \ldots + \left(t_n^n - \mu\right) x^n = 0. \end{cases} \tag{4.1.6}$$

СЛАУ (4.1.6) — однородная и, следовательно, нетривиально разрешима в том и только в том случае, если ранг её основной матрицы меньше числа неизвестных. Поэтому должно выполняться условие

$$T(\mu) \overset{def}{=} \det\left(t_j^i - \mu\delta_j^i\right) = \begin{vmatrix} t_1^1 - \mu & t_2^1 & \ldots & t_n^1 \\ t_1^2 & t_2^2 - \mu & \ldots & t_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ t_1^n & t_2^n & \ldots & t_n^n - \mu \end{vmatrix} = 0, \tag{4.1.7}$$

где $\left(\delta_j^i\right)$ — матрица единичного оператора. Раскрывая определитель (4.1.7) по элементам какого-либо столбца (строки), легко видеть, что $T(\mu)$ является ***многочленом относительно*** μ.

Определение 4.1.4. *Пусть* $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ *некоторый линейный оператор. Тогда многочлен*

$$T(\mu) = \det\left(t_j^i - \mu\delta_j^i\right) \equiv \det(T - \mu I)$$

называется ***характеристическим многочленом*** *оператора* $\hat{T}$. •

Теорема 4.1.2. *Характеристический многочлен* $T(\mu)$ *линейного оператора*

$$\hat{T} : X^n \to \hat{T}\left(X^n\right)$$

не зависит от выбора базиса.

Д о к а з а т е л ь с т в о. При переходе от старого базиса

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \in X^n$$

к новому базису

$$\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\} \in X^n$$

матрица оператора $\hat{T}$ преобразуется по формуле

$$T_{p'}^{q'} = \sum_{j=1}^{n} \sum_{i=1}^{n} A_j^{q'} T_i^j A_{p'}^i,$$

где $A_{p'}^i$ — транспонированная матрица перехода от старого базиса к новому базису, а $A_j^{q'}$ — транспонированная матрица перехода от нового базиса к старому. Теперь имеем:

$$\det\left(T_{p'}^{q'} - \mu I_{p'}^{q'}\right) = \det\left(\sum_{j=1}^{n} \sum_{i=1}^{n} A_j^{q'} T_i^j A_{p'}^i - \mu \sum_{j=1}^{n} \sum_{i=1}^{n} A_j^{q'} I_i^j A_{p'}^i\right) =$$

$$= \det\left(\sum_{j=1}^{n} \sum_{i=1}^{n} A_j^{q'} \left(T_i^j - \mu I_i^j\right) A_{p'}^i\right).$$

Отмечая, что для единичной матрицы справедливо свойство $I_{i'}^{j'} = I_i^j$, учитывая следующую из свойств определителей формулу

$$\det A = \frac{1}{\det A^{-1}}$$

для взаимно обратных матриц $A_{p'}^i$ и $A_j^{q'}$, получаем:

$$\det\left(T_{p'}^{q'} - \mu \cdot I_{p'}^{q'}\right) = \det\left(A_j^{q'} \left(T_i^j - \mu \cdot I_i^j\right) A_{p'}^i\right) =$$

$$= \det\left(A_j^{q'}\right) \cdot \det\left(T_i^j - \mu \cdot I_i^j\right) \cdot \det\left(A_{p'}^i\right) = \det\left(T_i^j - \mu \cdot I_i^j\right).$$

Так как базисы выбраны произвольно, характеристический многочлен действительно не зависит от выбора базиса. ••

Отметим, что доказательство теоремы можно провести, используя более простые матричные обозначения. Действительно, матрица оператора при переходе к новому базису преобразуется по формуле

$$T' = \left(A^{-1}\right)^T TA^T.$$

Для матрицы оператора $\hat{T} - \mu \cdot \hat{I}$ аналогично имеем:

$$T' - \mu \cdot I' = \left(A^{-1}\right)^T \left(T - \mu \cdot I\right) A^T.$$

Далее, для определителя матрицы оператора $\hat{T} - \mu \cdot \hat{I}$ в новом базисе получаем

$$\det\left(T' - \mu \cdot I'\right) = \det\left(\left(A^{-1}\right)^T \left(T - \mu \cdot I\right) A^T\right) =$$

$$= \det\left(A^{-1}\right)^T \cdot \det\left(T - \mu \cdot I\right) \cdot \det A^T = \det\left(T - \mu \cdot I\right),$$

что и доказывает утверждение теоремы.

Пример 4.1.2. В каноническом базисе $\left\{\vec{e}_1, \vec{e}_2\right\} \subset R^2$ оператор $\hat{A}$ задан матрицей

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}.$$

Найти собственные значения и собственные векторы оператора.

Р е ш е н и е. 1) Составляем характеристическое уравнение:

$$\det\left(A^i_j - \mu \cdot \delta^i_j\right) = 0 \Rightarrow \mu^2 - 7 \cdot \mu + 10 = 0.$$

Откуда $\mu_1 = 5$, $\mu_2 = 2$.

2) Составляем систему для нахождения собственного вектора, соответствующего собственному значению $\mu_1 = 5$:

$$\begin{pmatrix} -2 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1^1 \\ x_1^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Её решение $x_1^2 = c$, $x_1^1 = c/2$, то есть $\left|x_1\right\rangle = c\left|a_1\right\rangle$, где

$$\left|a_1\right\rangle = \begin{pmatrix} 1/2 \\ 1 \end{pmatrix}, c \in R^1.$$

Аналогично для $\mu_2 = 2$ находим $|x_2\rangle = c|a_2\rangle$, где

$$|a_2\rangle = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Оператор имеет два одномерных собственных подпространства $L\left\{\vec{a}_1\right\}$ и $L\left\{\vec{a}_2\right\}$, где

$$\vec{a}_1 = \frac{1}{2}\vec{e}_1 + \vec{e}_2,\ \vec{a}_2 = -\vec{e}_1 + \vec{e}_2 \otimes$$

Понятие спектра линейного оператора. Итак, все собственные значения линейного оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ являются корнями его характеристического многочлена. Если пространство X^n является комплексным линейным пространством, то по следствию 2 из основной теоремы алгебры многочленов (теорема 6.1.7) характеристический многочлен $T(\mu)$ оператора $\hat{T}$ имеет ровно *n* корней, если каждый корень считать столько раз, какова его кратность.

Определение 4.1.5. *Множество $S_\mu\left(\hat{T}\right)$ всех собственных значений линейного оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ называется его* ***спектром***. •

Спектр в определённом смысле слова ***характеризует данный линейный оператор***.

Действительно, можно показать (рекомендуется сделать это самостоятельно с помощью теоремы Виета), что для характеристического многочлена $T(\mu)$ линейного оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ имеют место следующие представления:

$$T(\mu) = (-1)^n\left[\mu^n - \left(t_1^1 + t_2^2 + \ldots + t_n^n\right)\mu^{n-1} + \ldots\right];$$

$$T(\mu) = \left[\mu^n - \left(\mu_1 + \mu_2 + \ldots + \mu_n\right)\mu^{n-1} + \ldots\right]. \qquad (4.1.8)$$

Здесь $\mu_1, \mu_2, \ldots, \mu_n$ — корни многочлена $T(\mu)$, а величина

$$Sp\,T \overset{def}{=} t_1^1 + t_2^2 + \ldots + t_n^n = \sum_{k=1}^{n} t_k^k$$

называется ***следом матрицы оператора*** $\hat{T} : X^n \to \hat{T}\left(X^n\right)$. След матрицы оператора является одним из коэффициентов его характеристического многочлена и, следовательно, согласно теореме 4.1.2 не зависит от выбора базиса. Поэтому след матрицы T оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ называется также ***следом оператора*** $\hat{T}$.

Из сравнения представлений (4.1.8) вытекает следующая теорема.

Теорема 4.1.3. *Для суммы всех собственных значений линейного оператора* $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ *справедлива формула:*

$$\sum_{k=1}^{n} \mu_k = (-1)^n \, Sp\,\hat{T}.$$

Справедлива также следующая теорема.

Теорема 4.1.4. *След суммы и композиции линейных операторов*

$$\hat{A} : X^n \to \hat{A}\left(X^n\right),\ \hat{B} : X^n \to \hat{B}\left(X^n\right)$$

обладает следующими свойствами:

$$1)\ Sp\left(\hat{A} + \hat{B}\right) = Sp\,\hat{A} + Sp\,\hat{B};\ 2)\ Sp\left(\hat{A}\,\hat{B}\right) = Sp\left(\hat{B}\,\hat{A}\right).$$

Д о к а з а т е л ь с т в о. Первое свойство следа оператора очевидно. Поэтому приведём доказательство только второго свойства.

Введём обозначения:

$$A = \left(a_i^k\right),\ B = \left(b_i^k\right)\ D = AB = \left(d_j^k\right),\ C = BA = \left(c_j^k\right).$$

Тогда, используя формулу для элемента произведения матриц

$$c_j^k = \sum_{i=1}^{n} b_i^k a_j^i,$$

получаем:

$$Sp(AB) = Sp\,D = \sum_{k=1}^{n} d_k^k = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_i^k b_k^i\right) =$$

$$= \sum_{i=1}^{n}\left(\sum_{k=1}^{n} b_k^i a_i^k\right) = \sum_{i=1}^{n} c_i^i = Sp\,C = Sp(BA). \bullet\bullet$$

Следствие из теоремы 4.1.4. *Для любых операторов*

$$\hat{A} : X^n \to \hat{A}\left(X^n\right),\ \hat{B} : X^n \to \hat{B}\left(X^n\right),$$

где $\hat{A}$ — невырожденный оператор, справедливо равенство:

$$Sp\left(\hat{A}^{-1}\left(\hat{B}\hat{A}\right)\right)=Sp\,\hat{B}.$$

Д о к а з а т е л ь с т в о. По теореме 4.1.4 получаем:

$$Sp\left(A^{-1}(BA)\right)=Sp\left((BA)A^{-1}\right)=Sp\left(B\left(AA^{-1}\right)\right)=SpB.\bullet\bullet$$

Определение 4.1.6. *Пусть* $\hat{T}:X^n\to\hat{T}\left(X^n\right)$ *и*

$$F_m(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m\in C[x]$$

— некоторый многочлен. Тогда оператор

$$F_m\left(\hat{T}\right)=a_0\hat{I}+a_1\hat{T}+a_2\hat{T}^2+\ldots+a_m\hat{T}^m \tag{4.1.9}$$

называется ***многочленом от оператора*** $\hat{T}$ ***или операторным многочленом***. •

Так как композиция (степень) и сумма линейных операторов являются линейными операторами, ***операторный многочлен*** $F_m\left(\hat{T}\right)$ ***сам является линейным оператором***.

Построим множество всех операторных многочленов $F_m\left(\hat{T}\right)$ для фиксированного оператора $\hat{T}$. Множество всех многочленов является коммутативным кольцом. Поэтому ***множество всех операторных многочленов также является коммутативным кольцом***. В частности, справедлива важная формула

$$F_m\left(\hat{T}\right)\hat{T}=\hat{T}\,F_m\left(\hat{T}\right). \tag{4.1.10}$$

Лемма 4.1.3. *Пусть* $\vec{x}$ *— собственный вектор линейного оператора*

$$\hat{T}:X^n\to\hat{T}\left(X^n\right),$$

соответствующий собственному значению μ. *Тогда* $\vec{x}$ *является собственным вектором операторного многочлена* $F_m\left(\hat{T}\right)$*, соответствующим собственному значению*

$$a_0+a_1\mu+a_2\mu^2+\ldots+a_m\mu^m.$$

Д о к а з а т е л ь с т в о. Так как $\hat{T}\vec{x} = \mu \cdot \vec{x}$, то

$$F_m\left(\hat{T}\right)\vec{x} = \left(a_0\hat{I} + a_1\hat{T} + a_2\hat{T}^2 + \ldots + a_m\hat{T}^m\right)\vec{x} =$$

$$= a_0\hat{I}\vec{x} + a_1\hat{T}\vec{x} + a_2\hat{T}^2\vec{x} + \ldots + a_m\hat{T}^m\vec{x} =$$

$$= a_0\vec{x} + a_1\mu\vec{x} + a_2\mu^2\vec{x} + \ldots + a_m\mu^m\vec{x} =$$

$$= \left(a_0 + a_1\mu + a_2\mu^2 + \ldots + a_m\mu^m\right)\vec{x}. \bullet\bullet$$

Лемма 4.1.4. *Пусть $\vec{x}$ — собственный вектор невырожденного линейного оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$, соответствующий собственному значению μ, то есть*

$$\hat{T}\vec{x} = \mu \cdot \vec{x}.$$

Тогда вектор $\vec{x}$ является собственным вектором оператора $\hat{T}^{-1}$, соответствующим собственному значению μ^{-1}.

Д о к а з а т е л ь с т в о. Действуя на обе части равенства

$$\hat{T}\vec{x} = \mu \cdot \vec{x}$$

оператором $\hat{T}^{-1}$, получаем

$$\hat{T}^{-1}\left(\hat{T}\vec{x}\right) = \hat{T}^{-1}\left(\mu\vec{x}\right) \Rightarrow \vec{x} = \mu\hat{T}^{-1}\vec{x} \Rightarrow \hat{T}^{-1}\vec{x} = \mu^{-1}\vec{x}. \bullet\bullet$$

Из приведённых результатов видно, что ***действиям с линейными операторами соответствуют действия с их собственными значениями***. Поэтому набор собственных значений оператора действительно характеризует этот оператор.

Инвариантные подпространства линейного оператора. По определению 4.1.1 для собственных векторов и собственных значений линейного оператора имеем

$$\hat{T}\vec{x} = \mu \cdot \vec{x},$$

откуда следует, что как прообраз $\vec{x}$, так и образ $\mu \cdot \vec{x}$ принадлежат одному и тому же множеству векторов, образующему (при добавлении нулевого вектора) некоторое подпространство $X_{\mu} \subset X^{n}$. Следовательно, образы собственных векторов принадлежат тому же собственному подпространству оператора $\hat{T}$. Эта ситуация возможна и с векторами, входящими в подпространства, которые не являются собственными подпространствами.

Определение 4.1.7. *Подпространство* $X_{inv} \subset X^{n}$ *называется* ***инвариантным подпространством оператора*** $\hat{T} : X^{n} \to \hat{T}\left(X^{n}\right)$, *если*

$$\left(\forall \vec{x} \in X_{inv}\right) \vec{y} = \hat{T}\vec{x} \in X_{inv}. \bullet$$

Инвариантные подпространства оператора обладают некоторыми специальными свойствами.

Лемма 4.1.5. *Пересечение и сумма инвариантных подпространств линейного оператора* $\hat{T}$ *снова являются инвариантными подпространствами.*

Доказательство. Пусть $X_{inv}^{(1)}$, $X_{inv}^{(2)}$ — инвариантные подпространства оператора $\hat{T}$ и

$$\vec{x} \in X_{inv}^{(1)} \bigcap X_{inv}^{(2)}$$

— их пересечение. Тогда по определению пересечения подпространств $\vec{x} \in X_{inv}^{(1)}$ и $\vec{x} \in X_{inv}^{(2)}$, а по определению инвариантного подпространства

$$\hat{T}\vec{x} \in X_{inv}^{(1)} \wedge \hat{T}\vec{x} \in X_{inv}^{(2)}.$$

Поэтому снова по определению пересечения подпространств

$$\hat{T}\vec{x} \in X_{inv}^{(1)} \bigcap X_{inv}^{(2)}.$$

Пусть теперь

$$\vec{x} \in X_{inv}^{(1)} + X_{inv}^{(2)}.$$

Тогда по определению суммы подпространств $\vec{x} = \vec{x}_1 + \vec{x}_2$, где

$$\vec{x}_1 \in X_{inv}^{(1)}, \vec{x}_2 \in X_{inv}^{(2)},$$

откуда получаем

$$\hat{T}\vec{x} = \hat{T}\vec{x}_1 + \hat{T}\vec{x}_2 \in X_{ivn}^{(1)} + X_{inv}^{(2)},$$

так как $\hat{T}\vec{x}_1 \in X_{ivn}^{(1)}$ и $\hat{T}\vec{x}_2 \in X_{inv}^{(2)}$. ••

Теорема 4.1.5. *Если оператор* $\hat{T}$ *невырожденный, то его инвариантное подпространство* X_{inv}^m *является инвариантным подпространством и для обратного оператора* $\hat{T}^{-1}$.

Доказательство. Пусть $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \subset X_{inv}^m$ — базис в X_{inv}^m, тогда каждый из векторов $\hat{T}\vec{a}_1, \hat{T}\vec{a}_2, \ldots, \hat{T}\vec{a}_m$ принадлежит тому же инвариантному подпространству, а система $\left\{\hat{T}\vec{a}_1, \hat{T}\vec{a}_2, \ldots, \hat{T}\vec{a}_m\right\}$ образует в X_{inv}^m другой базис. Следовательно,

$$\left(\forall \vec{x} \in X_{inv}^m\right) \vec{x} = x^1 \hat{T}\vec{a}_1 + x^2 \hat{T}\vec{a}_2 + \ldots + x^m \hat{T}\vec{a}_m.$$

Поэтому вектор

$$\hat{T}^{-1}\vec{x} = x^1 \hat{T}^{-1}\hat{T}\vec{a}_1 + x^2 \hat{T}^{-1}\hat{T}\vec{a}_2 + \ldots + x^m \hat{T}^{-1}\hat{T}\vec{a}_m =$$

$$= x^1 \vec{a}_1 + x^2 \vec{a}_2 + \ldots + x^m \vec{a}_m \in X_{inv}^m,$$

так как он разложен по базису X_{inv}. ••

Пусть оператор $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ имеет инвариантное подпространство $X_{inv}^m \subset X^n$. Найдём матрицу оператора, для чего выберем в пространстве X^n базис

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m, \vec{a}_{m+1}, \ldots, \vec{a}_n\right\} \subset X^n$$

так, чтобы первые его m векторов лежали в подпространстве X_{inv}^m

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \subset X_{inv}^m,$$

и найдём образы базисных векторов. Образы первых m базисных векторов

$$\left\{\hat{T}\vec{a}_1, \hat{T}\vec{a}_2, \ldots, \hat{T}\vec{a}_m\right\} \subset X_{inv}^m.$$

Разложим их по базису $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ подпространства X_{inv}^m, добавляя нулевые слагаемые, содержащие оставшиеся базисные векторы:

$$\hat{T}\vec{a}_1 = t_1^1\vec{a}_1 + t_1^2\vec{a}_2 + \ldots + t_1^m\vec{a}_m + 0\vec{a}_{m+1} + \ldots + 0\vec{a}_n,$$

$$\hat{T}\vec{a}_2 = t_2^1\vec{a}_1 + t_2^2\vec{a}_2 + \ldots + t_2^m\vec{a}_m + 0\vec{a}_{m+1} + \ldots + 0\vec{a}_n,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\hat{T}\vec{a}_m = t_m^1\vec{a}_1 + t_m^2\vec{a}_2 + \ldots + t_m^m\vec{a}_m + 0\vec{a}_{m+1} + \ldots + 0\vec{a}_n.$$

Образы оставшихся векторов

$$\left\{\hat{T}\vec{a}_{m+1}, \hat{T}\vec{a}_{m+2}, \ldots, \hat{T}\vec{a}_n\right\}$$

разложим по отмеченному базису всего пространства

$$\hat{T}\vec{a}_{m+1} = t_{m+1}^1\vec{a}_1 + t_{m+1}^2\vec{a}_2 + \ldots + t_{m+1}^m\vec{a}_m + t_{m+1}^{m+1}\vec{a}_{m+1} + \ldots + t_{m+1}^n\vec{a}_n;$$

$$\hat{T}\vec{a}_{m+2} = t_{m+2}^1\vec{a}_1 + t_{m+2}^2\vec{a}_2 + \ldots + t_{m+2}^m\vec{a}_m + t_{m+2}^{m+1}\vec{a}_{m+1} + \ldots + t_{m+2}^n\vec{a}_n;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;$$

$$\hat{T}\vec{a}_n = t_n^1\vec{a}_1 + t_n^2\vec{a}_2 + \ldots + t_n^m\vec{a}_m + t_n^{m+1}\vec{a}_{m+1} + \ldots + t_n^n\vec{a}_n.$$

Теперь матрица оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ принимает следующий вид:

$$T = \begin{pmatrix} t_1^1 & t_2^1 & \ldots & t_m^1 & t_{m+1}^1 & t_{m+2}^1 & \ldots & t_n^1 \\ t_1^2 & t_2^2 & \ldots & t_m^2 & t_{m+1}^2 & t_{m+2}^2 & \ldots & t_n^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ t_1^m & t_2^m & \ldots & t_m^m & t_{m+1}^m & t_{m+2}^m & \ldots & t_n^m \\ 0 & 0 & \ldots & 0 & t_{m+1}^{m+1} & t_{m+2}^{m+1} & \ldots & t_n^{m+1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & t_{m+1}^n & t_{m+2}^n & \ldots & t_n^n \end{pmatrix}. \qquad (4.1.11)$$

Такие матрицы называют ***клеточными*** и записывают в виде:

$$T = \begin{pmatrix} T_1^1 & T_2^1 \\ O & T_2^2 \end{pmatrix}. \qquad (4.1.12)$$

Здесь ***клетки*** матрицы T имеют следующее строение:
T_1^1 — квадратная матрица порядка m;
T_2^2 — квадратная матрица порядка $n - m$;

O — нулевая матрица размерами $(n-m)\times m$;

T_2^1 — матрица размерами $m\times(n-m)$.

Если пространство X^n разложено в прямую сумму инвариантных подпространств X_{inv}^m и X_{inv}^{n-m} оператора $\hat{T}$, то вид матрицы оператора ещё более упрощается.

Теорема 4.1.6. *Пусть* $\hat{T}: X^n \to \hat{T}\left(X^n\right)$ *— некоторый линейный оператор и*

$$X^n = X_{inv}^m \oplus X_{inv}^{n-m},$$

где X_{inv}^m и X_{inv}^{n-m}— инвариантные подпространства оператора $\hat{T}$. Тогда в некотором базисе $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$ матрица оператора имеет вид

$$T = \begin{pmatrix} T_1^1 & O_2^1 \\ O_1^2 & T_2^2 \end{pmatrix}, \tag{4.1.13}$$

где T_1^1 *— квадратная матрица порядка* m, T_2^2 *— квадратная матрица порядка* $n-m$, O_2^1 *— нулевая матрица размерами* $m\times(n-m)$, O_1^2 *— нулевая матрица размерами* $(n-m)\times m$.

Доказательство. Выбирая в пространстве X^n базис

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m, \vec{a}_{m+1}, \ldots, \vec{a}_n\right\}$$

так, чтобы первые его m векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ принадлежали подпространству X_{inv}^m, а оставшиеся $n-m$ векторов $\left\{\vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n\right\}$ принадлежали подпространству X_{inv}^{n-m}, разложим образы

$$\left\{\hat{T}\vec{a}_1, \hat{T}\vec{a}_2, \ldots, \hat{T}\vec{a}_m\right\} \subset X_{inv}^m$$

по базису $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ подпространства X_{inv}^m, а оставшиеся образы

$$\left\{\hat{T}\vec{a}_{m+1}, \hat{T}\vec{a}_{m+2}, \ldots, \hat{T}\vec{a}_n\right\} \subset X_{inv}^{n-m}$$

— по базису $\left\{\vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n\right\}$ подпространства X_{inv}^{n-m}. Очевидно, что матрица оператора $\hat{T}$ принимает указанный вид

$$T = \begin{pmatrix} T_1^1 & O \\ O & T_2^2 \end{pmatrix}. \bullet\bullet$$

Отметим, что каждый линейный оператор $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ имеет, по крайней мере, два инвариантных подпространства — само пространство X^n и нулевое подпространство, которые называются ***тривиальными***.

В частности, для каждого оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ инвариантными подпространствами являются его множество значений и ядро, которые будут тривиальными в том и только в том случае, если оператор невырожденный.

Рассмотрим ещё одно важное понятие.

Определение 4.1.8. *Пусть $X_{inv} \subset X^n$ — инвариантное подпространство оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$. Тогда* ***сужение***

$$\hat{T}\left|X_{inv} : X_{inv} \to X_{inv}\right.$$

оператора $\hat{T}$ на инвариантное подпространство X_{inv}, действие которого определено правилом

$$\left(\hat{T}\left|X_{inv}\right.\right)\vec{x} = \left\{\hat{T}\vec{x} : \vec{x} \in X_{inv}\right\},$$

называется ***индуцированным оператором****, порождённым оператором $\hat{T}$.* •

Оператор $\hat{T}$ называется ***порождающим оператором*** для индуцированного оператора $\hat{T}\left|X_{inv}\right.$. Индуцированный оператор является линейным, совпадает с порождающим оператором на инвариантном подпространстве и не определён на дополнении $C_{X^n}\left(X_{inv}\right)$ инвариантного подпространства до всего пространства.

Индуцированный оператор, как и любой другой линейный оператор, имеет, по крайней мере, один собственный вектор и совпадает с порождающим оператором на своём множестве определения. Получаем следующее утверждение.

Лемма 4.1.6. *Любой линейный оператор*

$$\hat{T} : X^n \to \hat{T}\left(X^n\right)$$

в каждом инвариантном подпространстве имеет, по крайней мере, один собственный вектор.

Если пространство X^n разложено в прямую сумму инвариантных подпространств $X_{inv}^{(k)}$ $(k = 1, 2, \ldots, m)$ линейного оператора $\hat{T} : X^n \to \hat{T}\left(X^n\right)$, то есть

$$X^n = X_{inv}^{(1)} \oplus X_{inv}^{(2)} \oplus \ldots \oplus X_{inv}^{(m)},$$

то оператор имеет, по крайней мере, m собственных векторов, образующих линейно независимую систему. Очевидно, любой собственный вектор и любое собственное значение индуцированного оператора являются соответственно собственным вектором и собственным значением порождающего оператора.

Теорема 4.1.7. *Пусть* $\hat{T}\big|X_{inv}$ *— индуцированный оператор, порождённый оператором* $\hat{T} : X^n \to \hat{T}\left(X^n\right)$ *на нетривиальном инвариантном подпространстве* $X_{inv} \subset X^n$. *Тогда характеристический многочлен* $\left(T\big|X_{inv}\right)(\mu)$ *индуцированного оператора является делителем характеристического многочлена* $T(\mu)$ *порождающего оператора.*

Д о к а з а т е л ь с т в о. В условиях теоремы матрица порождающего оператора имеет клеточный вид (4.1.12)

$$T = \begin{pmatrix} T_1^1 & T_2^1 \\ O & T_2^2 \end{pmatrix},$$

причём матрица индуцированного оператора является клеткой T_1^1. Характеристический многочлен порождающего оператора

$$T(\mu) = \det\left(T - \mu I\right),$$

а характеристический многочлен индуцированного оператора равен

$$\left(T\big|X_{inv}\right)(\mu) = \det\left(T_1^1 - \mu I_1^1\right).$$

Запишем характеристический многочлен порождающего оператора в виде

$$T(\mu) = \det(T - \mu I) = \det\begin{pmatrix} T_1^1 - \mu I_1^1 & T_2^1 \\ O_1^2 & T_2^2 - \mu I_2^2 \end{pmatrix}.$$

Напомним, что здесь I_1^1 — единичная матрица размера $m \times m$, а I_2^2 — единичная матрица размера $(n-m)\times(n-m)$. Учитывая вид матрицы (4.1.11), разложим полученный определитель последовательно по первым m столбцам. В результате несложных преобразований получим произведение двух сомножителей. Первый сомножитель является суммой всевозможных произведений ненулевых элементов клетки $T_1^1 - \mu \cdot I_1^1$, взятых по одному из каждой строки и каждого столбца и снабжённых соответствующими общему определению определителя знаками, то есть определитель $\det\left(T_1^1 - \mu \cdot I_1^1\right)$ матрицы индуцированного оператора. Второй сомножитель — определитель $\det\left(T_2^2 - \mu \cdot I_2^2\right)$ клетки $T_2^2 - \mu \cdot I_2^2$. Итак, для характеристического многочлена порождающего оператора получаем представление

$$T(\mu) = \det(T - \mu I) = \det\left(T_1^1 - \mu I_1^1\right) \cdot \det\left(T_2^2 - \mu I_2^2\right),$$

которое и доказывает утверждение теоремы. ••

Следствие из теоремы 4.1.7. *Пространство X^n можно представить в виде прямой суммы*

$$X^n = X_{inv}^{k_1} \oplus X_{inv}^{k_2} \oplus \ldots \oplus X_{inv}^{k_m}$$

инвариантных подпространств оператора $\hat{T} : X^n \to X^n$ в том и только в том случае, если выполняются следующие два условия:

1) *характеристический многочлен оператора раскладывается на линейные множители, то есть*

$$T(\mu) = (\mu - \mu_1)^{k_1} (\mu - \mu_2)^{k_2} \ldots (\mu - \mu_m)^{k_m};$$

2) *размерность каждого подпространства $X_{inv}^{k_j}$ $(j = 1, 2, \ldots, m)$ равна кратности соответствующего корня характеристического многочлена.*

Доказательство. Из теоремы следует, что $\dim X_{inv}^{k_j} \le k_j$. Поэтому

$$\sum_j \dim X_{inv}^{k_j} \le \sum_j k_j \le n.$$

Представление

$$X^n = X_{inv}^{k_1} \oplus X_{inv}^{k_2} \oplus \ldots \oplus X_{inv}^{k_m}$$

справедливо в том и только в том случае, если базис X^n является объединением базисов подпространств $X_{inv}^{k_j}$, то есть в том и только в том случае, если

$$\sum \dim X_{inv}^{k_j} = n.$$

В силу предыдущего неравенства имеем

$$\sum_j k_j = n$$

и

$$\dim X_{inv}^{k_j} = k_j \left(j = 1, 2, \ldots, m\right).$$

Первое равенство означает выполнение первого условия утверждения, а второе — второго. ••

Треугольная форма матрицы оператора. Дальнейшего упрощения матрицы оператора можно достигнуть, используя теорию инвариантных подпространств.

Лемма 4.1.7. *Множество значений* $F_p\left(\hat{T}\right)\left(X^n\right)$ *и ядро* $K\left(F_p\left(\hat{T}\right)\right)$ *операторного многочлена* $F_p\left(\hat{T}\right)$ *являются инвариантными подпространствами оператора* $\hat{T}$.

Д о к а з а т е л ь с т в о. Пусть вектор $\vec{y}$ принадлежит множеству значений операторного многочлена

$$\vec{y} \in F_p\left(\hat{T}\right)\left(X^n\right).$$

Это означает, что

$$\left(\exists \vec{x} \in X^n\right): \vec{y} = F_p\left(\hat{T}\right)\vec{x}.$$

В силу коммутативности операторов $\hat{T}$ и $F_p\left(\hat{T}\right)$ (формула (4.1.10)) имеем

$$\hat{T}\vec{y} = \hat{T} F_p\left(\hat{T}\right)\vec{x} = F_p\left(\hat{T}\right)\left(\hat{T}\vec{x}\right).$$

Таким образом, вектор $\hat{T}\vec{y} \in \hat{T}(X^n)$ является результатом применения оператора $F_p\left(\hat{T}\right)$ к вектору $\hat{T}\vec{x} \in X^n$, то есть

$$\hat{T}\vec{y} \in F_p\left(\hat{T}\right)(X^n).$$

Если теперь $\vec{x} \in K\left(F_p\left(\hat{T}\right)\right)$, то $F_p\left(\hat{T}\right)\vec{x} = \vec{0}$ и

$$F_p\left(\hat{T}\right)\left(\hat{T}\vec{x}\right) = \hat{T}\left(F_p\left(\hat{T}\right)\vec{x}\right) = \hat{T}\vec{0} = \vec{0},$$

то есть

$$\hat{T}\vec{x} \in K\left(F_p\left(\hat{T}\right)\right). \bullet\bullet$$

Лемма 4.1.8. *Если собственное значение* μ *оператора* $\hat{T} : X^n \to \hat{T}(X^n)$ *является (не является) корнем многочлена* $F_n(z)$, *то все собственные векторы оператора* $\hat{T}$, *соответствующие этому собственному значению, принадлежат ядру (множеству значений) оператора* $F_n\left(\hat{T}\right)$.

Д о к а з а т е л ь с т в о. Пусть $\hat{T}\vec{x} = \mu\vec{x}$. Тогда по лемме 4.1.3 вектор $\vec{x}$ является собственным вектором операторного многочлена $F_n\left(\hat{T}\right)$, соответствующим собственному значению $F_n(\mu)$, то есть

$$F_n\left(\hat{T}\right)\vec{x} = \left(a_0 + a_1\mu + a_2\mu^2 + \ldots + a_n\mu^n\right)\vec{x} \equiv F_n(\mu)\vec{x}.$$

Если μ является корнем многочлена $F_n(z)$, то $F_n(\mu) = 0$ и

$$F_n\left(\hat{T}\right)\vec{x} = \vec{0},$$

то есть $\vec{x}$ принадлежит ядру оператора $F_n\left(\hat{T}\right)$. Если же μ не является корнем $F_n(z)$, то $F_n(\mu) \neq 0$ и вектор

$$F_n\left(\hat{T}\right)\vec{x} = F_n(\mu)\vec{x} \neq \vec{0},$$

следовательно, вектор

$$\vec{x} \in F_n\left(\hat{T}\right)\left(X^n\right),$$

то есть принадлежит множеству значений оператора $F_n\left(\hat{T}\right)$. ••

Теперь мы можем доказать следующую важную теорему.

Теорема 4.1.8. *Каждый линейный оператор $\hat{T}$, действующий в комплексном векторном пространстве X^n, имеет, по крайней мере, одно инвариантное подпространство размерности $n-1$.*

Доказательство. Оператор $\hat{T}$ имеет, по крайней мере, один (ненулевой!) собственный вектор $\vec{x}$, соответствующий, например, собственному значению μ. По лемме 4.1.7 множество значений $\left(\hat{T}-\mu\hat{I}\right)\left(X^n\right)$ операторного многочлена $\hat{T}-\mu\hat{I}$ является инвариантным подпространством оператора $\hat{T}$. Так как $\vec{x} \in K\left(\hat{T}-\mu\hat{I}\right)$, то оператор $\hat{T}-\mu\hat{I}$ вырожденный и, следовательно,

$$\dim\left(\hat{T}-\mu\hat{I}\right)\left(X^n\right) \le n-1.$$

Пусть теперь выполняется включение

$$\left(\hat{T}-\mu\hat{I}\right)\left(X^n\right) \subset L \subset X^n.$$

Тогда

$$\left(\forall \vec{x} \in X^n\right) \vec{y} = \left(\hat{T}-\mu\hat{I}\right)\vec{x} \in \left(\hat{T}-\mu\hat{I}\right)\left(X^n\right) \in L.$$

Следовательно, подпространство $L \subset X^n$ является инвариантным подпространством оператора $\hat{T}-\mu\hat{I}$ и, очевидно, оператора $\hat{T}$. ••

Определение 4.1.9. *Система подпространств*

$$F\left\{X^0, X^1, X^2, \ldots, X^{n-1}, X^n\right\}$$

(абстрактного) векторного пространства X^n называется ***флагом подпространств*** *$X^0, X^1, X^2, \ldots, X^{n-1}, X^n$* ***на*** *$X^n$, если её элементы (подпространства) удовлетворяют следующим двум условиям:* $\left(\forall k = 0, 1, 2, \ldots, n-1\right)$

1) $X^k \subset X^{k+1}$;
2) $\dim X^k = k$. •

Теорема 4.1.9. *Для любого линейного оператора $\hat{T}$, действующего на пространстве X^n, существует флаг*

$$F\left\{X^0_{inv}, X^1_{inv}, X^2_{inv}, \ldots, X^{n-1}_{inv}, X^n_{inv}\right\}$$

инвариантных подпространств.

Доказательство. Очевидно, подпространства X^0_{inv} и X^n_{inv} существуют как тривиальные инвариантные подпространства. По теореме 4.1.8 оператор $\hat{T}$ имеет инвариантное подпространство X^{n-1}_{inv}. Рассмотрим на подпространстве X^{n-1}_{inv} индуцированный оператор $\hat{T}\big|X^{n-1}_{inv}$. По теореме 4.1.8 он также имеет инвариантное подпространство X^{n-2}_{inv}, которое является инвариантным и для оператора $\hat{T}$. Рассматривая на подпространстве X^{n-2}_{inv} индуцированный оператор $\hat{T}\big|X^{n-2}_{inv}$, аналогично предыдущему докажем существование инвариантного подпространства X^{n-3}_{inv}, и так далее. ••

Основной интерес для приложений представляют два следствия из только что доказанной теоремы.

Следствие 1 из теоремы 4.1.9. *Пусть*

$$F\left\{X^0_{inv}, X^1_{inv}, X^2_{inv}, \ldots, X^n_{inv}\right\}$$

— флаг инвариантных подпространств оператора $\hat{T}$. Тогда в базисе

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} : \left(\forall k = 1, 2, \ldots, n\right) \vec{a}_k \in X^k_{inv},\ \vec{a}_k \neq \vec{0},$$

матрица оператора $\hat{T}$ имеет вид

$$T = \begin{pmatrix} t^1_1 & t^1_2 & \ldots & t^1_{n-1} & t^1_n \\ 0 & t^2_2 & \ldots & t^2_{n-1} & t^2_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & t^{n-1}_{n-1} & t^{n-1}_n \\ 0 & 0 & \ldots & 0 & t^n_n \end{pmatrix}. \qquad (4.1.14)$$

Д о к а з а т е л ь с т в о. Пусть $\vec{a}_k \in X_{inv}^k$ $(k=1, 2, \ldots, n)$. Так как X_{inv}^k — инвариантное подпространство оператора $\hat{T}$, то в разложении вектора $\hat{T}\vec{a}_k$ по векторам базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ останутся только первые k слагаемых:

$$\hat{T}\vec{a}_k = t_k^1 \vec{a}_1 + t_k^2 \vec{a}_2 + \ldots + t_k^k \vec{a}_k \,(k=1, 2, \ldots, n).$$

Что и приводит к матрице вида (4.8.14). ••

Квадратная матрица, все элементы которой, расположенные ниже (выше) главной диагонали, равны нулю, называется ***правой*** (***левой***) ***треугольной матрицей***.

Следствие 2 из теоремы 4.1.9. *Если оператор*

$$\hat{T} : X^n \to \hat{T}\left(X^n\right)$$

имеет в некотором базисе $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$ *треугольную матрицу T, то диагональные элементы матрицы T равны собственным значениям оператора* $\hat{T}$ *с учётом их кратности.*

Д о к а з а т е л ь с т в о. В силу того, что матрица оператора $\hat{T}$ треугольная, характеристический многочлен оператора $\det\left(\hat{T} - \mu \hat{I}\right)$ после последовательного разложения определителя по элементам столбцов принимает вид

$$\det\left(\hat{T} - \mu \hat{I}\right) = \left(t_1^1 - \mu\right)\left(t_2^2 - \mu\right)\ldots\left(t_n^n - \mu\right) = \prod_{k=1}^{n}\left(t_k^k - \mu\right),$$

откуда получаем, что

$$\left(\forall k = 1, 2, \ldots, n\right) t_k^k - \mu = 0 \Rightarrow t_k^k = \mu. \;••$$

4.2. Линейные операторы в евклидовых пространствах

Операторы в пространствах со скалярным произведением, в частности в евклидовых и унитарных пространствах, обладают значительно более богатыми свойствами, чем операторы в пространствах без скалярного произведения.

Линейные функционалы и сопряжённое пространство[*]. Рассмотрим сначала частный случай оператора, действующего на произвольном векторном пространстве X^n — так называемый ***функционал***.

Определение 4.2.1. *Линейный оператор* $\hat{A}: X^n \to P$, *где* P *— числовое поле, над которым определено пространство* X^n, *называется* ***линейным функционалом***. •

Таким образом, линейный функционал — это ***линейное отображение пространства*** X^n ***в поле*** P. Будем обозначать линейные функционалы малыми буквами латинского алфавита, но без «шляпки», например, $f\left(\vec{x}\right)$, что согласуется с общепринятым обозначением для абстрактной функции, которой, по существу, и является функционал.

Выше мы видели, что представителем линейного оператора в фиксированных базисах является матрица. Возникает вопрос: *что является представителем линейного функционала*?

В пространстве X^n зафиксируем базис $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$. Тогда

$$\left(\forall \vec{x} \in X^n\right) \vec{x} = x^1 \vec{a}_1 + x^2 \vec{a}_2 + \ldots + x^n \vec{a}_n;$$

$$f\left(\vec{x}\right) = f\left(x^1 \vec{a}_1 + x^2 \vec{a}_2 + \ldots + x^n \vec{a}_n\right) =$$

$$= x^1 f\left(\vec{a}_1\right) + x^2 f\left(\vec{a}_2\right) + \ldots + x^n f\left(\vec{a}_n\right). \tag{4.2.1}$$

Вводя обозначение $\left(\forall k = 1, 2, \ldots, n\right) f\left(\vec{a}_k\right) = b_k$, получаем

$$f\left(\vec{x}\right) = x^1 b_1 + x^2 b_2 + \ldots + x^n b_n. \tag{4.2.2}$$

Таким образом, линейная комбинация образов базисных векторов в правой части (4.2.1) является ***линейной формой***, то есть однородным многочленом первой степени ***относительно*** n ***неизвестных*** $x^1, x^2, \ldots, x^n$.

Если векторное пространство является евклидовым пространством E^n с фиксированным ортонормированным базисом $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ (например, пространство R^n), то линейная форма (4.2.2) может быть представлена как скалярное произведение

$$f\left(\vec{x}\right)=\left(\vec{x},\vec{b}\right)=x_1b_1+x_2b_2+\ldots+x_nb_n$$

вектора $\vec{x}$ и некоторого ***фиксированного вектора*** $\vec{b}$, зависящего только от вида функционала f, но не от вектора $\vec{x}$.

Здесь использована запись вектора через ковариантные координаты $x_1, x_2, \ldots, x_n$, которые совпадают в ортонормированном базисе с контравариантными координатами $x^1=x_1, x^2=x_2, \ldots, x^n=x_n$.

Пусть теперь в евклидовом пространстве E^n задан ортонормированный базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ и фиксированный вектор

$$\vec{b}=b_1\vec{e}_1+b_2\vec{e}_2+\ldots+b_n\vec{e}_n.$$

Тогда, в силу свойств линейности, $\left(\forall\vec{x}\in E^n\right)$ скалярное произведение

$$f\left(\vec{x}\right)=\left(\vec{x},\vec{b}\right)$$

является линейным функционалом.

Итак, мы доказали следующее утверждение.

Лемма 4.2.1. *Пусть в евклидовом пространстве E^n задан ортонормированный базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$. Тогда линейный функционал $f: E^n\to P$ однозначно реализуется в виде скалярного произведения $f\left(\vec{x}\right)=\left(\vec{x},\vec{b}\right)$ произвольного вектора $\vec{x}\in E^n$ на фиксированный вектор $\vec{b}\in E^n$, зависящий только от вида функционала f.*

Из леммы следует, что каждый линейный функционал в ортонормированном базисе евклидова пространства E^n может быть записан в виде линейной формы. Кроме этого, из определения 4.2.1 очевидным образом следует, что *множество всех линейных функционалов, определённых на евклидовом пространстве E^n, образует линейное векторное пространство, которое называется **сопряжённым пространству** E^n.*

В сопряжённом пространстве имеется базис.

Определим линейные функционалы $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$, ставящие в соответствие каждому вектору $\vec{x}\in E^n$ его координату в фиксированном базисе $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, то есть

$$\left(\forall \vec{x} \in E^n \wedge \forall k = 1, 2, \ldots, n\right).\ \varphi_k\left(\vec{x}\right) = x_k.$$

Очевидно, что

$$\varphi_k\left(\vec{e}_j\right) = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

Тогда по лемме 4.2.1 имеем:

$$f\left(\vec{x}\right) = b_1x_1 + b_2x_2 + \ldots + b_nx_n = b_1\varphi_1\left(\vec{x}\right) + b_2\varphi_2\left(\vec{x}\right) + \ldots + b_n\varphi_n\left(\vec{x}\right) =$$
$$= \left(b_1\varphi_1 + b_2\varphi_2 + \ldots + b_n\varphi_n\right)\left(\vec{x}\right).$$

Таким образом, коэффициенты $\{b_1, b_2, \ldots, b_n\}$ можно рассматривать как координаты функционала f в базисе $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ сопряжённого пространства. По этой причине функционалы называют иногда ***ковекторами***, а базис $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ — ***дуальным*** к базису $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ пространства E^n. Сопряжённое к E^n пространство линейных функционалов часто обозначают символом E^{n*}.

Сопряжённый оператор. Рассмотрим линейный оператор

$$\hat{T} : E^m \to E^n,$$

действующий из евклидова (в общем случае, комплексного) пространства E^m в евклидово пространство E^n.

Определение 4.2.2. *Пусть $\hat{T} : E^m \to E^n$ — некоторый линейный оператор. Оператор $\hat{T}^* : E^n \to E^m$ называется **сопряжённым по отношению к оператору** $\hat{T}$, если $\left(\forall \vec{x} \in E^m, \forall \vec{y} \in E^n\right)$ выполняется условие*

$$\left(\hat{T}\vec{x}, \vec{y}\right) \overset{def}{=} \left(\vec{x}, \hat{T}^*\vec{y}\right). \bullet \tag{4.2.3}$$

Теорема 4.2.1. *Для любого линейного оператора $\hat{T} : E^m \to E^n$ существует единственный сопряжённый оператор $\hat{T}^* : E^n \to E^m$, который также является линейным.*

Д о к а з а т е л ь с т в о. Пусть оператор $\hat{T}^*$ существует, докажем его ***единственность***. Для этого предположим, что $\hat{T}^*$ неединственный, то есть существуют, например, два оператора $\hat{T}_1^*$ и $\hat{T}_2^*$, удовлетворяющие определению 4.2.2. Тогда по формуле (4.2.3) имеем:

$$\left(\hat{T}\vec{x},\vec{y}\right)=\left(\vec{x},\hat{T}_1^*\vec{y}\right),\ \left(\hat{T}\vec{x},\vec{y}\right)=\left(\vec{x},\hat{T}_2^*\vec{y}\right), \tag{4.2.4}$$

откуда получаем

$$\left(\vec{x},\hat{T}_1^*\vec{y}\right)=\left(\vec{x},\hat{T}_2^*\vec{y}\right)\Rightarrow\left(\vec{x},\hat{T}_1^*\vec{y}-\hat{T}_2^*\vec{y}\right)=0. \tag{4.2.5}$$

В силу того, что в определении 4.2.2, то есть в формуле (4.2.3), вектор $\vec{x}\in E^m$ произволен, полагаем в равенстве (4.2.5) $\vec{x}=\hat{T}_1^*\vec{y}-\hat{T}_2^*\vec{y}$, получим

$$\left(\hat{T}_1^*\vec{y}-\hat{T}_2^*\vec{y},\hat{T}_1^*\vec{y}-\hat{T}_2^*\vec{y}\right)=0.$$

Так как скалярное произведение удовлетворяет аксиоме невырожденности, из последнего равенства имеем

$$\hat{T}_1^*\vec{y}-\hat{T}_2^*\vec{y}=\vec{0}\Rightarrow\hat{T}_1^*\vec{y}=\hat{T}_2^*\vec{y},$$

откуда в силу произвольности вектора $\vec{y}$ следует, что $\hat{T}_1^*=\hat{T}_2^*$. Единственность сопряжённого оператора доказана.

Докажем ***линейность*** сопряжённого оператора, предполагая его существование. Используя определение (4.2.3) и свойства скалярного произведения, получаем:

$$\left(\forall\vec{x}\in E^m\right),\left(\forall\vec{y},\vec{z}\in E^n\right)\text{ и }\left(\forall\alpha,\beta\in C\right)$$

а) $\left(\hat{T}\vec{x},\alpha\vec{y}+\beta\vec{z}\right)=\left(\vec{x},\hat{T}^*\left(\alpha\vec{y}+\beta\vec{z}\right)\right)$;

б) $\left(\hat{T}\vec{x},\alpha\vec{y}+\beta\vec{z}\right)=\left(\hat{T}\vec{x},\alpha\vec{y}\right)+\left(\hat{T}\vec{x},\beta\vec{z}\right)=$

$$= \alpha \cdot \left(\hat{T} \vec{x}, \vec{y} \right) + \beta \left(\hat{T} \vec{x}, \vec{z} \right) = \alpha \left(\vec{x}, \hat{T}^* \vec{y} \right) + \beta \left(\vec{x}, \hat{T}^* \vec{z} \right) =$$

$$= \left(\vec{x}, \alpha \hat{T}^* \vec{y} \right) + \left(\vec{x}, \beta \hat{T}^* \vec{z} \right) = \left(\vec{x}, \alpha \hat{T}^* \vec{y} + \beta \hat{T}^* \vec{z} \right).$$

Из сравнения формул а) и б) следует, что выполняется условие

$$\hat{T}^* \left(\alpha \vec{y} + \beta \vec{z} \right) = \alpha \hat{T}^* \vec{y} + \beta \hat{T}^* \vec{z},$$

то есть сопряжённый оператор $\hat{T}^*$ линеен в соответствии с определением.

Докажем теперь ***существование*** сопряжённого оператора. Для этого найдём матрицу сопряжённого оператора $\hat{T}^*$ в ортонормированных базисах

$$\left\{ \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m \right\} \subset E^m, \left\{ \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n \right\} \subset E^n.$$

В этих базисах имеем разложения

$$\vec{x} = \sum_{k=1}^{m} x^k \vec{e}_k, \; \vec{y} = \sum_{j=1}^{n} y^j \vec{g}_j.$$

Найдём образы $\hat{T} \vec{x}$ и $\hat{T}^* \vec{y}$:

$$\hat{T} \vec{x} = \hat{T} \sum_{k=1}^{m} x^k \vec{e}_k = \sum_{k=1}^{m} x^k \hat{T} \vec{e}_k = \sum_{k=1}^{m} x^k \sum_{j=1}^{n} t_k^j \vec{g}_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{m} t_k^j x^k \right) \vec{g}_j, \quad (4.2.6')$$

$$\hat{T}^* \vec{y} = \hat{T}^* \sum_{j=1}^{n} y^j \vec{g}_j = \sum_{j=1}^{n} y^j \hat{T}^* \vec{g}_j = \sum_{j=1}^{n} y^j \sum_{k=1}^{m} t_j^{k*} \vec{e}_k = \sum_{k=1}^{m} \left(\sum_{j=1}^{n} t_j^{k*} y^j \right) \vec{e}_k. \quad (4.2.6'')$$

Используем соотношения (4.2.6’) и (4.2.6’’) для вычисления левой и правой частей определения (4.2.3).

Для левой части получаем:

$$\left(\hat{T} \vec{x}, \vec{y} \right) = \left(\sum_{j=1}^{n} \left(\sum_{k=1}^{m} t_k^j x^k \right) \vec{g}_j, \sum_{i=1}^{n} y^i \vec{g}_i \right) =$$

$$= \sum_{i,\, j=1}^{n} \sum_{k=1}^{m} \overline{t_k^j} \cdot \overline{x^k} \cdot y^i \cdot \left(\vec{g}_j, \vec{g}_i \right) = \sum_{j=1}^{n} \sum_{k=1}^{m} \overline{t_k^j} \cdot \overline{x^k} \cdot y^j. \quad (4.2.7)$$

Для правой части получаем:

$$\left(\vec{x}, \hat{T}^{*}\vec{y}\right) = \left(\sum_{i=1}^{m} x^{i}\vec{e}_{i}, \sum_{k=1}^{m}\left(\sum_{j=1}^{n} t_{j}^{k*} y^{j}\right)\vec{e}_{k}\right) = \sum_{i,k=1}^{m} \overline{x^{i}} \cdot \left(\sum_{j=1}^{n} t_{j}^{k*} y^{j}\right)\left(\vec{e}_{i}, \vec{e}_{k}\right) =$$

$$= \sum_{k=1}^{m}\sum_{j=1}^{n} t_{j}^{k*} \cdot y^{j} \cdot \overline{x^{k}} = \sum_{j=1}^{n}\sum_{k=1}^{m} t_{j}^{k*} \cdot \overline{x^{k}} \cdot y^{j}. \qquad (4.2.8)$$

Из выражений (4.2.7) и (4.2.8) с учётом формулы (4.2.3) определения 4.2.2 получаем

$$t_{j}^{k*} = \overline{t_{k}^{j}}. \qquad (4.2.9)$$

Таким образом, в любых ортонормированных базисах

$$\left\{\vec{e}_{1}, \vec{e}_{2}, \ldots, \vec{e}_{m}\right\} \subset E^{m}, \left\{\vec{g}_{1}, \vec{g}_{2}, \ldots, \vec{g}_{n}\right\} \subset E^{n}$$

элементы матрицы сопряжённого оператора $\hat{T}^{*}$ вычисляются по формуле (4.2.9), что и доказывает существование сопряжённого оператора. ••

Теорема 4.2.2. *Если* $\hat{T}: E^{m} \to E^{n}$, $\hat{R}: E^{m} \to E^{n}$ *— некоторые операторы и* $\beta \in C$ *— произвольное число, то справедливы следующие пять свойств сопряжённого оператора:*

$$1) \left(\hat{T}^{*}\right)^{*} = \hat{T};\ 2) \left(\hat{T}^{*}\right)^{-1} = \left(\hat{T}^{-1}\right)^{*};\ 3) \left(\beta\hat{T}\right)^{*} = \overline{\beta}\hat{T}^{*};$$

$$4) \left(\hat{T} + \hat{R}\right)^{*} = \hat{T}^{*} + \hat{R}^{*};\ 5) \left(\hat{T}\hat{R}\right)^{*} = \hat{R}^{*}\hat{T}^{*}. \qquad (4.2.10)$$

(Здесь черта над числом β означает комплексное сопряжение).

Д о к а з а т е л ь с т в о. Докажем первое соотношение. Пусть $\hat{T}$ — произвольный линейный оператор. Для сопряжённого оператора $\hat{T}^{*}$ сопряжённым будет оператор $\left(\hat{T}^{*}\right)^{*}$. Тогда $\left(\forall \vec{x} \in E^{m}\right)$ и $\left(\forall \vec{y} \in E^{n}\right)$

$$\left(\hat{T}\vec{x}, \vec{y}\right) = \left(\vec{x}, \hat{T}^{*}\vec{y}\right) = \overline{\left(\hat{T}^{*}\vec{y}, \vec{x}\right)} = \overline{\left(\vec{y}, \left(\hat{T}^{*}\right)^{*}\vec{x}\right)} = \left(\left(\hat{T}^{*}\right)^{*}\vec{x}, \vec{y}\right).$$

Левая часть равна правой части при любом векторе $\vec{x}$. Следовательно,

$$\left(\forall \vec{x} \in E^m\right)\left(\hat{T}^*\right)^* \vec{x} = \hat{T}\vec{x},$$

откуда имеем $\left(\hat{T}^*\right)^* = \hat{T}$. ••

Матрицы операторов $\hat{T}$ и $\hat{T}^*$ в любых ортонормированных базисах

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\right\} \subset E^m, \left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset E^n$$

связаны посредством соотношений (4.2.9). То есть

$$T = \begin{pmatrix} t_1^1 & t_2^1 & \ldots & t_m^1 \\ t_1^2 & t_2^2 & \ldots & t_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ t_1^n & t_2^n & \ldots & t_m^n \end{pmatrix}, T^* = \begin{pmatrix} \overline{t_1^1} & \overline{t_1^2} & \ldots & \overline{t_1^n} \\ \overline{t_2^1} & \overline{t_2^2} & \ldots & \overline{t_2^n} \\ \ldots & \ldots & \ldots & \ldots \\ \overline{t_m^1} & \overline{t_m^2} & \ldots & \overline{t_m^n} \end{pmatrix}.$$

Операторы $\hat{T}$ и $\hat{T}^*$ называются также ***взаимно сопряжёнными***.

Определение 4.2.3. *Матрица T^* размеров $m \times n$ с элементами $t^{*i}_{\ j}$ называется **сопряжённой по отношению к матрице** T размеров $n \times m$ с элементами t^i_j, если $t^{*i}_{\ j} = \overline{t^j_i}$ для всех i, j.* •

Теорема 4.2.3. *В любых ортонормированных базисах сопряжённым операторам соответствуют сопряжённые матрицы.*

Доказательство. Доказательство теоремы является прямым следствием теоремы 4.2.1 и определения 4.2.3.••

Очевидно, сопряжённая матрица удовлетворяет всем соотношениям (4.2.10). Легко видеть также, что в случае оператора, действующего в пространстве E^n, то есть в случае преобразования, сопряжённая матрица T^* связана с матрицей T операциями транспонирования и комплексного сопряжения элементов (в вещественных пространствах — только операцией транспонирования):

$$T^* = \overline{\left(T^T\right)} = \left(\overline{T}\right)^T. \tag{4.2.11}$$

В соотношении (4.2.11) буквой T обозначена операция транспонирования, а чертой — операция комплексного сопряжения.

Теорема 4.2.4. *Пусть* $\hat{T}: E^n \to E^n$ *и* E_{inv} *— некоторое инвариантное подпространство оператора* $\hat{T}$*, тогда его ортогональное дополнение* $E_{inv}^{\perp}$ *инвариантно относительно сопряжённого оператора* $\hat{T}^*$*.*

Д о к а з а т е л ь с т в о. Так как

$$\left(\forall \vec{x} \in E_{inv}\right) \hat{T}\vec{x} \in E_{inv},$$

то

$$\left(\forall \vec{y} \in E_{inv}^{\perp}\right)\left(\hat{T}\vec{x}, \vec{y}\right) = 0,$$

откуда

$$\left(\hat{T}\vec{x}, \vec{y}\right) = \left(\vec{x}, \hat{T}^*\vec{y}\right) = 0.$$

Поэтому $\hat{T}^*\vec{y} \in E_{inv}^{\perp}$. ••

Теорема 4.2.5. *Пусть* $T(\mu)$ *— характеристический многочлен линейного оператора* $\hat{T}: E^n \to E^n$*. Тогда характеристический многочлен сопряжённого оператора равен* $\overline{T(\mu)}$.

Д о к а з а т е л ь с т в о. Доказательство следует из теоремы 4.2.1 и инвариантности характеристического многочлена относительно выбора базиса. ••

Следствие из теоремы 4.2.5. *Если* μ_1 *— собственное значение оператора* $\hat{T}: E^n \to E^n$ *кратности k, то* $\overline{\mu_1}$ *— собственное значение сопряжённого оператора* $\hat{T}^*$ *той же кратности.*

В частности, в вещественном пространстве E^n характеристические многочлены оператора $\hat{T}: E^n \to E^n$ и ему сопряжённого оператора $\hat{T}^*$ совпадают и, следовательно, совпадают и их спектры.

Самосопряжённые операторы и их свойства. Для приложений большое значение имеют ***самосопряжённые операторы***.

Определение 4.2.4. *Линейный оператор* $\hat{T}: E^n \to E^n$ *называется* ***самосопряжённым (эрмитовым)****, если* $\hat{T}^* = \hat{T}$. •

В силу определения 4.2.4 и теоремы 4.2.2 элементы матрицы самосопряжённого оператора удовлетворяют соотношению $t^i_j = \overline{t^j_i}$. Сама матрица в этом случае называется ***эрмитовой***. Таким образом, если матрица эрмитова, то её элементы, симметричные относительно главной диагонали, являются ***комплексно-сопряжёнными***. Самосопряжённый оператор $\hat{T} : E^n \to E^n$, действующий в вещественном пространстве E^n, имеет матрицу, элементы которой симметричные относительно главной диагонали равны, то есть удовлетворяют равенству $t^i_j = t^j_i$. Такая матрица называется ***симметрической***. По этой причине самосопряжённый оператор, действующий в вещественном пространстве, иногда называется ***симметрическим***.

Самосопряжённые операторы обладают рядом свойств, которые нетрудно доказать, используя определение и свойства сопряжённого оператора (4.2.10).

1) *Единичный оператор* $\hat{I}$ *является самосопряжённым.*

2) *Сумма самосопряжённых операторов является самосопряжённым оператором.*

Д о к а з а т е л ь с т в о. Действительно, если $\hat{T}^* = \hat{T}$ и $\hat{R}^* = \hat{R}$, то

$$\left(\hat{T} + \hat{R}\right)^* = \hat{T}^* + \hat{R}^* = \hat{T} + \hat{R}. \bullet\bullet$$

3) *Композиция самосопряжённых операторов является самосопряжённым оператором в том и только в том случае, если эти операторы коммутативны.*

Д о к а з а т е л ь с т в о. Напомним, что операторы называются коммутативными, если $\hat{T}\hat{R} = \hat{R}\hat{T}$, или в другой записи

$$\hat{T}\hat{R} - \hat{R}\hat{T} = \hat{O},$$

где $\hat{O}$ — нулевой оператор.

Пусть $\hat{T}^* = \hat{T}$, $\hat{R}^* = \hat{R}$. Тогда оператор, сопряжённый композиции операторов в порядке $\hat{T}\hat{R}$

$$\left(\hat{T}\hat{R}\right)^* = \hat{R}^*\hat{T}^* = \hat{R}\hat{T},$$

равен композиции операторов в порядке $\hat{T}\hat{R}$ в том и только в том случае, если операторы являются коммутативными. ••

4) *Оператор* $\hat{T}^{-1}$*, обратный к невырожденному, самосопряжённому оператору* $\hat{T}: E^n \to E^n$*, является самосопряжённым.*

Доказательство. Действительно, если $\hat{T}^* = \hat{T}$, то

$$\left(\hat{T}^{-1}\right)^* = \left(\hat{T}^*\right)^{-1} = \hat{T}^{-1}. \bullet\bullet$$

5) *Если* $\hat{T}$ *— самосопряжённый оператор, то произведение этого оператора на некоторое число* α *является самосопряжённым оператором в том и только в том случае, если это число является вещественным (действительным).*

Доказательство. Действительно, из третьего свойства (4.2.10) имеем:

$$\left(\alpha\hat{T}\right)^* = \bar{\alpha}\hat{T}^* = \alpha\hat{T} \Leftrightarrow \alpha \in R^1. \bullet\bullet$$

Теорема 4.2.6. *Собственные значения самосопряжённого (эрмитова) оператора* $\hat{T}: E^n \to E^n$ *вещественны.*

Доказательство. Пусть $\vec{x}$ — собственный вектор самосопряжённого оператора $\hat{T}$, соответствующий собственному значению μ. Тогда из определения самосопряжённого оператора следует цепочка преобразований:

$$\left(\hat{T}\vec{x}, \vec{x}\right) = \left(\vec{x}, \hat{T}\vec{x}\right) \Rightarrow \left(\mu\vec{x}, \vec{x}\right) = \left(\vec{x}, \mu\vec{x}\right) \Rightarrow$$

$$\Rightarrow \bar{\mu}\left(\vec{x}, \vec{x}\right) = \mu\left(\vec{x}, \vec{x}\right) \Rightarrow \left(\bar{\mu} - \mu\right)\left(\vec{x}, \vec{x}\right) = 0.$$

Так как $\left(\vec{x}, \vec{x}\right) \neq 0$, то $\bar{\mu} = \mu$, то есть μ — вещественно. ••

В случае вещественных пространств справедлива аналогичная теорема. Прежде чем сформулировать и доказать эту теорему, докажем следующее вспомогательное утверждение.

Лемма 4.2.2. *Для любого линейного оператора* $\hat{T}$*, действующего в вещественном пространстве* X^n *размерности* $n > 2$*, существует одномерное или двумерное инвариантное подпространство.*

Доказательство. Если характеристический многочлен линейного оператора имеет, по крайней мере, один вещественный корень, то этот оператор имеет собственный вектор и, следовательно, одномерное инвариантное подпространство.

Если характеристический многочлен не имеет вещественных корней, то по основной теореме алгебры он имеет, по крайней мере, один комплексный корень вида $\mu = a + ib$. Решая для этого корня СЛАУ вида (4.1.6)

$$\begin{cases} \left(t_1^1 - \mu\right) z^1 + t_2^1 z^2 + \ldots + t_n^1 z^n = 0, \\ t_1^2 z^1 + \left(t_2^2 - \mu\right) z^2 + \ldots + t_n^2 z^n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ t_1^n z^1 + t_2^n z^2 + \ldots + \left(t_n^n - \mu\right) z^n = 0, \end{cases}$$

найдём соответствующие комплексные решения

$$z^1 = x^1 + iy^1,\, z^2 = x^2 + iy^2, \ldots, z^n = x^n + iy^n.$$

Подстановка в СЛАУ $z^k = x^k + iy^k \left(k = 1, 2, \ldots, n\right)$ и $\mu = a + ib$ даёт:

$$\begin{cases} t_1^1\left(x^1 + iy^1\right) + t_2^1\left(x^2 + iy^2\right) + \ldots + t_n^1\left(x^n + iy^n\right) = (a + ib)\left(x^1 + iy^1\right), \\ t_1^2\left(x^1 + iy^1\right) + t_2^2\left(x^2 + iy^2\right) + \ldots + t_n^2\left(x^n + iy^n\right) = (a + ib)\left(x^2 + iy^2\right), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ t_1^n\left(x^1 + iy^1\right) + t_2^n\left(x^2 + iy^2\right) + \ldots + t_n^n\left(x^n + iy^n\right) = (a + ib)\left(x^n + iy^n\right). \end{cases}$$

Приравнивая вещественные и мнимые части, получаем две системы равенств:

$$\begin{cases} t_1^1 x^1 + t_2^1 x^2 + \ldots + t_n^1 x^n = ax^1 - by^1, \\ t_1^2 x^1 + t_2^2 x^2 + \ldots + t_n^2 x^n = ax^2 - by^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ t_1^n x^1 + t_2^n x^2 + \ldots + t_n^n x^n = ax^n - by^n \end{cases} \tag{4.2.12}$$

и

$$\begin{cases} t_1^1 y^1 + t_2^1 y^2 + \ldots + t_n^1 y^n = a y^1 + b x^1, \\ t_1^2 y^1 + t_2^2 y^2 + \ldots + t_n^2 y^n = a y^2 + b x^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ t_1^n y^1 + t_2^n y^2 + \ldots + t_n^n y^n = a y^n + b x^n. \end{cases} \tag{4.2.13}$$

Рассмотрим в пространстве X^n два вещественных вектора:

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^n \vec{e}_n,\ \vec{y} = y^1 \vec{e}_1 + y^2 \vec{e}_2 + \ldots + y^n \vec{e}_n.$$

Из выражения (4.2.12) следует:

$$\hat{T}\vec{x} = a\vec{x} - b\vec{y},$$

а из выражения (4.2.13) следует:

$$\hat{T}\vec{y} = b\vec{x} + a\vec{y}.$$

Векторы $\vec{x}$ и $\vec{y}$ порождают подпространство $X^2 \subset X^n$. Действительно, предположим, что эти векторы образуют линейно зависимую систему, то есть $\vec{y} = \alpha \cdot \vec{x}$. Тогда

$$\hat{T}\vec{x} = a\vec{x} - b\vec{y} = (a - b\alpha)\vec{x},$$

откуда видно, что вектор $\vec{x}$ являлся бы собственным вектором оператора $\hat{T}$ с вещественным собственным значением $a - b\alpha$, что противоречит комплексности корня μ. Подпространство X^2, порождённое векторами $\vec{x}$ и $\vec{y}$, является инвариантным относительно оператора $\hat{T}$. Действительно,

$$\left(\forall \vec{u} \in X^2\right) : \vec{u} = \alpha \cdot \vec{x} + \beta \cdot \vec{y}$$

получаем:

$$\hat{T}\vec{u} = \alpha \cdot \hat{T}\vec{x} + \beta \cdot \hat{T}\vec{y} = \alpha \cdot \left(a\vec{x} - b\vec{y}\right) + \beta \cdot \left(b\vec{x} + a\vec{y}\right) =$$

$$= (\alpha a + \beta b)\vec{x} + (\beta a - \alpha b)\vec{y}.$$

Этот вектор, очевидно, также принадлежит X^2, то есть X^2 — инвариантное подпространство оператора $\hat{T}$. ••

Теорема 4.2.7. *Все корни характеристического многочлена самосопряжённого оператора $\hat{T} : E^n \to E^n$, действующего в вещественном евклидовом пространстве E^n, вещественны.*

Доказательство. Пусть $\mu = a + ib$ — комплексный корень характеристического многочлена самосопряжённого оператора $\hat{T}$. Тогда из доказательства леммы 4.2.2 следует, что в пространстве E^n имеется двумерное подпространство, порождённое векторами $\vec{x}$ и $\vec{y}$, такими, что

$$\begin{aligned} \hat{T}\vec{x} &= a \cdot \vec{x} - b \cdot \vec{y}, \\ \hat{T}\vec{y} &= b \cdot \vec{x} + a \cdot \vec{y}, \end{aligned} \tag{4.2.14}$$

причём

$$a \neq 0,\ b \neq 0,\ \vec{x} \neq \vec{0},\ \vec{y} \neq \vec{0}.$$

Находя скалярное произведение обеих частей первого из равенств (4.2.14) и вектора $\vec{y}$ и скалярное произведение обеих частей второго равенства (4.2.14) и вектора $\vec{x}$, получим:

$$\left(\hat{T}\vec{x}, \vec{y}\right) = a \cdot \left(\vec{x}, \vec{y}\right) - b \cdot \left(\vec{y}, \vec{y}\right),$$

$$\left(\hat{T}\vec{y}, \vec{x}\right) = b \cdot \left(\vec{x}, \vec{x}\right) + a \cdot \left(\vec{y}, \vec{x}\right).$$

Так как пространство E^n вещественно, то

$$\left(\hat{T}\vec{x}, \vec{y}\right) = \left(\vec{x}, \hat{T}\vec{y}\right).$$

Поэтому имеем:

$$a\left(\vec{x}, \vec{y}\right) - b\left(\vec{y}, \vec{y}\right) = b\left(\vec{x}, \vec{x}\right) + a\left(\vec{y}, \vec{x}\right) \Rightarrow$$

$$\Rightarrow b\left(\vec{x}, \vec{x}\right) + b\left(\vec{y}, \vec{y}\right) = 0 \Rightarrow b\left(\left\|\vec{x}\right\|^2 + \left\|\vec{y}\right\|^2\right) = 0.$$

Так как $\left\|\vec{x}\right\|^2 + \left\|\vec{y}\right\|^2 \neq 0$, то $b = 0$ и корень μ вещественный. ••

Доказательство следующей теоремы проводится одинаково для вещественного и комплексного случая.

Теорема 4.2.8. *Пусть* $\hat{T} : E^n \to E^n$ *— самосопряжённый оператор. Тогда в пространстве* E^n *можно выбрать ортонормированный базис* $\left\{ \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n \right\}$ *так, чтобы матрица оператора* $\hat{T}$ *в этом базисе была диагональной.*

Доказательство. Заметим, что если $\vec{x}$ — собственный вектор оператора $\hat{T}$, соответствующий собственному значению μ, то орт этого вектора

$$\vec{e}_x = \left\| \vec{x} \right\|^{-1} \cdot \vec{x}$$

также является собственным вектором оператора $\hat{T}$ с тем же собственным значением μ.

Итак, пусть μ_1 — один из корней характеристического многочлена самосопряжённого оператора $\hat{T}$. По теоремам 4.2.6 и 4.2.7 μ_1 вещественно. Пусть $\vec{e}_1$ — единичный собственный вектор оператора $\hat{T}$, соответствующий собственному значению μ_1, то есть

$$\hat{T}\vec{e}_1 = \mu_1 \vec{e}_1.$$

Обозначим через

$$E^1_{inv} = L\left\{ \vec{e}_1 \right\}$$

— одномерное инвариантное подпространство, порождённое вектором $\vec{e}_1$. Так как $\hat{T} = \hat{T}^*$, то по теореме 4.2.4 ортогональное дополнение $E^{1\,\perp}_{inv}$ также будет инвариантным подпространством оператора $\hat{T}$.

Пусть μ_2 — собственное значение оператора $\hat{T}$ в подпространстве $E^{1\,\perp}_{inv}$, а $\vec{e}_2$ — единичный собственный вектор, соответствующий этому собственному значению, то есть

$$\hat{T}\vec{e}_2 = \mu_2 \vec{e}_2.$$

Обозначим через

$$E_{inv}^2 = L\left\{\vec{e}_1, \vec{e}_2\right\}$$

— инвариантное подпространство, порождённое системой векторов $\left\{\vec{e}_1, \vec{e}_2\right\}$. Его ортогональное дополнение $E_{inv}^{2\perp}$ также будет инвариантным подпространством оператора $\hat{T}$. Продолжая этот процесс, получим ортонормированную систему, состоящую из n собственных векторов $\left\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\right\}$ оператора $\hat{T}$, которая может быть выбрана за базис пространства E^n. В этом базисе матрица оператора $\hat{T}$, очевидно, имеет вид

$$T = \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mu_n \end{pmatrix}. \bullet\bullet$$

Следствие из теоремы 4.2.8. *Все корни характеристического многочлена самосопряжённого оператора различные, а соответствующие собственные векторы ортогональны.*

Д о к а з а т е л ь с т в о. Первое утверждение непосредственно следует из теоремы 4.2.8. Поэтому докажем только второе утверждение.

Пусть $\vec{x}, \vec{y} \in R^n$ — некоторые собственные векторы самосопряжённого оператора $\hat{T}$, соответствующие различным собственным значениям, то есть

$$(\exists\lambda, \mu): \hat{T}\vec{x} = \lambda \cdot \vec{x},\ \hat{T}\vec{y} = \mu \cdot \vec{y},$$

причём $\lambda \neq \mu$. Тогда:

$$\left(\hat{T}\vec{x}, \vec{y}\right) = \left(\lambda \cdot \vec{x}, \vec{y}\right) = \lambda \cdot \left(\vec{x}, \vec{y}\right);$$

$$\left(\vec{x}, \hat{T}\vec{y}\right) = \left(\vec{x}, \mu \cdot \vec{y}\right) = \mu \cdot \left(\vec{x}, \vec{y}\right).$$

Так как оператор самосопряжённый, то

$$\left(\hat{T}\vec{x}, \vec{y}\right) = \left(\vec{x}, \hat{T}\vec{y}\right)$$

и в силу $\lambda \neq \mu$ получаем:

$$(\lambda-\mu)\cdot\left(\vec{x},\vec{y}\right)=0\Rightarrow\left(\vec{x},\vec{y}\right)=0. \bullet\bullet$$

Геометрический смысл ***преобразования пространства под действием самосопряжённого оператора*** $\hat{T}$ вытекает из теоремы 4.2.8. Пусть

$$\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right\}\subset E^n$$

— некоторый базис из теоремы 4.2.8. Тогда

$$\left(\forall\vec{x}\in E^n\right)\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2+\ldots+x_n\vec{e}_n.$$

Действуя на вектор $\vec{x}$ оператором $\hat{T}$, получаем

$$\hat{T}\vec{x}=x_1\mu_1\vec{e}_1+x_2\mu_2\vec{e}_2+\ldots+x_n\mu_n\vec{e}_n.$$

Следовательно, преобразование пространства E^n под действием самосопряжённого оператора $\hat{T}$ сводится к ***гомотетии*** — n растяжениям вдоль координатных осей с коэффициентами $\mu_1,\mu_2,\ldots,\mu_n$ соответственно.

Ортогональные операторы и их свойства. Рассмотрим определение и свойства ещё одного важного класса операторов, действующих в вещественном евклидовом пространстве E^n.

Определение 4.2.5. *Оператор $\hat{T}$, действующий в вещественном евклидовом пространстве E^n, называется **ортогональным**, если он сохраняет скалярное произведение, то есть*

$$\left(\forall\vec{x},\vec{y}\in E^n\right)\left(\hat{T}\vec{x},\hat{T}\vec{y}\right)=\left(\vec{x},\vec{y}\right). \bullet \tag{4.2.15}$$

Из определения следует, что ***ортогональный оператор сохраняет нормы векторов и углы между векторами***.

Лемма 4.2.3. *Оператор $\hat{T}:E^n\to E^n$ является ортогональным в том и только в том случае, если $\hat{T}^*=\hat{T}^{-1}$.*

Д о к а з а т е л ь с т в о. Пусть $\hat{T}$ — ортогональный оператор. Тогда

$$\left(\forall\vec{x},\vec{y}\in E^n\right)\left(\vec{x},\vec{y}\right)=\left(\hat{T}\vec{x},\hat{T}\vec{y}\right)=\left(\vec{x},\hat{T}^*\hat{T}\vec{y}\right),$$

откуда имеем:

$$\left(\vec{x}, \vec{y} - \hat{T}^{*}\hat{T}\vec{y}\right) = 0.$$

Полагая $\vec{x} = \vec{y} - \hat{T}^{*}\hat{T}\vec{y}$, получаем:

$$\left(\vec{y} - \hat{T}^{*}\hat{T}\vec{y}, \vec{y} - \hat{T}^{*}\hat{T}\vec{y}\right) = 0 \Rightarrow \vec{y} - \hat{T}^{*}\hat{T}\vec{y} = \vec{0} \Rightarrow$$

$$\Rightarrow \vec{y} = \hat{T}^{*}\hat{T}\vec{y} \Rightarrow \hat{T}^{*} = \hat{T}^{-1}.$$

Пусть $\hat{T}^{*} = \hat{T}^{-1}$. Тогда имеем:

$$\left(\vec{x}, \vec{y}\right) = \left(\vec{x}, \hat{I}\vec{y}\right) = \left(\vec{x}, \hat{T}^{-1}\hat{T}\vec{y}\right) = \left(\vec{x}, \hat{T}^{*}\hat{T}\vec{y}\right) = \left(\hat{T}\vec{x}, \hat{T}\vec{y}\right). \bullet\bullet$$

Из леммы 4.2.3 следует, что ***ортогональный оператор невырожден***.

Теорема 4.2.9 (о свойствах ортогональных операторов). *Ортогональные операторы* $\hat{T}: E^n \to E^n$ *в вещественном евклидовом пространстве* E^n *обладают следующими свойствами:*

1) *единичный оператор является ортогональным;*

2) *композиция ортогональных операторов также является ортогональным оператором;*

3) *оператор, обратный ортогональному оператору, также является ортогональным;*

4) *если* $\hat{T}: E^n \to E^n$ *— ортогональный оператор, то оператор* $\alpha \cdot \hat{T}$ *является ортогональным в том и только в том случае, если* $\alpha = \pm 1$.

Доказательство. 1) Доказательство этого свойства очевидно:

$$\left(\forall \vec{x}, \vec{y} \in E^n\right) \hat{I}\vec{x} = \vec{x} \wedge \hat{I}\vec{y} = \vec{y} \Rightarrow \left(\vec{x}, \vec{y}\right) = \left(\hat{I}\vec{x}, \hat{I}\vec{y}\right).$$

2) Пусть $\hat{T}: E^n \to E^n$ и $\hat{S}: E^n \to E^n$ — два ортогональных оператора. Тогда:

$$\left(\forall \vec{x}, \vec{y} \in E^n\right)\left(\hat{T}\hat{S}\vec{x}, \hat{T}\hat{S}\vec{y}\right) = \left(\hat{S}\vec{x}, \hat{S}\vec{y}\right) = \left(\vec{x}, \vec{y}\right).$$

3) Пусть $\hat{T}: E^n \to E^n$ — ортогональный оператор, тогда

$$\left(\hat{T}^{-1}\vec{x}, \hat{T}^{-1}\vec{y}\right) = \left(\vec{x}, \left(\hat{T}^{-1}\right)^* \hat{T}^{-1}\vec{y}\right) = \left(\vec{x}, \left(\hat{T}^*\right)^{-1} \hat{T}^{-1}\vec{y}\right) =$$

$$= \left(\vec{x}, \left(\hat{T}^{-1}\right)^{-1} \hat{T}^{-1}\vec{y}\right) = \left(\vec{x}, \hat{T}\hat{T}^{-1}\vec{y}\right) = \left(\vec{x}, \vec{y}\right).$$

4) Пусть $\hat{T}: E^n \to E^n$ — ортогональный оператор. Тогда

$$\left(\alpha \cdot \hat{T}\vec{x}, \alpha \cdot \hat{T}\vec{y}\right) = \alpha^2 \left(\hat{T}\vec{x}, \hat{T}\vec{y}\right) = \alpha^2 \left(\vec{x}, \vec{y}\right) \Rightarrow$$

$$\left(\alpha \cdot \hat{T}\vec{x}, \alpha \cdot \hat{T}\vec{y}\right) = \left(\vec{x}, \vec{y}\right) \Leftrightarrow \alpha = \pm 1. \bullet\bullet$$

Следующая теорема даёт критерий ортогональности оператора.

Теорема 4.2.10. *Оператор* $\hat{T}: E^n \to E^n$ *является ортогональным в том и только в том случае, если он переводит хотя бы один ортонормированный базис снова в ортонормированный базис.*

Доказательство. Пусть $\hat{T}: E^n \to E^n$ — ортогональный оператор. Тогда он, сохраняя скалярное произведение, переводит ортонормированный базис в новый ортонормированный базис.

Пусть теперь оператор $\hat{T}: E^n \to E^n$ переводит ортонормированный базис

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$$

в новый ортонормированный базис $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset E^n$. Тогда

$$\left(\forall \vec{x}, \vec{y} \in E^n\right)$$

$$\hat{T}\vec{x} = \sum_{k=1}^{n} x^k \hat{T}\vec{e}_k = \sum_{k=1}^{n} x^k \vec{a}_k, \quad \hat{T}\vec{y} = \sum_{j=1}^{n} y^j \hat{T}\vec{e}_j = \sum_{j=1}^{n} y^j \vec{a}_j.$$

Откуда

$$\left(\hat{T}\vec{x}, \hat{T}\vec{y}\right) = \sum_{k=1}^{n}\sum_{j=1}^{n} x^k y^j \left(\vec{a}_k, \vec{a}_j\right) = \sum_{k=1}^{n} x^k y^k = \left(\vec{x}, \vec{y}\right). \bullet\bullet$$

Рассмотрим свойства матрицы ортогонального оператора.

Теорема 4.2.11. *Система вектор-столбцов (строк) матрицы ортогонального оператора* $\hat{T}: E^n \to E^n$ *в любом ортонормированном базисе*

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$$

является ортонормированной.

Доказательство. Пусть $\hat{T}: E^n \to E^n$ — некоторый ортогональный оператор и $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$ — некоторый ортонормированный базис. Так как система образов базисных векторов сама является ортонормированной, то $\left(\hat{T}\vec{e}_i, \hat{T}\vec{e}_k\right) = \delta_{ik}$. Поэтому для столбцов матрицы оператора

$$T = \begin{pmatrix} t_1^1 & t_2^1 & \ldots & t_n^1 \\ t_1^2 & t_2^2 & \ldots & t_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ t_1^n & t_2^n & \ldots & t_n^n \end{pmatrix}$$

(как векторов арифметического пространства R^n) имеем:

$$\begin{pmatrix} t_k^1 & t_k^2 & \ldots & t_k^n \end{pmatrix} \begin{pmatrix} t_i^1 \\ t_i^2 \\ \ldots \\ t_i^n \end{pmatrix} = t_k^1 t_i^1 + t_k^2 t_i^2 + \ldots + t_k^n t_i^n = \delta_{ik}. \qquad (4.2.16)$$

Аналогичное свойство справедливо и для строк матрицы T. Действительно, для ортогонального оператора по лемме 4.2.3 сопряжённый оператор $\hat{T}^* = \hat{T}^{-1}$ и по теореме 4.2.9 также является ортогональным. Поэтому, по доказанному выше, столбцы матрицы T^* (строки матрицы T) тоже образуют ортонормированную систему, то есть

$$t_1^k t_1^i + t_2^k t_2^i + \ldots + t_n^k t_n^i = \delta_{ik}. \bullet\bullet \qquad (4.2.17)$$

Теорема 4.2.12. *Матрица ортогонального оператора* $\hat{T}: E^n \to E^n$ *в любом ортонормированном базисе удовлетворяет соотношению*

$$T^T = T^{-1}. \qquad (4.2.18)$$

Д о к а з а т е л ь с т в о. Пусть $\hat{T}: E^n \to E^n$ — ортогональный оператор. Так как матрицы операторов $\hat{T}$ и $\hat{T}^*$ в случае вещественного евклидова пространства связаны соотношениями $(\forall i, k = 1, 2, \ldots, n)\ t_k^i = t_i^{*k}$, то по лемме 4.2.3 получаем: матрица оператора $\hat{T}$ удовлетворяет соотношению (4.2.18).

Обратно пусть выполнено соотношение (4.2.18). Тогда $T^{-1} = T^*$, откуда по лемме 4.2.3 оператор $\hat{T}$ является ортогональным. ••

Определение 4.2.6. *Матрица T, для которой справедливо свойство (4.2.18), называется ортогональной.*•

Приведём некоторые теоремы о свойствах ортогонального оператора.

Теорема 4.2.13. *Пусть оператор $\hat{T}: E^n \to E^n$ является ортогональным и $E_{inv} \subset E^n$ — инвариантное подпространство оператора $\hat{T}$. Тогда ортогональное дополнение $E_{inv}^{\perp}$ также инвариантно относительно $\hat{T}$.*

Д о к а з а т е л ь с т в о. Так как $\hat{T}$ — ортогональный оператор, то по лемме 4.2.3 $\hat{T}^{-1} = \hat{T}^*$. По теореме 4.2.4 подпространство $E_{inv}^{\perp}$ инвариантно относительно сопряжённого оператора $\hat{T}^* = \hat{T}^{-1}$. Но тогда в силу невырожденности ортогонального оператора и по теореме 4.1.5 это подпространство инвариантно и относительно оператора $\left(\hat{T}^{-1}\right)^{-1} = \hat{T}$. ••

Теорема 4.2.14. *Собственные значения ортогонального оператора $\hat{T}: E^n \to E^n$ удовлетворяют условию* $\mu = \pm 1$.

Д о к а з а т е л ь с т в о. Пусть $\hat{T}\vec{x} = \mu\vec{x}$. Тогда

$$\left(\vec{x}, \vec{x}\right) = \left(\hat{T}\vec{x}, \hat{T}\vec{x}\right) = \left(\mu\vec{x}, \mu\vec{x}\right) = \mu^2\left(\vec{x}, \vec{x}\right).$$

Так как по определению $\vec{x} \neq 0$, то $\mu^2 = 1 \Rightarrow \mu = \pm 1$. ••

Теорема 4.2.15. *Если матрица T ортогональная, то* $\det T = \pm 1$.

Д о к а з а т е л ь с т в о. Для ортогональной матрицы выполняется равенство $T^T = T^{-1}$, следовательно, имеем $TT^T = I$. Поэтому

$$\left(\det T\right)^2 = \det T \det T^T = \det\left(TT^T\right) = \det I = 1 \Rightarrow \det T = \pm 1. \bullet\bullet$$

Интерпретацию действия ортогонального оператора даёт следующая теорема.

Теорема 4.2.16. *Пусть* $\hat{T} : E^n \to E^n$ *— ортогональный оператор. Тогда найдётся такой ортонормированный базис* $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$*, в котором матрица оператора* $\hat{T}$ *имеет вид:*

$$T = \begin{pmatrix} T_1^1 & O_2^1 & \ldots & O_k^1 & \ldots & O_m^1 \\ O_1^2 & T_2^2 & \ldots & O_k^2 & \ldots & O_m^2 \\ \multicolumn{6}{c}{\dotfill} \\ O_1^k & O_2^k & \ldots & T_k^k & \ldots & O_m^k \\ \multicolumn{6}{c}{\dotfill} \\ O_1^m & O_2^m & \ldots & O_k^m & \ldots & T_m^m \end{pmatrix}.$$

Здесь «диагональные» клетки с номерами, соответственно, $i = 1, 2, \ldots, k$ имеют вид:

$$T_i^i = \begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix},$$

а с номерами $i = k+1, \ldots, m$ соответственно

$$T_i^i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Остальные клетки — нулевые матрицы.

Доказательство. Докажем теорему по индукции, используя теорему 4.2.13. Отметим сначала, что в случае ортогонального преобразования на прямой, то есть $\hat{T} : E^1 \to E^1$, для вектора $\vec{e} \in E^1$, $\vec{e} \neq \vec{0}$, имеем $\hat{T}\vec{e} = \mu\vec{e}$ и по теореме 4.2.14 $\mu = \pm 1$. Следовательно, матрица оператора имеет вид $T = (\pm 1)$.

1) Пусть $k = 1$. Тогда имеем ортогональное преобразование плоскости E^2. Матрица оператора $\hat{T}$ в некотором ортонормированном базисе имеет вид

$$T = \begin{pmatrix} t_1^1 & t_2^1 \\ t_1^2 & t_2^2 \end{pmatrix},$$

где по теореме 4.2.11 (формулы (4.2.16) и (4.2.17))

$$\left(t_1^1\right)^2+\left(t_1^2\right)^2=1,\ \left(t_2^1\right)^2+\left(t_2^2\right)^2=1,\ t_1^1\cdot t_2^1+t_1^2\cdot t_2^2=0.$$

В силу двух первых равенств существуют такие φ и ψ, что

$$t_1^1=\cos\varphi,\ t_1^2=\sin\varphi \text{ и } t_2^1=\cos\psi,\ t_2^2=\sin\psi.$$

Тогда по третьему равенству имеем:

$$\cos\varphi\cdot\cos\psi+\sin\varphi\cdot\sin\psi=\cos(\psi-\varphi)=0,$$

откуда $\psi-\varphi=\dfrac{\pi}{2}$ или $\psi-\varphi=\dfrac{3\pi}{2}$.

В первом случае

$$t_2^1=\cos\psi=-\sin\varphi,\ t_2^2=\sin\psi=\cos\varphi$$

и матрица имеет вид

$$T_1=\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}. \tag{4.2.19}$$

Во втором случае аналогично имеем:

$$t_2^1=\cos\psi=\cos\left(\frac{3\pi}{2}+\varphi\right)=\sin\varphi,\ t_2^2=\sin\psi=\sin\left(\frac{3\pi}{2}+\varphi\right)=-\cos\varphi,$$

$$T_2=\begin{pmatrix}\cos\varphi & \sin\varphi\\ \sin\varphi & -\cos\varphi\end{pmatrix}. \tag{4.2.20}$$

Матрица (4.2.20) симметрическая, то есть она соответствует самосопряжённому оператору. Тогда по теореме 4.2.8 в некотором ортонормированном базисе (новом) матрица приводится к диагональному виду. Характеристический многочлен равен

$$\det\begin{pmatrix}\cos\varphi-\mu & \sin\varphi\\ \sin\varphi & -\cos\varphi-\mu\end{pmatrix}=\mu^2-\left(\sin^2\varphi+\cos^2\varphi\right)=\mu^2-1.$$

Откуда имеем $\mu_{1,2}=\pm1$. Поэтому по той же теореме 4.2.8 имеем

$$T_2=\begin{pmatrix}\mu_1 & 0\\ 0 & \mu_2\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{4.2.21}$$

2) Допустим, что теорема верна для $k\le m-1$.

3) Покажем, что теорема верна и для $k = m$. Пусть характеристическое уравнение имеет, по крайней мере, один вещественный корень μ_m. Тогда по теореме 4.2.14 $\mu_m = \pm 1$. В этом случае собственный вектор $\vec{x}_m$, соответствующий собственному значению μ_m, принимаем за базисный вектор и в ортогональном дополнении $L^{\perp}\left\{\vec{e}_m\right\}$ приводим матрицу к требуемому виду.

Если теперь характеристическое уравнение имеет только комплексные корни и $a + ib$ — один из них, то по лемме 4.2.2 для этого корня имеется двумерное инвариантное подпространство $L^2_{inv} = L\left\{\vec{u}, \vec{v}\right\}$ оператора $\hat{T}$. Выбирая в этом подпространстве ортонормированный базис, получим, что $L^{2\ \perp}_{inv}$ также инвариантное подпространство оператора $\hat{T}$. Поэтому матрица оператора имеет клеточный вид:

$$T = \begin{pmatrix} t_1^1 & t_2^1 & O_2^1 \\ t_1^2 & t_2^2 & \\ O_1^2 & & T_2^2 \end{pmatrix}.$$

Далее имеем:

$$\left(t_1^1\right)^2 + \left(t_2^1\right)^2 = 1, \left(t_1^2\right)^2 + \left(t_2^2\right)^2 = 1, t_1^1 \cdot t_1^2 + t_2^1 \cdot t_2^2 = 0.$$

Полагая $t_1^1 = \cos\varphi$, $t_2^1 = -\sin\varphi$, получаем

$$\begin{pmatrix} t_1^1 & t_2^1 \\ t_1^2 & t_2^2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \pm\sin\varphi & \pm\cos\varphi \end{pmatrix}.$$

Так как характеристическое уравнение

$$\det\begin{pmatrix} \cos\varphi - \mu & -\sin\varphi \\ -\sin\varphi & -\cos\varphi - \mu \end{pmatrix} = 0 \Rightarrow \mu^2 - 1 = 0$$

имеет вещественные корни, то предположение об их отсутствии неверно. Комплексных корней нет. Следовательно, остаётся случай

$$\begin{pmatrix} t_1^1 & t_2^1 \\ t_1^2 & t_2^2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

рассмотренный выше. Приводя клетку T_2^2 к требуемому виду, что возможно по предположению индукции, получаем доказательство теоремы. ••

Геометрическая интерпретация ортогональных преобразований вытекает из доказательства теоремы 4.2.16. Действительно, как отмечено выше, ортогональный оператор $\hat{T}: E^1 \to E^1$ является или тождественным оператором, или оператором отражения относительно центра O. Для установления геометрического смысла оператора $\hat{T}: E^2 \to E^2$ приведём решение следующего примера.

Пример 4.2.1. В пространстве R^2 найти матрицы линейных операторов:

1) поворота на угол φ против часовой стрелки;

2) зеркального отражения относительно оси OX^1.

Р е ш е н и е. Найдём, сначала, матрицу оператора поворота на конечный угол φ против часовой стрелки. Для этого воспользуемся геометрическим смыслом элементов матрицы оператора и рис. 4.1.

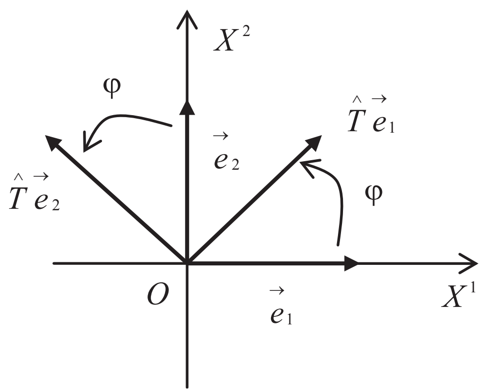


Рис. 4.1

Действие оператора поворота на базисные векторы декартовой системы координат задаётся формулами:

$$\hat{T}\vec{e}_1 = \cos\varphi \cdot \vec{e}_1 + \sin\varphi \cdot \vec{e}_2, \hat{T}\vec{e}_2 = -\sin\varphi \cdot \vec{e}_1 + \cos\varphi \cdot \vec{e}_2.$$

И для матрицы оператора поворота получаем:

$$T = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Действие оператора зеркального отражения относительно оси OX^1 определяется формулами:

$$\hat{R}\vec{e}_1 = \vec{e}_1, \ \hat{R}\vec{e}_2 = -\vec{e}_2.$$

Поэтому его матрица имеет вид:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \otimes$$

Теперь видно, что клетка (4.2.19) является матрицей оператора поворота плоскости R^2 на конечный угол φ против часовой стрелки. Аналогично, клетка (4.2.20) или, что то же самое, (4.2.21), является матрицей оператора зеркального отражения относительно оси OX^2 в плоскости R^2.

Унитарные операторы*). Унитарные операторы являются аналогом ортогональных операторов в случае комплексного евклидова (унитарного) пространства.

Определение 4.2.7. *Оператор $\hat{T} : U^n \to U^n$, действующий в комплексном евклидовом (унитарном) пространстве U^n, называется* ***унитарным****, если он сохраняет скалярное произведение, то есть*

$$\left(\forall \vec{x}, \vec{y} \in U^n\right)\left(\hat{T}\vec{x}, \hat{T}\vec{y}\right) = \left(\vec{x}, \vec{y}\right) \bullet \qquad (4.2.22)$$

Унитарный оператор сохраняет норму вектора и переводит ортогональные векторы в ортогональные векторы. Таким образом, унитарный оператор каждый ортонормированный базис переводит в ортонормированный базис. Справедлива теорема, аналогичная теореме 4.2.10.

Теорема 4.2.17. *Оператор $\hat{T} : U^n \to U^n$ является унитарным в том и только в том случае, если он переводит хотя бы один ортонормированный базис в ортонормированный базис.*

Свойства 1–3 ортогональных операторов из теоремы 4.2.9 переносятся на унитарные операторы без изменений. С естественными изменениями справедливо и свойство 4.

Теорема 4.2.18. *Если $\hat{T} : U^n \to U^n$ — унитарный оператор, то оператор $\alpha \cdot \hat{T}$, где $\alpha \in C$, является унитарным в том и только в том случае, если $|\alpha| = 1$.*

Д о к а з а т е л ь с т в о. Действительно,

$$\left(\alpha\hat{T}\vec{x}, \alpha\hat{T}\vec{y}\right) = \alpha\bar{\alpha}\left(\hat{T}\vec{x}, \hat{T}\vec{y}\right) = |\alpha|^2\left(\vec{x}, \vec{y}\right),$$

откуда $|\alpha|^2 = 1 \Rightarrow |\alpha| = 1$. ••

Пусть

$$T = \begin{pmatrix} t_1^1 & t_2^1 & \dots & t_n^1 \\ t_1^2 & t_2^2 & \dots & t_n^2 \\ \dots & \dots & \dots & \dots \\ t_1^n & t_2^n & \dots & t_n^n \end{pmatrix} \tag{4.2.23}$$

— матрица унитарного оператора $\hat{T}: U^n \to U^n$ в некотором ортонормированном базисе $\left\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\right\} \subset U^n$. В силу определения унитарного оператора образы базисных векторов ортогональны, то есть

$$\left(\hat{T}\vec{e}_i, \hat{T}\vec{e}_k\right) = \delta_{ik}. \tag{4.2.24}$$

Так как столбцы матрицы (4.2.23) составлены из координат разложений образов базисных векторов по векторам базиса, то получаем, что они также ортогональны:

$$\overline{t_i^1} \cdot t_k^1 + \overline{t_i^2} \cdot t_k^2 + \dots + \overline{t_i^n} \cdot t_k^n = \delta_{ik}. \tag{4.2.25}$$

Справедливо утверждение, что *если* $\hat{T}: U^n \to U^n$ *— унитарный оператор, то* $\hat{T}^* = \hat{T}^{-1}$*, причём оператор* $\hat{T}^* = \hat{T}^{-1}$ *также унитарный*. Поэтому для строк матрицы унитарного оператора справедливо свойство, аналогичное (4.2.25):

$$\overline{t_1^i} \cdot t_1^k + \overline{t_2^i} \cdot t_2^k + \dots + \overline{t_n^i} \cdot t_n^k = \delta_{ik}. \tag{4.2.26}$$

Из формул (4.2.25) и (4.2.26) имеем:

$$\left|t_1^k\right|^2 + \left|t_2^k\right|^2 + \dots + \left|t_n^k\right|^2 = 1; \left|t_k^1\right|^2 + \left|t_k^2\right|^2 + \dots + \left|t_k^n\right|^2 = 1. \tag{4.2.27}$$

Определение 4.2.8. *Матрица называется унитарной, если* $T^* = T^{-1}$. •

Получаем следующее утверждение.

Теорема 4.2.19. *Оператор* $\hat{T}: U^n \to U^n$ *является унитарным в том и только в том случае, если в каждом ортонормированном базисе*

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset U^n$$

его матрица является унитарной.

На унитарные операторы переносится теорема 4.2.13.

Теорема 4.2.20. *Если U_{inv} — инвариантное подпространство унитарного оператора $\hat{T}: U^n \to U^n$, то ортогональное дополнение $U_{inv}{}^{\perp}$ также является инвариантным подпространством оператора $\hat{T}$.*

Теорема 4.2.14 принимает следующий вид.

Теорема 4.2.21. *Собственные значения унитарного оператора $\hat{T}: U^n \to U^n$ по модулю равны 1.*

Доказательство. Так как $\hat{T}$ унитарный оператор, то по определению

$$\left(\forall \vec{x} \in U^n\right): \hat{T}\vec{x} = \mu \cdot \vec{x}$$

и

$$\left(\hat{T}\vec{x}, \hat{T}\vec{x}\right) = \left(\mu \cdot \vec{x}, \mu \cdot \vec{x}\right) = \mu\bar{\mu}\left(\vec{x}, \vec{x}\right) = |\mu|^2\left(\vec{x}, \vec{x}\right).$$

Так как $\vec{x} \neq \vec{0}$, то $|\mu|^2 = 1 \Rightarrow |\mu| = 1$. ••

Обозначая $\mu = \alpha + i\beta$, имеем $|\mu|^2 = \alpha^2 + \beta^2 = 1$. Учитывая геометрическую интерпретацию комплексных чисел, получаем, *что* ***спектр унитарного оператора находится на единичной окружности в комплексной плоскости с центром в нуле***.

Теорема 4.2.22. *Пусть $\hat{T}: U^n \to U^n$ — унитарный оператор. Тогда существует ортонормированный базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset U^n$, в котором матрица T оператора $\hat{T}$ приводится к диагональному виду.*

Доказательство. Пусть μ_1 — одно из собственных значений унитарного оператора $\hat{T}$. По теореме 4.2.21 $|\mu_1| = 1$. Орт собственного вектора, соответствующего собственному значению μ_1, обозначим $\vec{e}_{\mu_1}$. Тогда

$$\hat{T}\vec{e}_{\mu_1} = \mu_1 \cdot \vec{e}_{\mu_1}.$$

Одномерное подпространство $L\left\{\vec{e}_{\mu_1}\right\}$ инвариантно относительно оператора $\hat{T}$. Ортогональное дополнение $L^{\perp}\left\{\vec{e}_{\mu_1}\right\}$ также инвариантно относительно $\hat{T}$.

Если μ_2, где $|\mu_2| = 1$ — собственное значение унитарного оператора $\hat{T}$ в $L^{\perp}\left\{\vec{e}_{\mu_1}\right\}$, а $\vec{e}_{\mu_2}$ — орт соответствующего собственного вектора, то

$$\hat{T}\vec{e}_{\mu_2} = \mu_2 \cdot \vec{e}_{\mu_2}.$$

Обозначим $L^2_{inv}\left\{\vec{e}_{\mu_1}, \vec{e}_{\mu_2}\right\}$ — инвариантное подпространство оператора $\hat{T}$, порождённое векторами $\left\{\vec{e}_{\mu_1}, \vec{e}_{\mu_2}\right\}$. Тогда подпространство $L^{2\ \perp}_{inv}\left\{\vec{e}_{\mu_1}, \vec{e}_{\mu_2}\right\}$ также будет инвариантным относительно оператора $\hat{T}$. Продолжая этот процесс, построим флаг инвариантных относительно оператора $\hat{T}$ подпространств, порождённых n попарно ортогональными собственными векторами оператора

$$\left\{\vec{e}_{\mu_1}, \vec{e}_{\mu_2}, \ldots, \vec{e}_{\mu_n}\right\}.$$

В этом базисе матрица оператора $\hat{T}$ имеет диагональный вид

$$T = \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu_n \end{pmatrix}.$$

Все элементы этой матрицы, стоящие на главной диагонали, равны по модулю 1, то есть $(\forall k = 1, 2, \ldots, n)\ |\mu_k| = 1$. ••

Следствие из теоремы 4.2.22. *Пусть $\hat{T} : U^n \to U^n$ — унитарный оператор. Тогда определитель его матрицы в любом базисе по модулю равен 1.*

Доказательство. Доказательство следует из теоремы 4.2.22 и инвариантности определителя матрицы при изменении базиса. ••

Общие свойства операторов в евклидовых пространствах*). Рассмотрим некоторые общие свойства и типы линейных операторов в евклидовых пространствах.

Определение 4.2.9. *Эрмитов (самосопряжённый) оператор*

$$\hat{T}: U^n \to U^n$$

называется ***неотрицательным (положительно определённым)****, если*

$$\left(\forall \vec{x} \in U^n\right): \vec{x} \neq \vec{0}$$

выполняется неравенство $\left(\hat{T}\vec{x}, \vec{x}\right) \geq 0 \left(\left(\hat{T}\vec{x}, \vec{x}\right) > 0\right)$. •

Теорема 4.2.23. *Эрмитов оператор* $\hat{T}: U^n \to U^n$ *является неотрицательным (положительно определённым) в том и только в том случае, если все его собственные значения неотрицательны (положительны).*

Доказательство. По теореме 4.2.8 всегда можно выбрать в пространстве U^n ортонормированный базис из собственных векторов эрмитова оператора $\hat{T}$, пусть это будет базис $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$. Тогда $\left(\forall \vec{x} \in U^n\right)$ из разложения

$$\vec{x} = x^1\vec{a}_1 + x^2\vec{a}_2 + \ldots + x^n\vec{a}_n$$

следует (все μ_k вещественны!)

$$\left(\hat{T}\vec{x}, \vec{x}\right) = \mu_1\left|x^1\right|^2 + \mu_2\left|x^2\right|^2 + \ldots + \mu_n\left|x^n\right|^2.$$

Очевидно, что если все собственные значения неотрицательные (положительные), то $\left(\hat{T}\vec{x}, \vec{x}\right) \geq 0$. Следовательно, оператор $\hat{T}$ — неотрицательный (положительный).

Пусть теперь $\vec{x} = \vec{a}_k$. Тогда

$$\left(\forall k = 1, 2, \ldots, n\right)\left(\hat{T}\vec{a}_k, \vec{a}_k\right) = \mu_k.$$

Очевидно, что $\left(\forall k = 1, 2, \ldots, n\right)\left(\hat{T}\vec{a}_k, \vec{a}_k\right) \geq 0 \Rightarrow \mu_k \geq 0$. ••

Следствие из теоремы 4.2.23. *Положительно определённый оператор является невырожденным неотрицательным оператором.*

Приведём теорему о некоторых свойствах неотрицательных и положительно определённых операторов.

Теорема 4.2.24. *Пусть* $\hat{T}: U^n \to U^n$ *и* $\hat{S}: U^n \to U^n$ *— положительно определённые (неотрицательные) операторы, тогда:*

1) *оператор* $\alpha\hat{T}+\beta\hat{S}$ *положительно определён при любых неотрицательных числах* α *и* β*, не равных нулю одновременно;*

2) *оператор* $\hat{T}^{-1}$ *также положительно определён;*

3) *для любого невырожденного оператора* $\hat{A}: U^n \to U^n$ *операторы* $\hat{A}^*\hat{T}\hat{A}$ *и* $\hat{A}\hat{T}\hat{A}^*$ *положительно определены;*

4) *для любого неотрицательного оператора* $\hat{T}: U^n \to U^n$ *существует такой неотрицательный оператор* $\hat{S}: U^n \to U^n$*, что* $\hat{S}^2 = \hat{T}$.

Доказательство. 1) Из третьего и четвёртого свойств (4.2.10) сопряжённого оператора следует, что оператор $\alpha\hat{T}+\beta\hat{S}$ является эрмитовым при любых действительных числах α и β. Если эти числа неотрицательны, то

$$\left(\forall \vec{x} \in U^n\right): \vec{x} \neq \vec{0}$$

имеем

$$\left(\left(\alpha\hat{T}+\beta\hat{S}\right)\vec{x}, \vec{x}\right) = \alpha\left(\hat{T}\vec{x}, \vec{x}\right) + \beta\left(\hat{S}\vec{x}, \vec{x}\right) > 0.$$

2) Так как $\hat{T} = \hat{T}^{\bullet}$, то

$$\hat{T}^{-1} = \left(\hat{T}^{\bullet}\right)^{-1} = \left(\hat{T}^{-1}\right)^{*},$$

то есть оператор $\hat{T}^{-1}$ является эрмитовым. Собственные значения оператора $\hat{T}^{-1}$ являются обратными величинами по отношению к собственным значениям оператора $\hat{T}$. Поэтому они положительны и оператор $\hat{T}^{-1}$ — положительно определённый.

3) Операторы $\hat{A}^{\bullet}\hat{T}\hat{A}$ и $\hat{A}\hat{T}\hat{A}^{\bullet}$ положительны. Действительно, так как оператор $\hat{A}$ невырожден, то

$$\left(\forall \vec{x} \in U^n\right): \vec{x} \neq \vec{0} \text{ образы } \hat{A}\vec{x} \neq \vec{0} \text{ и } \hat{A}^{*}\vec{x} \neq \vec{0}.$$

Поэтому $\left(\forall \vec{x} \in U^n\right): \vec{x} \neq \vec{0}$ имеем

$$\left(\hat{A}^{\bullet}\hat{T}\hat{A}\vec{x}, \vec{x}\right) = \left(\hat{T}\hat{A}\vec{x}, \hat{A}\vec{x}\right) > 0$$

и

$$\left(\hat{A}\hat{T}\hat{A}^{\bullet}\vec{x}, \vec{x}\right) = \left(\hat{T}\hat{A}^{\bullet}\vec{x}, \hat{A}^{\bullet}\vec{x}\right) > 0.$$

Таким образом, для любого невырожденного оператора $\hat{A}: U^n \to U^n$ операторы $\hat{A}^{\bullet}\hat{T}\hat{A}$ и $\hat{A}\hat{T}\hat{A}^{\bullet}$ положительно определенные. Если оператор $\hat{A}: U^n \to U^n$ — вырожденный, то $\hat{A}^{\bullet}\hat{T}\hat{A}$ и $\hat{A}\hat{T}\hat{A}^{\bullet}$ неотрицательные операторы.

4) Пусть $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ — ортонормированная система собственных векторов неотрицательного оператора $\hat{T}: U^n \to U^n$, соответствующих различным неотрицательным собственным значениям $\mu_1, \mu_2, \ldots, \mu_n$ (такие векторы существуют по следствию из теоремы 4.2.8). Тогда

$$\left(\forall k = 1, 2, \ldots, n\right) \hat{T}\vec{a}_k = \mu_k \cdot \vec{a}_k.$$

Оператор $\hat{S}$ определим равенством

$$\hat{S}\vec{a}_k = \sqrt{\mu_k}\,\vec{a}_k.$$

Этот оператор неотрицательный, так как он имеет базисную ортонормированную систему собственных векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$, соответствующих неотрицательным собственным значениям $\sqrt{\mu_1}, \sqrt{\mu_2}, \ldots, \sqrt{\mu_n}$. Далее имеем

$$\hat{S}^2\vec{a}_k = \mu_k \cdot \vec{a}_k = \hat{T}\vec{a}_k.$$

Таким образом операторы $\hat{S}^2$ и $\hat{T}$ совпадают на векторах базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$, а поэтому они равны: $\hat{S}^2 = \hat{T}$.

Все собственные векторы операторов $\hat{T}$ *и* $\hat{S}$ *совпадают*. Действительно, предположим, что $\mu_1, \mu_2, \ldots, \mu_m$ и $\sqrt{\mu_1}, \sqrt{\mu_2}, \ldots, \sqrt{\mu_m}$ — попарно различные собственные значения операторов $\hat{T}$ и $\hat{S}$. Если обозначить через X_k и Y_k $(k = 1, 2, \ldots, m)$ собственные подпространства операторов $\hat{T}$ и $\hat{S}$, содержащие, соответственно, все собственные векторы, соответствующие собственным значениям $\mu_1, \mu_2, \ldots, \mu_n$ и $\sqrt{\mu_1}, \sqrt{\mu_2}, \ldots, \sqrt{\mu_m}$, то прямые суммы собственных подпространств совпадают со всем пространством и, следовательно, имеем

$$\dim X_1 + \dim X_2 + \ldots + \dim X_m = \dim Y_1 + \dim Y_2 + \ldots + \dim Y_m.$$

Так как

$$(\forall k = 1, 2, \ldots, m)\, Y_k \subset X_k,$$

то

$$\dim Y_k \le \dim X_k.$$

Поэтому последнее равенство может иметь место лишь в случае, когда

$$(\forall k = 1, 2, \ldots, m) \dim Y_k = \dim X_k,$$

то есть $Y_k = X_k$.••

Определение 4.2.10. *Пусть* $\hat{T} : U^n \to U^n$ *— неотрицательный оператор. Тогда неотрицательный оператор* $\hat{S}$*, такой, что*

$$\hat{S}^2 = \hat{T},$$

называется ***арифметическим квадратным*** *корнем из оператора* $\hat{T}$. •

Пусть, как и прежде, $\hat{T} : U^n \to \hat{T}(U^n)$ — произвольный линейный оператор, действующий в комплексном евклидовом (унитарном) пространстве. Тогда в пространстве U^n определены операторы $\hat{T}^* \hat{T}$ и $\hat{T} \hat{T}^*$. Из первого и пятого свойств (4.2.10) сопряжённого оператора следует, ***что операторы*** $\hat{T}\hat{T}^*$ ***и*** $\hat{T}^*\hat{T}$ ***являются эрмитовыми***. ***Операторы*** $\hat{T}\hat{T}^*$ ***и*** $\hat{T}^*\hat{T}$ ***неотрицательные***, так как

$$\left(\forall \vec{x}\in U^n\right)\left(\hat{T}^*\hat{T}\vec{x},\vec{x}\right)=\left(\hat{T}\vec{x},\hat{T}\vec{x}\right)\geq 0.$$

Поэтому в пространстве U^n по четвёртому утверждению теоремы 4.2.24 существует единственный неотрицательный оператор $\hat{S}$, такой, что $\hat{T}^*\hat{T}=\hat{S}^2$. В силу того, что оператор $\hat{T}^*\hat{T}$ является эрмитовым, у него имеется ортонормированная система собственных векторов (теорема 4.2.8) $\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n\right\}$.

Теорема 4.2.25. *Ортонормированная система собственных векторов оператора* $\hat{T}^*\hat{T}$ *переводится оператором* $\hat{T}$ *в некоторую ортогональную систему.*

Доказательство. Пусть $\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n\right\}$ — *система собственных векторов, причём* $\hat{T}^*\hat{T}\vec{a}_k=\mu_k^2\cdot\vec{a}_k$, где $\left(\forall k=1,2,\ldots,n\right)\ \mu_k\geq 0$. Тогда

$$\left(\hat{T}\vec{a}_k,\hat{T}\vec{a}_p\right)=\left(\hat{T}^*\hat{T}\vec{a}_k,\vec{a}_p\right)=\mu_k^2\left(\vec{a}_k,\vec{a}_p\right)=\mu_k^2\delta_{kp}.\ \bullet\bullet \qquad (4.2.28)$$

Теорема 4.2.26. *Если* $\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n\right\}$ — *система собственных векторов оператора* $\hat{T}^*\hat{T}$, *соответствующих собственным значениям* $\{\mu_1,\mu_2,\ldots,\mu_n\}$, *то ненулевой вектор* $\hat{T}\vec{a}_k$ *является собственным вектором оператора* $\hat{T}\hat{T}^*$, *соответствующим собственному значению* μ_k^2.

Доказательство. Действительно, в соответствии с равенством (4.2.28) имеем

$$\hat{T}\hat{T}^*\left(\hat{T}\vec{a}_k\right)=\hat{T}\left(\hat{T}^*\hat{T}\vec{a}_k\right)=\hat{T}\left(\mu_k^2\vec{a}_k\right)=\mu_k^2\hat{T}\vec{a}_k.\ \bullet\bullet$$

Из этой теоремы следует, что все ненулевые собственные значения оператора $\hat{T}^*\hat{T}$ являются собственными значениями оператора $\hat{T}\hat{T}^*$

и наоборот. Таким образом, ***ненулевые собственные значения операторов*** $\hat{T}^*\hat{T}$ ***и*** $\hat{T}\hat{T}^*$ ***всегда совпадают***.

Рассмотрим теперь представления произвольного оператора в виде суммы эрмитовых операторов и произведения положительно определённого и унитарного операторов.

Теорема 4.2.27. *Пусть* $\hat{T}:U^n \to \hat{T}\left(U^n\right)$ *— произвольный линейный оператор, действующий в комплексном евклидовом (унитарном) пространстве. Тогда справедливо представление* $\hat{T}=\hat{A}+i\,\hat{B}$*, где* $\hat{A}$ *и* $\hat{B}$ *— эрмитовы операторы.*

Доказательство. Пусть такое представление справедливо, тогда получаем:

$$\hat{T}^*=\hat{A}^*+\left(i\,\hat{B}\right)^*=\hat{A}^*-i\,\hat{B}^*=\hat{A}-i\,\hat{B},$$

так как $\hat{A}$ и $\hat{B}$ — эрмитовы операторы, то есть $\hat{A}^*=\hat{A}$ и $\hat{B}^*=\hat{B}$. Из равенств

$$\hat{T}=\hat{A}+i\,\hat{B},\ \hat{T}^*=\hat{A}-i\,\hat{B}$$

имеем:

$$\hat{A}=\frac{1}{2}\left(\hat{T}+\hat{T}^*\right),\ \hat{B}=\frac{i}{2}\left(\hat{T}^*-\hat{T}\right).$$

Очевидно, операторы $\hat{A}$ и $\hat{B}$ действительно эрмитовы и $\hat{T}=\hat{A}+i\,\hat{B}$. ••

Теорема 4.2.28. *Каждый невырожденный линейный оператор* $\hat{T}:U^n\to U^n\left(\hat{T}:E^n\to E^n\right)$ *можно представить в виде* $\hat{T}=\hat{G}\,\hat{S}$*, где* $\hat{S}$ *— эрмитов (самосопряжённый) положительно определённый оператор, а* $\hat{G}$ *— унитарный (ортогональный) оператор.*

Доказательство. Пусть утверждение теоремы справедливо, то есть $\hat{T}=\hat{G}\,\hat{S}$. Тогда (лемма 4.2.3 и теорема 4.2.19) $\hat{G}^*\hat{G}=\hat{I}$ и

$$\hat{T}^*\hat{T}=\left(\hat{G}\,\hat{S}\right)^*\hat{G}\,\hat{S}=\hat{S}^*\hat{G}^*\hat{G}\,\hat{S}=\hat{S}^*\hat{S}=\hat{S}^2.$$

Так как оператор $\hat{T}^*\hat{T}$ эрмитов (самосопряжённый), то его матрица по теореме 4.2.8 приводится к диагональному виду

$$T^*T = \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mu_n \end{pmatrix},$$

где в силу положительной определённости оператора $\hat{T}^* \hat{T}$ все μ_k положительны. Тогда по доказанному выше арифметический квадратный корень $\hat{S}$ из оператора $\hat{T}^* \hat{T}$ имеет матрицу

$$S = \begin{pmatrix} \sqrt{\mu_1} & 0 & \dots & 0 \\ 0 & \sqrt{\mu_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sqrt{\mu_n} \end{pmatrix},$$

причём оператор $\hat{S}$ неотрицательный. Таким образом, $\hat{S}^2 = \hat{T}^* \hat{T}$. Если положить $\hat{T} = \hat{G}\hat{S}$, то $\hat{G} = \hat{T}\hat{S}^{-1}$. Легко видеть, что оператор $\hat{G}$ является унитарным. Действительно:

$$\hat{G}^* \hat{G} = \left(\hat{T}\hat{S}^{-1}\right)^* \hat{T}\hat{S}^{-1} = \left(\hat{S}^*\right)^{-1} \hat{T}^* \hat{T}\hat{S}^{-1} = \hat{S}^{-1} \hat{S}^2 \hat{S}^{-1} = \hat{I}.$$

Здесь использован факт того, что оператор $\hat{G}$ является унитарным в том и только в том случае, если $\hat{G}^* = \hat{G}^{-1}$. Теорема доказана. ••

4.3. Операторы проектирования *)

Прямая сумма линейных операторов. Покажем, что любой линейный оператор $\hat{T} : X \to X$, где X — некоторое векторное пространство, можно сконструировать из линейных операторов специального вида, все собственные значения которых одинаковы.

Предположим, что векторное пространство X представлено в виде прямой суммы двух своих подпространств $L \subset X$ и $M \subset X$, то есть

$$X = L \oplus M.$$

Тогда по определению для любого вектора $\vec{x} \in X$ справедливо единственное разложение вида

$$\vec{x} = \vec{x}_L + \vec{x}_M,$$

где $\vec{x}_L \in L, \vec{x}_M \in M$.

Определение 4.3.1. *Пусть* $X = L \oplus M$ и $\hat{B} : L \to L, \hat{C} : M \to M$ *— некоторые линейные операторы. Тогда оператор* $\hat{T}$*, действие которого определено правилом*

$$\left(\forall \vec{x} \in X\right) : \vec{x} = \vec{x}_L + \vec{x}_M, \left(\vec{x}_L \in L, \vec{x}_M \in M\right),$$

$$\hat{T}\vec{x} = \hat{B}\vec{x}_L + \hat{C}\vec{x}_M, \tag{4.3.1}$$

*называется **прямой суммой операторов** $\hat{B}$ **и** $\hat{C}$.* •

Нетрудно показать, что оператор $\hat{T}$ из определения 4.3.1 является линейным оператором, действующим на пространстве *X*, причём представление (4.3.1) единственно.

Лемма 4.3.1. *Пусть* $\hat{T} : X \to X$ *— некоторый произвольный линейный оператор. Тогда, если* $X = L \oplus M$*, где* $L \subset X$ *и* $M \subset X$ *— инвариантные подпространства оператора* $\hat{T}$*, то оператор* $\hat{T}$ *можно представить в виде прямой суммы некоторых линейных операторов, действующих на подпространствах L и M соответственно.*

Доказательство. Для доказательства рассмотрим индуцированные операторы $\hat{T}|L$ и $\hat{T}|M$.

Так как $\left(\forall \vec{x} \in X\right)$ имеет место единственное разложение вида

$$\vec{x} = \vec{x}_L + \vec{x}_M,$$

где $\vec{x}_L \in L, \vec{x}_M \in M$, то справедливо представление

$$\hat{T}\vec{x} = \left(\hat{T}|L\right)\vec{x}_L + \left(\hat{T}|M\right)\vec{x}_M.$$

Причём по условию теоремы *L* и *M* — инвариантные подпространства оператора $\hat{T}$ и, следовательно, образы

$$\left(\hat{T}\middle|L\right)\vec{x}_L \in L \wedge \left(\hat{T}\middle|M\right)\vec{x}_M \in M,$$

что и доказывает утверждение теоремы. ••

Характеристический многочлен $T(\mu)$ оператора $\hat{T}$ в этом случае представляется в виде произведения характеристических многочленов индуцированных операторов, то есть справедлива следующая формула:

$$T(\mu) = \left(T\middle|L\right)(\mu)\cdot\left(T\middle|M\right)(\mu). \tag{4.3.2}$$

Имеет место теорема, обеспечивающая возможность представления ***произвольного*** линейного оператора $\hat{T}$ в виде прямой суммы (4.3.1).

Теорема 4.3.1. *Произвольный линейный оператор $\hat{T}: X \to X$ можно разложить в прямую сумму линейных операторов вида (4.3.1) при помощи любого операторного многочлена $F\left(\hat{T}\right)$.*

Д о к а з а т е л ь с т в о. Для доказательства рассмотрим ядро

$$K_m \equiv K\left(F^m\left(\hat{T}\right)\right)$$

степени $F^m\left(\hat{T}\right)$ операторного многочлена $F\left(\hat{T}\right)$. Ядро K_m является, очевидно, инвариантным относительно оператора $\hat{T}$ подпространством пространства X. Кроме этого, понятно, что $K_1 \subset K_2 \subset \ldots$.

Покажем, что

$$(\exists m \in N): K_m = K_{m+1} \Rightarrow (\forall p > m)\ K_m = K_p.$$

Так как

$$\left(\forall \vec{x} \in K_p\right) F^p\left(\hat{T}\right)\vec{x} = \vec{0},$$

получаем следующее равенство:

$$F^p\left(\hat{T}\right)\vec{x} = F^{m+1}\left(\hat{T}\right)\left(F^{p-m-1}\left(\hat{T}\right)\vec{x}\right) = \vec{0}.$$

Из этого равенства следует, что вектор $F^{p-m-1}\left(\hat{T}\right)\vec{x}$ принадлежит K_{m+1}. Но так как $K_m = K_{m+1}$, то $F^{p-m-1}\left(\hat{T}\right)\vec{x} \in K_m$. Поэтому:

$$F^m\left(\hat{T}\right)\left(F^{p-m-1}\left(\hat{T}\right)\vec{x}\right) = F^{p-1}\left(\hat{T}\right)\vec{x} = \vec{0},$$

откуда имеем $\vec{x} \in K_{p-1}$. Теперь справедливость утверждения

$$K_m = K_{m+1} \Rightarrow (\forall p > m)\; K_m = K_p$$

устанавливается индукцией по p:

$$\vec{x} \in K_{p-1} \Rightarrow \vec{x} \in K_{p-2} \Rightarrow \ldots \Rightarrow \vec{x} \in K_m.$$

Итак, получаем, что $K_m = K_p$.

Так как пространство X конечномерное, размерности ядер K_m операторов $F^m\left(\hat{T}\right)$ не могут возрастать неограниченно. Обозначая через q наименьшее целое положительное число, для которого $K_q = K_{q+1}$, предположим, что пересечение множества значений $S_q = F^q\left(\hat{T}\right)(X)$ и ядра K_q оператора $F^q\left(\hat{T}\right)$ не пустое множество, то есть

$$\left(\exists \vec{x} \in X\right) : \vec{x} \in K_q \cap S_q.$$

Получаем, что

$$\left(\exists \vec{z} \in X\right) : \vec{x} = F^q\left(\hat{T}\right)\vec{z},$$

причём

$$F^q\left(\hat{T}\right)\vec{x} = \vec{0}.$$

Откуда следует

$$F^{2q}\left(\hat{T}\right)\vec{z} = \vec{0},$$

то есть $\vec{z} \in K_{2q}$. По доказанному выше имеем $K_q = K_{2q}$. Следовательно, $\vec{z} \in K_q$, то есть

$$\vec{x} = F^q\left(\hat{T}\right)\vec{z} = \vec{0}.$$

Итак, пересечение $K_q \cap S_q$ содержит только нуль-вектор. А это и означает, что $X = K_q \oplus S_q$. Так как подпространства K_q и S_q — инвариантные подпространства относительно оператора $\hat{T}$, справедливость его разложения в прямую сумму операторов, индуцированных на K_q и S_q, следует из леммы 4.3.1. ●●

Из доказанной теоремы вытекает следствие.

Следствие из теоремы 4.3.1. *Каждый (ни один) из корней характеристического многочлена оператора, индуцированного на подпространстве K_q (на подпространстве S_q), является (не является) корнем многочлена $F(z)$.*

Доказательство. Все собственные векторы оператора $\hat{T}$ находятся в подпространствах K_q и S_q. Причём в K_q находятся собственные векторы, соответствующие собственным значениям, совпадающим с какими-либо корнями многочлена F, а в S_q находятся собственные векторы, соответствующие собственным значениям, не совпадающим ни с одним из корней многочлена F. Доказательство теперь следует из того, что каждому собственному значению соответствует, по крайней мере, один собственный вектор. ●●

Прежде чем доказать возможность разложения оператора в прямую сумму, полученную с помощью характеристических многочленов, докажем следующую лемму.

Лемма 4.3.2. *Пусть L — некоторое произвольное инвариантное подпространство оператора $\hat{T}: X \to X$. Если все собственные значения индуцированного оператора $\hat{T}|L$ являются корнями многочлена $F(z)$, то, начиная с достаточно больших целых положительных k, подпространство L является подмножеством ядра каждого из операторов $F^k\left(\hat{T}\right)$.*

Доказательство. Пусть на подпространстве L с помощью операторных многочленов $F^k\left(\hat{T}\right)$ индуцированы операторы $\hat{T}_1, \hat{T}_2, \ldots$. Обозначим множества значений этих операторов через $\hat{T}_1(L), \hat{T}_2(L), \ldots$. Так как ядро операторного многочлена $F\left(\hat{T}\right)$ состоит, по крайней мере, из всех собственных векторов оператора $\hat{T}$, принадлежащих L, операторный многочлен $F\left(\hat{T}\right)$ является вырожденным на подпространстве L. Поэтому

$$\hat{T}_1(L) \subset L, \dim \hat{T}_1(L) < \dim L.$$

Подпространство $\hat{T}_1(L)$ инвариантно относительно $\hat{T}$. Если $\hat{T}_1(L)$ состоит не только из нуль-вектора, то, как показано выше, характеристический многочлен индуцированного оператора $\hat{T}\big|\hat{T}_1(L)$ является делителем характеристического многочлена индуцированного оператора $\hat{T}|L$. Поэтому все собственные значения оператора $\hat{T}\big|\hat{T}_1(L)$ являются корнями многочлена $F(z)$. Отсюда следует, что

$$\hat{T}_2(L) \subset \hat{T}_1(L), \dim \hat{T}_2(L) < \dim \hat{T}_1(L),$$

и так далее. Так как размерности подпространств $\hat{T}_1(L), \hat{T}_2(L), \ldots$ не могут убывать неограниченно, начиная с некоторого k подпространства $\hat{T}_1(L), \hat{T}_2(L), \ldots$ будут являться пустыми множествами, то есть начиная с этого номера k будут выполняться включения

$$L \subset K\left(\hat{T}_k(L)\right),$$

что и доказывает лемму. ••

Докажем теорему, дающую критерий разложения оператора в прямую сумму операторов.

Теорема 4.3.2. *Пусть характеристический многочлен $T(\mu)$ оператора $\hat{T}: X^n \to X^n$ разложен в произведение многочленов $\Phi(\mu)$ и $\Psi(\mu)$, не имеющих равных корней. Тогда оператор $\hat{T}$ единственным образом можно разложить в прямую сумму операторов $\hat{\Phi}$ и $\hat{\Psi}$ с характеристическими многочленами $\Phi(\mu)$ и $\Psi(\mu)$.*

Доказательство. По лемме 4.3.1 существует по крайней мере одно разложение оператора $\hat{T}$ в прямую сумму операторов с характеристическими многочленами, определяющими прямую сумму подпространств.

Предположим, что пространство X разложено в прямую сумму подпространств L и M некоторым другим способом, без использования разложения характеристического многочлена. Пусть индуцированные операторы на L и M имеют, соответственно, характеристические многочлены $\Phi(\mu)$ и $\Psi(\mu)$. По лемме 4.3.2 для достаточно больших значений k, то есть начиная с некоторого $k = q$, имеем

$$L \subset K_q = K\left(\Phi^q\left(\hat{T}\right)\right).$$

Оператор $\Phi\left(\hat{T}\right)$ на подпространстве M невырожденный. Следовательно, множество значений

$$\Phi\left(\hat{T}\right)(M) = M.$$

Это означает, что

$$M \subset \Phi^k\left(\hat{T}\right)(M)$$

для всех k, начиная с $k = q$, то есть

$$M \subset \Phi^q\left(\hat{T}\right)(M).$$

Для рассмотренных подпространств имеем:

$$X = L \oplus M \wedge X = K\left(\Phi^q\left(\hat{T}\right)\right) \oplus \Phi^q\left(\hat{T}\right)(M).$$

Поэтому полученные включения

$$K\left(\hat{T}\right) \subset K\left(\Phi^q\left(\hat{T}\right)\right) \wedge M \subset \Phi^q\left(\hat{T}\right)(M)$$

выполняются только в том случае, если

$$L = K\left(\Phi^q\left(\hat{T}\right)\right) \wedge M = \Phi^q\left(\hat{T}\right)(M). \bullet\bullet$$

Оператор проектирования на подпространство. Операторы проектирования встречались нам выше при рассмотрении примеров конкретных задач. Причём эти операторы не определялись строго и понимались на интуитивном уровне. В этом пункте дано строгое определение операторов проектирования в общем случае аффинного и евклидова пространств и рассмотрены некоторые их свойства.

Определение 4.3.2. *Пусть* $X = L \oplus M$*, то есть* $\left(\forall \vec{x} \in X\right)$ *справедливо единственное разложение вида*

$$\vec{x} = \vec{y} + \vec{z},$$

где $\vec{y} \in L$ и $\vec{z} \in M$. Рассмотрим операторы $\hat{P}_L$ и $\hat{P}_M$, действие которых определим правилами

$$\left(\forall \vec{x} \in X\right) \hat{P}_L \vec{x} = \vec{y} \wedge \hat{P}_M \vec{x} = \vec{z}.$$

Операторы $\hat{P}_L$ *и* $\hat{P}_M$ *называются, соответственно,* ***оператором проектирования на подпространство L параллельно подпространству M и оператором проектирования на подпространство M параллельно подпространству L.*** •

Лемма 4.3.3. *Оператор проектирования является линейным оператором.*

Доказательство. Действительно, пусть

$$\vec{x} = \alpha \cdot \vec{x}_1 + \beta \cdot \vec{x}_2,\ \vec{x}_1 = \vec{y}_1 + \vec{z}_1,\ \vec{x}_2 = \vec{y}_2 + \vec{z}_2,$$

где $\vec{y}_1, \vec{y}_2 \subset L$ и $\vec{z}_1, \vec{z}_2 \subset M$. Для вектора $\vec{x}$, очевидно, имеем

$$\vec{x} = \alpha \cdot \vec{x}_1 + \beta \cdot \vec{x}_2 = \alpha \cdot \left(\vec{y}_1 + \vec{z}_1\right) + \beta \cdot \left(\vec{y}_2 + \vec{z}_2\right) =$$

$$= \left(\alpha \cdot \vec{y}_1 + \beta \cdot \vec{y}_2\right) + \left(\alpha \cdot \vec{z}_1 + \beta \cdot \vec{z}_2\right). \qquad (4.3.3)$$

Здесь $\alpha\cdot\vec{y}_1+\beta\cdot\vec{y}_2\in L$ и $\alpha\cdot\vec{z}_1+\beta\cdot\vec{z}_2\in M$. Применяя к правой части (4.3.3), например, оператор проектирования на подпространство L, получаем

$$\hat{P}_L\left[\left(\alpha\cdot\vec{y}_1+\beta\cdot\vec{y}_2\right)+\left(\alpha\cdot\vec{z}_1+\beta\cdot\vec{z}_2\right)\right]=\alpha\cdot\vec{y}_1+\beta\cdot\vec{y}_2.$$

Предположим, что оператор $\hat{P}_L$ линеен. Тогда, действуя оператором $\hat{P}_L$ на левую часть представления (4.3.3), имеем

$$\hat{P}_L\left(\alpha\cdot\vec{x}_1+\beta\cdot\vec{x}_2\right)=\alpha\cdot\hat{P}_L\vec{x}_1+\beta\cdot\hat{P}_L\vec{x}_2=\alpha\cdot\vec{y}_1+\beta\cdot\vec{y}_2.$$

Из сравнения двух последних равенств получаем доказательство леммы. ••

Пусть $X=L\oplus M$ и вектор $\vec{x}\in L$, тогда $\hat{P}_L\vec{x}=\vec{x}$. Если же $\vec{x}\in M$, то $\hat{P}_L\vec{x}=\vec{0}$. Таким образом, в соответствии с определениями 4.3.1 и 4.3.2 оператор проектирования на подпространство $L\subset X$ является прямой суммой единичного оператора, действующего на подпространстве L, и нулевого оператора, действующего на подпространстве M: $\hat{P}_L=\hat{I}_L\oplus\hat{O}_M$. Так как подпространство L инвариантно относительно единичного оператора, а подпространство M инвариантно относительно нулевого оператора, то, как показано выше, матрица оператора проектирования на подпространство L является клеточной и имеет вид:

$$P_L=\begin{pmatrix} I_1^1 & O_2^1 \\ O_1^2 & O_2^2 \end{pmatrix}. \tag{4.3.4}$$

Здесь I_1^1 — единичная матрица порядка $\dim L$, O_2^2 — нулевая матрица порядка $\dim M$, а O_2^1 и O_1^2 — нулевые матрицы размерами, соответственно,

$$O_2^1\Rightarrow\dim L\times\left(n-\dim L\right) \text{ и } O_1^2\Rightarrow\left(n-\dim L\right)\times\dim L.$$

Определение 4.3.3. *Оператор* $\hat{T}:X\to X$ *называется* ***идемпотентным,*** *если* $\hat{T}^2=\hat{T}$. •

Теорема 4.3.3. *Оператор* $\hat{P}:X\to X$ *является оператором проектирования на подпространство* $L\subset X$ *в том и только в том случае, если он* *является идемпотентным оператором, отличным от единичного и нулевого оператора.*

Доказательство. Пусть $X = L \oplus M$ и оператор $\hat{P} : X \to X$, для определённости, является оператором проектирования на подпространство L параллельно подпространству M, то есть $\hat{P} = \hat{P}_L$. Тогда

$$\left(\forall \vec{x} \in X\right) \hat{P}_L \vec{x} \in L.$$

Поэтому

$$\hat{P}_L\left(\hat{P}_L \vec{x}\right) = \hat{P}_L \vec{x}.$$

Отсюда следует, что $\hat{P}_L^2 = \hat{P}_L$, то есть оператор $\hat{P}_L$ является идемпотентным.

Пусть теперь оператор $\hat{P}$ является идемпотентным, причём $\hat{P} \neq \hat{I}$ и $\hat{P} \neq \hat{O}$. Обозначим $\hat{P}(X) = L$ и $\left(\hat{I} - \hat{P}\right)(X) = M$. Тогда, очевидно,

$$\left(\forall \vec{x} \in X\right) \vec{x} = \hat{I}\vec{x} = \hat{P}\vec{x} + \left(\hat{I} - \hat{P}\right)\vec{x} = \vec{y} + \vec{z},$$

где $\vec{y} \in L$ и $\vec{z} \in M$. Поэтому $X = L + M$. Если мы покажем, что эта сумма прямая, то доказательство теоремы будет завершено. Пусть $\vec{v} \in L \cap M$. Тогда для некоторых векторов $\vec{x}_1, \vec{x}_2 \in X$ можем записать следующие представления:

$$\vec{v} = \hat{P}\vec{x}_1,\ \vec{v} = \left(\hat{I} - \hat{P}\right)\vec{x}_2.$$

Так как оператор $\hat{P}$ является идемпотентным, из первого и второго представлений, соответственно, получаем:

$$\hat{P}\vec{v} = \hat{P}^2 \vec{x}_1 = \hat{P}\vec{x}_1 = \vec{v};\ \hat{P}\vec{v} = \left(\hat{P} - \hat{P}^2\right)\vec{x}_2 = \vec{0}.$$

Таким образом, $\vec{v} = \vec{0}$ и сумма $X = L + M$ является прямой суммой. ••

Оператор ортогонального проектирования. Рассмотрим случай евклидова пространства E. Если $E = L \oplus M$, где $M = L^{\perp}$, то оператор проектирования на подпространство $\hat{P}_L$ называется оператором ***ортого-***

нального проектирования. Оператор ортогонального проектирования является самосопряжённым, так как имеет вещественную диагональную матрицу в ортонормированном базисе, являющимся объединением ортонормированных базисов подпространств L и M.

Теорема 4.3.4. *Любой самосопряжённый идемпотентный оператор является оператором ортогонального проектирования.*

Доказательство. Пусть $\hat{P}$ — самосопряжённый идемпотентный оператор. Выше показано, что любой идемпотентный оператор является оператором проектирования пространства E на подпространство $L = \hat{P}(E)$ параллельно подпространству $M = \left(\hat{I} - \hat{P}\right)(E)$. Покажем, что подпространства L и M ортогональны. Действительно, пусть $\vec{y} = \hat{P}\vec{u} \in L$ и $\vec{z} = \vec{v} - \hat{P}\vec{v} \in M$. Так как оператор $\hat{P}$ самосопряжённый и идемпотентный, имеем

$$\left(\vec{y}, \vec{z}\right) = \left(\hat{P}\vec{u}, \vec{v} - \hat{P}\vec{v}\right) = \left(\vec{u}, \hat{P}\vec{v} - \hat{P}^2\vec{v}\right) = \left(\vec{u}, \vec{0}\right) = 0,$$

что и доказывает теорему. ••

Пример 4.3.1. Записать матрицу оператора $\hat{P}^{\perp}_{x^1ox^2}$ ортогонального проектирования на координатную плоскость X^1OX^2 в пространстве R^3.

Решение. Этот оператор любому вектору пространства R^3 ставит в соответствие его проекцию на координатную плоскость X^1OX^2 параллельно координатной оси OX^3. Согласно изложенной выше теории оператор $\hat{P}^{\perp}_{x^1ox^2}$ является прямой суммой единичного оператора

$$\hat{I}_{x^1ox^2} : R^2_{x^1ox^2} \to R^2_{x^1ox^2}$$

и нулевого оператора

$$\hat{O}_{ox^3} : R^1_{ox^3} \to R^1_{ox^3}.$$

На главной диагонали его матрицы расположены 2×2 клетка вида

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

и 1×1 клетка вида (0). Остальные элементы равны нулю:

$$P^{\perp}_{x^1ox^2}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}.$$

Легко видеть, что оператор $\hat{P}^{\perp}_{x^1ox^2}$ является самосопряжённым. ⊗

Глава 5.
Геометрия векторных пространств

5.1. Некоторые задачи геометрии в *n*-мерном собственно евклидовом пространстве

В этом параграфе рассмотрены некоторые типичные геометрические задачи в ***собственно евклидовом пространстве***.

Критерий Грама линейной зависимости системы векторов. В собственно евклидовом пространстве E^n вопрос о линейной зависимости системы векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\} \subset E^n$ может быть решён без разложения векторов по базису с последующим решением СЛАУ.

Определение 5.1.1. *Пусть дана система векторов*

$$\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\} \subset E^n.$$

Определитель вида

$$G\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right) \stackrel{def}{=} \begin{vmatrix} \left(\vec{x}_1, \vec{x}_1\right) & \left(\vec{x}_1, \vec{x}_2\right) & \ldots & \left(\vec{x}_1, \vec{x}_m\right) \\ \left(\vec{x}_2, \vec{x}_1\right) & \left(\vec{x}_2, \vec{x}_2\right) & \ldots & \left(\vec{x}_2, \vec{x}_m\right) \\ \ldots & \ldots & \ldots & \ldots \\ \left(\vec{x}_m, \vec{x}_1\right) & \left(\vec{x}_m, \vec{x}_2\right) & \ldots & \left(\vec{x}_m, \vec{x}_m\right) \end{vmatrix} \qquad (5.1.1)$$

называется ***определителем Грама***. •

Теорема 5.1.1. *Система* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ *векторов евклидова пространства E^n линейно зависима в том и только в том случае, если её определитель Грама равен нулю.*

Доказательство. *Докажем необходимость*. Пусть система

$$\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$$

линейно зависима. Тогда найдутся такие, не равные нулю одновременно числа $\alpha_1, \alpha_2, \ldots, \alpha_m$, что выполняется условие

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \ldots + \alpha_m \vec{x}_m = \vec{0}. \qquad (5.1.2)$$

Находя последовательно скалярное произведение обеих частей векторного равенства (5.1.2) и векторов $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$, получим:

$$\begin{aligned}
&\alpha_1\left(\vec{x}_1, \vec{x}_1\right) + \alpha_2\left(\vec{x}_2, \vec{x}_1\right) + \ldots + \alpha_m\left(\vec{x}_m, \vec{x}_1\right) = 0,\\
&\alpha_1\left(\vec{x}_1, \vec{x}_2\right) + \alpha_2\left(\vec{x}_2, \vec{x}_2\right) + \ldots + \alpha_m\left(\vec{x}_m, \vec{x}_2\right) = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\\
&\alpha_1\left(\vec{x}_1, \vec{x}_m\right) + \alpha_2\left(\vec{x}_2, \vec{x}_m\right) + \ldots + \alpha_m\left(\vec{x}_m, \vec{x}_m\right) = 0.
\end{aligned} \qquad (5.1.3)$$

Система уравнений (5.1.3) — следствие линейной зависимости строк (столбцов) определителя Грама (5.1.1), поэтому одно из уравнений может быть линейно выражено через остальные. Вычтем из него остальные уравнения, умноженные на соответствующие коэффициенты, при этом определитель системы не изменится, но элементы соответствующей строки определителя обратятся в нули. Следовательно, имеем

$$G\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right) = 0.$$

Докажем достаточность. Пусть теперь

$$G\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right) = 0.$$

Тогда строки определителя Грама образуют линейно зависимую систему, то есть справедливы равенства (5.1.3). Перепишем их в следующем виде:

$$\left(\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \ldots + \alpha_m \vec{x}_m, \vec{x}_1\right) = 0,$$

$$\left(\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m, \vec{x}_2\right)=0,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\left(\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m, \vec{x}_m\right)=0.$$

Умножая обе части этих равенств, соответственно, на $\alpha_1,\ \alpha_2,\ldots,\alpha_m$ и относя коэффициенты ко второму аргументу, в силу линейности скалярного произведения получаем

$$\left(\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m, \alpha_1 \vec{x}_1\right)=0,$$

$$\left(\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m, \alpha_2 \vec{x}_2\right)=0,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\left(\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m, \alpha_m \vec{x}_m\right)=0.$$

Складывая последние равенства и снова используя свойства линейности скалярного произведения, получим

$$\left(\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m, \alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m\right)=0.$$

Так как

$$\left\|\vec{x}\right\| \stackrel{def}{=} +(\vec{x},\vec{x})^{1/2} \equiv \sqrt{\vec{x}^{\,2}},$$

из последнего равенства получаем

$$\left\|\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m\right\|^2=0,$$

откуда, очевидно, следует, что

$$\alpha_1 \vec{x}_1+\alpha_2 \vec{x}_2+\ldots+\alpha_m \vec{x}_m=\vec{0},$$

где коэффициенты $\alpha_1,\ \alpha_2,\ldots,\alpha_m$ не обращаются в нуль одновременно. Таким образом, система векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ линейно зависима. ••

Следствие из теоремы 5.1.1. *Система* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ *векторов евклидова пространства* E^n *линейно независима в том и только в том случае, если её определитель Грама не равен нулю.*

Следствие является, очевидно, логическим противопоставлением утверждению теоремы 5.1.1.

Наклонная, перпендикуляр и проекция в *n*-мерном евклидовом пространстве. Рассмотрим в n-мерном евклидовом пространстве E^n некоторое подпространство L^m размерности $m < n$. Пусть дан некоторый вектор $\vec{x} \in E^n$, причём $\vec{x} \notin L^m$. Покажем, что справедливо представление вектора $\vec{x}$ в виде следующего разложения:

$$\vec{x} = \vec{g}_L + \vec{h}^{\perp}, \tag{5.1.4}$$

где вектор $\vec{g}_L$ принадлежит подпространству L^m, а вектор $\vec{h}^{\perp}$ ортогонален к этому подпространству.

Определение 5.1.2. *В разложении (5.1.4) векторы* $\vec{x}$ *и* $\vec{g}_L$ *называются соответственно* ***наклонной к подпространству*** L^m *и* ***проекцией наклонной*** $\vec{x}$ ***на подпространство*** L^m. *Вектор* $\vec{h}^{\perp}$ *называется* ***перпендикуляром, опущенным из конца наклонной*** $\vec{x}$ ***на подпространство*** L^m. •

Эта терминология связана с известными понятиями из обычной трёхмерной геометрии. Здесь мы воспользуемся теорией СЛАУ для разработки аппарата, позволяющего проводить конкретные вычисления в произвольных пространствах конечной размерности.

Теорема 5.1.2. *Пусть в евклидовом пространстве* E^n *заданы произвольный вектор* $\vec{x}$ *и некоторое* m*-мерное подпространство* $L^m \subset E^n$ *с зафиксированным и, вообще говоря, неортогональным и ненормированным базисом* $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$. *Тогда вектор* $\vec{x}$, *не принадлежащий подпространству* L^m, *можно единственным образом представить в виде суммы*

$$\vec{x} = \vec{g}_L + \vec{h}^{\perp}, \ \vec{g}_L \in L^m, \ \vec{h}^{\perp} \perp L^m. \tag{5.1.5}$$

Доказательство. Представим $\vec{g}_L$ в виде разложения по векторам базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ подпространства L^m:

$$\vec{g}_L = g^1 \vec{a}_1 + g^2 \vec{a}_2 + \ldots + g^m \vec{a}_m, \tag{5.1.6}$$

где $g^1, g^2, \ldots, g^m$ — некоторые числа из поля R^1. Потребуем, чтобы вектор

$$\vec{h}^{\perp} = \vec{x} - \vec{g}_L$$

был ортогонален подпространству L^m. Как известно, это условие выполняется в том и только в том случае, если вектор $\vec{h}^{\perp}$ ортогонален всем векторам какого-либо базиса подпространства L^m, например, зафиксированного базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$. Поэтому должны выполняться соотношения

$$\left(\vec{h}^{\perp}, \vec{a}_k\right) = \left(\vec{x} - \vec{g}_L, \vec{a}_k\right) = 0, \tag{5.1.7}$$

где $k = 1, 2, \ldots, m$.

Подставляя в соотношение (5.1.7) разложение $\vec{g}_L$ из выражения (5.1.6), получаем, что

$$(\forall k = 1, 2, \ldots, m): \left(\vec{x} - \left(g^1 \vec{a}_1 + g^2 \vec{a}_2 + \ldots + g^m \vec{a}_m\right), \vec{a}_k\right) = 0.$$

Записывая последнее соотношение в развёрнутом виде последовательно для номеров $k = 1, 2, \ldots, m$, получаем

$$\begin{aligned}
&\left(\vec{a}_1, \vec{a}_1\right) g^1 + \left(\vec{a}_2, \vec{a}_1\right) g^2 + \ldots + \left(\vec{a}_m, \vec{a}_1\right) g^m = \left(\vec{x}, \vec{a}_1\right),\\
&\left(\vec{a}_1, \vec{a}_2\right) g^1 + \left(\vec{a}_2, \vec{a}_2\right) g^2 + \ldots + \left(\vec{a}_m, \vec{a}_2\right) g^m = \left(\vec{x}, \vec{a}_2\right),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\\
&\left(\vec{a}_1, \vec{a}_m\right) g^1 + \left(\vec{a}_2, \vec{a}_m\right) g^2 + \ldots + \left(\vec{a}_m, \vec{a}_m\right) g^m = \left(\vec{x}, \vec{a}_m\right).
\end{aligned} \tag{5.1.8}$$

Получили СЛАУ относительно неизвестных $g^1, g^2, \ldots, g^m$ — коэффициентов разложения (5.1.6) вектора $\vec{g}_L$ по векторам базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ подпространства L^m. Определитель этой СЛАУ

$$G^T\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right) = \begin{vmatrix} \left(\vec{a}_1, \vec{a}_1\right) & \left(\vec{a}_2, \vec{a}_1\right) & \ldots & \left(\vec{a}_m, \vec{a}_1\right) \\ \left(\vec{a}_1, \vec{a}_2\right) & \left(\vec{a}_2, \vec{a}_2\right) & \ldots & \left(\vec{a}_m, \vec{a}_2\right) \\ \ldots & \ldots & \ldots & \ldots \\ \left(\vec{a}_1, \vec{a}_m\right) & \left(\vec{a}_2, \vec{a}_m\right) & \ldots & \left(\vec{a}_m, \vec{a}_m\right) \end{vmatrix}$$

является транспонированным определителем Грама линейно независимой системы векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$. Потому он отличен от нуля.

Следовательно, СЛАУ (5.1.8) совместна и определённа. Решение её (значения неизвестных $g^1, g^2, \ldots, g^m$) можно найти по формулам Крамера, методом Гаусса или матричным методом. Например, по формулам Крамера получаем, что

$$g^k = \frac{D_k}{G^T\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right)}, \tag{5.1.9}$$

где определители D_k имеют вид:

$$D_k = \begin{vmatrix} \left(\vec{a}_1, \vec{a}_1\right) & \ldots & \left(\vec{a}_{k-1}, \vec{a}_1\right) & \left(\vec{x}, \vec{a}_1\right) & \left(\vec{a}_{k+1}, \vec{a}_1\right) & \ldots & \left(\vec{a}_m, \vec{a}_1\right) \\ \left(\vec{a}_1, \vec{a}_2\right) & \ldots & \left(\vec{a}_{k-1}, \vec{a}_2\right) & \left(\vec{x}, \vec{a}_2\right) & \left(\vec{a}_{k+1}, \vec{a}_2\right) & \ldots & \left(\vec{a}_m, \vec{a}_2\right) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \left(\vec{a}_1, \vec{a}_m\right) & \ldots & \left(\vec{a}_{k-1}, \vec{a}_m\right) & \left(\vec{x}, \vec{a}_m\right) & \left(\vec{a}_{k+1}, \vec{a}_m\right) & \ldots & \left(\vec{a}_m, \vec{a}_m\right) \end{vmatrix}.$$

Подставляя коэффициенты (5.1.9) в разложение (5.1.6), получаем проекцию $\vec{g}_L$ вектора $\vec{x}$ на подпространство $L^m \subset E^n$. Теперь перпендикуляр $\vec{h}^{\perp}$ находим из разложения (5.1.5):

$$\vec{h}^{\perp} = \vec{x} - \vec{g}_L. \bullet\bullet$$

Объём параллелепипеда в *n*-мерном собственно евклидовом пространстве. В одномерном случае ***объём*** — это просто ***норма*** (длина) вектора.

В двумерном случае ***объём*** — это ***площадь*** параллелограмма, которая вычисляется как произведение его основания на высоту. Пусть параллелограмм построен на двух векторах $\vec{x}_1, \vec{x}_2$. Тогда за основание параллелограмма можно принять, например, длину вектора $\vec{x}_1$, а за высоту — длину перпендикуляра, опущенного из конца вектора $\vec{x}_2$ на одномерное подпространство, образованное линейной оболочкой вектора $\vec{x}_1$ (рис. 5.1).

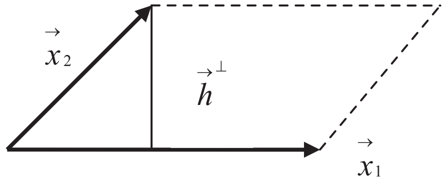


Рис. 5.1

В трёхмерном случае ***объём*** параллелепипеда, построенного на векторах $\vec{x}_1, \vec{x}_2, \vec{x}_3$, равен произведению площади параллелограмма, построенного на векторах $\vec{x}_1, \vec{x}_2$ (основание), на длину перпендикуляра (высоту), опущенного из конца вектора $\vec{x}_3$ на линейную оболочку (плоскость) векторов $\vec{x}_1, \vec{x}_2$ (рис. 5.2).

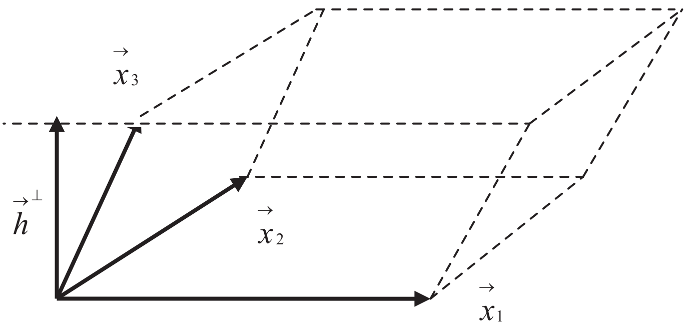


Рис. 5.2

Эти элементарные рассуждения приводят к естественному обобщению понятия объёма на случай ***параллелепипеда в конечномерном собственно евклидовом пространстве***.

Определение 5.1.3. *Пусть в евклидовом пространстве* E^n *зафиксирована система векторов* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}(m<n)$. *Обозначим* $\vec{h}_m^{\perp}$ *перпендикуляр, опущенный из конца вектора* $\vec{x}_{m+1}$ *на подпространство*

$$L\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}(m=1, 2, \ldots, n-1).$$

Положим по определению:

$V_1 \overset{def}{=} \left\|\vec{x}_1\right\|$ — ***одномерный объём*** — *длина вектора* $\vec{x}_1$*;*

$V_2 \overset{def}{=} V_1\left\|\vec{h}_1\right\|$ — ***двумерный объём*** — *площадь параллелограмма, построенного на векторах* $\vec{x}_1, \vec{x}_2$*;*

$V_3 \overset{def}{=} V_2\left\|\vec{h}_2\right\|$ — ***трёхмерный объём*** — *объём параллелепипеда, построенного на векторах* $\vec{x}_1, \vec{x}_2, \vec{x}_3$*;*

....;

$V_m \overset{def}{=} V_{m-1}\left\|\vec{h}_{m-1}\right\|$ — ***m-мерный объём*** — *объём параллелепипеда, построенного на векторах* $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$. •

Из определения непосредственно следует формула для вычисления объёма m-мерного параллелепипеда:

$$V_m = V\left[\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right] \overset{def}{=} \left\|\vec{x}_1\right\|\left\|\vec{h}_1\right\|\ldots\left\|\vec{h}_{m-1}\right\|. \tag{5.1.10}$$

Выражение (5.1.10) в значительной степени формальное и не даёт конструктивного алгоритма вычисления объёма m-мерного параллелепипеда. Для построения такого алгоритма сначала сформулируем и докажем обобщённый вариант ***теоремы об ортогонализации (Шмидта) системы векторов в евклидовом пространстве***.

Теорема 5.1.3. *Пусть* $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m, \ldots\right\}$ — *некоторая система векторов евклидова пространства* E^n *(конечная или бесконечная — всё равно). Обозначим через*

$$L^m = L\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$$

линейную оболочку первых m векторов этой системы. Тогда существует система векторов $\left\{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m, \ldots\right\}$, *обладающая следующими свойствами:*

1) *для любого натурального m линейная оболочка*

$$L^{m\,\prime} = L\left\{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m\right\}$$

векторов $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m$ совпадает с подпространством

$$L^m = L\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\};$$

2) *для любого натурального m вектор* $\vec{y}_{m+1}$ *ортогонален к подпространству* L^m.

Д о к а з а т е л ь с т в о. Положим $\vec{y}_1 = \vec{x}_1$. При этом, очевидно, выполняется условие

$$L^{1\,\prime}\left\{\vec{y}_1\right\} = L^1\left\{\vec{x}_1\right\}.$$

Воспользуемся индукцией по числу векторов. Пусть уже построено m векторов $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m$, удовлетворяющих условиям теоремы. Построим вектор $\vec{y}_{m+1}$ так, чтобы он тоже удовлетворял условиям теоремы.

Подпространство L^m конечномерно, поэтому в соответствии с разложением (5.1.5) можем записать разложение

$$\vec{x}_{m+1} = \vec{g}_m + \vec{h}^{\perp}_m,\ \vec{g}_m \in L^m,\ \vec{h}^{\perp}_m \perp L^m. \tag{5.1.11}$$

Положим теперь

$$\vec{y}_{m+1} = \vec{h}^{\perp}_m.$$

Проверим справедливость условий теоремы для построенного таким образом вектора $\vec{y}_{m+1}$.

По предположению индукции подпространство L^m содержит векторы $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m$, следовательно, более широкое подпространство L^{m+1}

содержит эти векторы. Из формулы (5.1.11) следует, что L^{m+1} содержит также и вектор $\vec{y}_{m+1} = \vec{h}_m^{\perp}$. Получаем, что подпространство L^{m+1} содержит все векторы $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m, \vec{y}_{m+1}$, а значит, и всю их линейную оболочку

$$L'_{m+1} = L\left\{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m, \vec{y}_{m+1}\right\}.$$

Обратно, как видно из разложения (5.1.11), подпространство $L^{m+1}{}'$ содержит и вектор $\vec{x}_{m+1}$. Следовательно, $L^{m+1}{}' \supset L^{m+1}$ и $L^{m+1}{}' = L^{m+1}$.

Первое условие теоремы выполняется. Второе условие выполняется по построению вектора $\vec{y}_{m+1} = \vec{h}_m^{\perp}$. Теорема доказана. ••

Прежде чем переходить к построению объёма m-мерного параллелепипеда, сделаем одно полезное замечание.

Применяя к разложению (5.1.5) теорему Пифагора, получаем

$$\left\|\vec{x}\right\|^2 = \left\|\vec{g}_L\right\|^2 + \left\|\vec{h}^{\perp}\right\|^2.$$

Из этого равенства следует справедливость неравенства

$$0 \le \left\|\vec{h}^{\perp}\right\| \le \left\|\vec{x}\right\|, \tag{5.1.12}$$

выражающего тот геометрический факт, что ***длина перпендикуляра не превосходит длины наклонной***.

Теорема 5.1.4. *Определитель Грама*

$$G\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right) = \begin{vmatrix} \left(\vec{x}_1, \vec{x}_1\right) & \left(\vec{x}_1, \vec{x}_2\right) & \ldots & \left(\vec{x}_1, \vec{x}_m\right) \\ \left(\vec{x}_2, \vec{x}_1\right) & \left(\vec{x}_2, \vec{x}_2\right) & \ldots & \left(\vec{x}_2, \vec{x}_m\right) \\ \ldots & \ldots & \ldots & \ldots \\ \left(\vec{x}_m, \vec{x}_1\right) & \left(\vec{x}_m, \vec{x}_2\right) & \ldots & \left(\vec{x}_m, \vec{x}_m\right) \end{vmatrix}$$

системы векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ *равен квадрату объёма m-мерного параллелепипеда, построенного на этих векторах.*

Доказательство. Пусть m-мерный параллелепипед построен на векторах системы $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$.

Для того чтобы вычислить определитель Грама (5.1.1), применим к векторам системы $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ процесс ортогонализации Шмидта. Положим $\vec{y}_1 = \vec{x}_1$ и предположим, что вектор $\vec{y}_2 = \alpha_1 \vec{y}_1 + \vec{x}_2$ ортогонален вектору $\vec{y}_1$ (то есть подбираем α_1 так, чтобы выполнялось это условие). Заменим все $\vec{x}_1$ в определителе Грама на $\vec{y}_1 = \vec{x}_1$:

$$\begin{vmatrix} \left(\vec{y}_1, \vec{y}_1\right) & \left(\vec{y}_1, \vec{x}_2\right) & \ldots & \left(\vec{y}_1, \vec{x}_m\right) \\ \left(\vec{x}_2, \vec{y}_1\right) & \left(\vec{x}_2, \vec{x}_2\right) & \ldots & \left(\vec{x}_2, \vec{x}_m\right) \\ \ldots & \ldots & \ldots & \ldots \\ \left(\vec{x}_m, \vec{y}_1\right) & \left(\vec{x}_m, \vec{x}_2\right) & \ldots & \left(\vec{x}_m, \vec{x}_m\right) \end{vmatrix}.$$

Умножим мысленно первый столбец определителя на α_1 и, относя α_1 ко вторым аргументам скалярных произведений, прибавим результат ко второму столбцу, получим:

$$\begin{vmatrix} \left(\vec{y}_1, \vec{y}_1\right) & \left(\vec{y}_1, \vec{x}_2 + \alpha_1 \vec{y}_1\right) & \ldots & \left(\vec{y}_1, \vec{x}_m\right) \\ \left(\vec{x}_2, \vec{y}_1\right) & \left(\vec{x}_2, \vec{x}_2 + \alpha_1 \vec{y}_1\right) & \ldots & \left(\vec{x}_2, \vec{x}_m\right) \\ \ldots & \ldots & \ldots & \ldots \\ \left(\vec{x}_m, \vec{y}_1\right) & \left(\vec{x}_m, \vec{x}_2 + \alpha_1 \vec{y}_1\right) & \ldots & \left(\vec{x}_m, \vec{x}_m\right) \end{vmatrix}.$$

Первую строку получившегося определителя мысленно умножим на α_1 и, относя α_1 к первым аргументам скалярных произведений, прибавим результат ко второй строке, получим:

$$\begin{vmatrix} \left(\vec{y}_1, \vec{y}_1\right) & \left(\vec{y}_1, \vec{x}_2 + \alpha_1 \vec{y}_1\right) & \dots & \left(\vec{y}_1, \vec{x}_m\right) \\ \left(\vec{x}_2 + \alpha_1 \vec{y}_1, \vec{y}_1\right) & \left(\vec{x}_2 + \alpha_1 \vec{y}_1, \vec{x}_2 + \alpha_1 \vec{y}_1\right) & \dots & \left(\vec{x}_2 + \alpha_1 \vec{y}_1, \vec{x}_m\right) \\ \dots & \dots & \dots & \dots \\ \left(\vec{x}_m, \vec{y}_1\right) & \left(\vec{x}_m, \vec{x}_2 + \alpha_1 \vec{y}_1\right) & \dots & \left(\vec{x}_m, \vec{x}_m\right) \end{vmatrix}.$$

Учитывая, что $\vec{y}_2 = \alpha_1 \vec{y}_1 + \vec{x}_2$, получаем, что на всех местах, где в исходном определителе стоял вектор $\vec{x}_2$, теперь стоит вектор $\vec{y}_2$.

Пусть вектор $\vec{y}_3 = \beta_1 \vec{y}_1 + \beta_2 \vec{y}_2 + \vec{x}_3$ ортогонален векторам $\vec{y}_1$ и $\vec{y}_2$. Прибавим к третьему столбцу первый, умноженный на β_1, и второй, умноженный на β_2; аналогичную операцию произведём со строками определителя. В получившемся определителе на всех местах, где раньше стоял вектор $\vec{x}_3$, будет стоять вектор $\vec{y}_3$. Продолжая преобразования, получаем:

$$G\left(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m\right) = \begin{vmatrix} \left(\vec{y}_1, \vec{y}_1\right) & \left(\vec{y}_1, \vec{y}_2\right) & \dots & \left(\vec{y}_1, \vec{y}_m\right) \\ \left(\vec{y}_2, \vec{y}_1\right) & \left(\vec{y}_2, \vec{y}_2\right) & \dots & \left(\vec{y}_2, \vec{y}_m\right) \\ \dots & \dots & \dots & \dots \\ \left(\vec{y}_m, \vec{y}_1\right) & \left(\vec{y}_m, \vec{y}_2\right) & \dots & \left(\vec{y}_m, \vec{y}_m\right) \end{vmatrix}.$$

Так как векторы $\left\{\vec{y}_1, \vec{y}_2, \dots, \vec{y}_m\right\}$ ортогональны по построению, то

$$G\left(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m\right) = \begin{vmatrix} \left(\vec{y}_1, \vec{y}_1\right) & 0 & \dots & 0 \\ 0 & \left(\vec{y}_2, \vec{y}_2\right) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \left(\vec{y}_m, \vec{y}_m\right) \end{vmatrix}.$$

Этот определитель равен произведению диагональных элементов, то есть

$$G\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right) = \left(\vec{y}_1, \vec{y}_1\right)\left(\vec{y}_2, \vec{y}_2\right)\ldots\left(\vec{y}_m, \vec{y}_m\right). \qquad (5.1.13)$$

Из формулы (5.1.13) следует, что

$$0 \le G\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right) \le \left(\vec{x}_1, \vec{x}_1\right)\left(\vec{x}_2, \vec{x}_2\right)\ldots\left(\vec{x}_m, \vec{x}_m\right). \qquad (5.1.14)$$

В неравенстве (5.1.14) крайнее левое значение достигается, согласно теореме 5.1.1 в том и только в том случае, если система векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ линейно зависима. Крайнее правое значение достигается в том и только в том случае, когда система векторов $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right\}$ ортогональна. Теперь квадрат объёма m-мерного параллелепипеда (5.1.10) с учётом формулы (5.1.13) можно представить в виде:

$$(V_m)^2 = \begin{vmatrix} \left(\vec{x}_1, \vec{x}_1\right) & \left(\vec{x}_1, \vec{x}_2\right) & \ldots & \left(\vec{x}_1, \vec{x}_m\right) \\ \left(\vec{x}_2, \vec{x}_1\right) & \left(\vec{x}_2, \vec{x}_2\right) & \ldots & \left(\vec{x}_2, \vec{x}_m\right) \\ \multicolumn{4}{c}{\dotfill} \\ \left(\vec{x}_m, \vec{x}_1\right) & \left(\vec{x}_m, \vec{x}_2\right) & \ldots & \left(\vec{x}_m, \vec{x}_m\right) \end{vmatrix}. \bullet\bullet$$

Выше отмечено, что в данном параграфе рассмотрены типичные геометрические задачи в собственно евклидовом пространстве. Общий случай вещественных евклидовых пространств рассмотрен в параграфе 6.3.

5.2. Квадратичные формы в пространстве R^n

Понятие квадратичной формы. В этом параграфе мы ограничимся теорией квадратичных форм в простом случае пространства R^n. Общий случай рассмотрен в главе 6, параграф 6.3.

Пусть

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset R^n$$

— ортонормированный базис в пространстве R^n и $\hat{T} : R^n \to R^n$ — линейный оператор. Учитывая, что для любого вектора $\vec{x} \in R^n$ справедливо разложение

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{e}_i,$$

найдём вид скалярного произведения в евклидовом пространстве R^n, взяв в качестве его аргументов вектор $\vec{x}$ и его образ $\hat{T}\vec{x}$:

$$\left(\vec{x}, \hat{T}\vec{x}\right) = \left(\sum_{i=1}^{n} x^i \vec{e}_i, \hat{T}\sum_{j=1}^{n} x^j \vec{e}_j\right) =$$

$$= \left(\sum_{i=1}^{n} x^i \vec{e}_i, \sum_{j=1}^{n} x^j \hat{T}\vec{e}_j\right) = \left(\sum_{i=1}^{n} x^i \vec{e}_i, \sum_{j=1}^{n} x^j \sum_{k=1}^{n} t_j^k \vec{e}_k\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x^i x^j \sum_{k=1}^{n} t_j^k \left(\vec{e}_i, \vec{e}_k\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i x^j \left(\sum_{k=1}^{n} t_j^k \delta_{ik}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij} x^i x^j =$$

$$= t_{11}\left(x^1\right)^2 + t_{12}x^1x^2 + t_{13}x^1x^3 + \ldots + t_{1n-1}x^1x^{n-1} + t_{1n}x^1x^n +$$

$$+ t_{21}x^2x^1 + t_{22}\left(x^2\right)^2 + t_{23}x^2x^3 + \ldots + t_{2n-1}x^2x^{n-1} + t_{2n}x^2x^n +$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots +$$

$$+ t_{n1}x^nx^1 + t_{n2}x^nx^2 + t_{n3}x^nx^3 + \ldots + t_{nn-1}x^nx^{n-1} + t_{nn}\left(x^n\right)^2 =$$

$$= t_{11}\left(x^1\right)^2 + 2\frac{\left(t_{12} + t_{21}\right)}{2}x^1x^2 + 2\frac{\left(t_{13} + t_{31}\right)}{2}x^1x^3 + \ldots +$$

$$+ t_{22}\left(x^2\right)^2 + \ldots + 2\frac{\left(t_{n\,n-1} + t_{n-1\,n}\right)}{2}x^{n-1}x^n + t_{nn}\left(x^n\right)^2 =$$

$$= t_{11}\left(x^1\right)^2 + \frac{t_{12} + t_{21}}{2}x^1x^2 + \frac{t_{13} + t_{31}}{2}x^1x^3 + \ldots +$$

$$+ \frac{t_{1n-1} + t_{n-11}}{2}x^1x^{n-1} + \frac{t_{1n} + t_{n1}}{2}x^1x^n +$$

$$+ \frac{t_{21} + t_{12}}{2}x^2x^1 + t_{22}\left(x^2\right)^2 + \frac{t_{23} + t_{32}}{2}x^2x^3 + \ldots +$$

$$+\frac{t_{2n-1}+t_{n-12}}{2}x^2x^{n-1}+\frac{t_{2n}+t_{n2}}{2}x^2x^n+$$

$$+\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots+$$

$$+\frac{t_{n1}+t_{1n}}{2}x^nx^1+\frac{t_{n2}+t_{2n}}{2}x^nx^2+\frac{t_{n3}+t_{3n}}{2}x^nx^3+\ldots+$$

$$+\frac{t_{nn-1}+t_{n-1n}}{2}x^nx^{n-1}+t_{nn}\left(x^n\right)^2.$$

Введём обозначения:

$$t_{ii}=\varphi_{ii},\ \varphi_{ij}=\frac{t_{ij}+t_{ji}}{2}=\frac{t_{ji}+t_{ij}}{2}=\varphi_{ji}.$$

Тогда получаем:

$$\varphi\left(\vec{x},\vec{x}\right)\overset{def}{=}\left(\vec{x},\hat{T}\vec{x}\right)=$$

$$=\varphi_{11}\left(x^1\right)^2+\varphi_{12}x^1x^2+\ldots+\varphi_{1n}x^1x^n+$$

$$+\varphi_{21}x^2x^1+\varphi_{22}\left(x^2\right)^2+\ldots+\varphi_{2n}x^2x^n+$$

$$+\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots+$$

$$+\varphi_{n1}x^nx^1+\varphi_{n2}x^nx^2+\ldots+\varphi_{nn}\left(x^n\right)^2.$$

Определение 5.2.1. ***Квадратичной формой*** $\varphi\left(\vec{x},\vec{x}\right)$ *от n переменных называется формальное выражение*

$$\varphi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^ix^j, \qquad (5.2.1)$$

где

$$\varphi_{ij}=\varphi_{ji}\left(i,j=1,2,\ldots,n\right)$$

— симметрические вещественные коэффициенты. •

В общей записи квадратичной формы (5.2.1) симметрическая матрица $\Phi=\left(\varphi_{ij}\right)$ называется ***матрицей квадратичной формы*** $\varphi\left(\vec{x},\vec{x}\right)$.

Преобразование матрицы квадратичной формы при изменении базиса. Пусть в пространстве R^n фиксированы два ортонормированных базиса: ***старый базис*** $\left\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\right\}$ и ***новый базис*** $\left\{\vec{e}_{1'},\vec{e}_{2'},\ldots,\vec{e}_{n'}\right\}$. Формула преобразования векторов старого базиса в векторы нового базиса имеет вид (3.3.1):

$$\vec{e}_{i'} = \sum_{j=1}^{n} A_{i'}^{j} \vec{e}_j, \tag{5.2.2}$$

где $\left(A_{i'}^{j}\right)$ — ортогональная матрица перехода от старого базиса к новому базису.

Координаты произвольного вектора $\vec{x}$ при переходе от ***нового базиса к старому базису*** преобразуются при помощи матрицы $\left(A_{j'}^{i}\right)$ по формулам

$$x^i = \sum_{k'=1}^{n'} A_{k'}^{i} x^{k'}. \tag{5.2.3}$$

Квадратичную форму запишем в старом и новом базисах:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} x^i x^j; \tag{5.2.4}$$

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{k'=1}^{n'} \sum_{p'=1}^{n'} \varphi_{k'p'} x^{k'} x^{p'}. \tag{5.2.5}$$

Преобразуем формулу (5.2.4), используя формулу (5.2.3):

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} x^i x^j = \sum_{k'=1}^{n'} \sum_{p'=1}^{n'} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} A_{k'}^{i} \varphi_{ij} A_{p'}^{j} \right) x^{k'} x^{p'} =$$

$$= \sum_{k'=1}^{n'} \sum_{p'=1}^{n'} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} A_{i}^{k'} \varphi_{ij} A_{p'}^{j} \right) x^{k'} x^{p'},$$

где $A_i^{k'}$ — элементы матрицы, транспонированной по отношению к матрице $A_{k'}^{i}$ (матрица $\left(A_{i'}^{j}\right)$ ортогональная, следовательно, $A^{-1} = A^* = A^T$). Сравнивая с формулой (5.2.5), получаем

$$\varphi_{k'p'} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{i}^{k'} \varphi_{ij} A_{p'}^{j}, \tag{5.2.6}$$

или

$$\Phi' = A^T \Phi A. \tag{5.2.7}$$

При изменении базиса матрица квадратичной формы преобразуется по формуле (5.2.7) и, так как справедливо равенство $A^T = A^{-1}$, *ранг* *матрицы* Φ *квадратичной формы* $\varphi\left(\vec{x},\vec{x}\right)$ *равен рангу матрицы* Φ', *то есть не зависит от выбора базиса. По этой причине ранг матрицы* Φ *называется* ***рангом квадратичной формы*** $\varphi\left(\vec{x},\vec{x}\right)$. *Квадратичная форма, имеющая ранг, совпадающий с размерностью пространства n, называется* ***невырожденной***.

Оператор $\hat{T}: R^n \to R^n$ с матрицей, равной матрице квадратичной формы $\varphi\left(\vec{x},\vec{x}\right)$, то есть $T = \Phi$, называется ***оператором***, ***ассоциированным*** с квадратичной формой.

При переходе к новому базису $\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\}$ с матрицей перехода $A = \left(A_{i'}^{j}\right)$ матрица квадратичной формы $\varphi\left(\vec{x},\vec{x}\right)$ преобразуется по формуле

$$\Phi' = A^T \Phi A,$$

а матрица оператора по формуле

$$T' = \left(A^{-1}\right)^T TA^T.$$

Таким образом, матрицы квадратичной формы и оператора преобразуются, вообще говоря, ***неодинаково***. Но так как матрица перехода ортогональная и, следовательно, для неё выполняется условие

$$A^T = A^{-1},$$

то с учётом симметричности матрицы квадратичной формы и, естественно, матрицы ассоциированного оператора, получаем

$$T' = \left(A^{-1}\right)^T TA^T = \left(A^T\right)^T TA^T = ATA^T = A^T TA.$$

Таким образом, закон преобразования матриц квадратичной формы и матрицы оператора, по существу, один и тот же.

В силу того, что оператор $\hat{T}$ самосопряжённый (симметрический), можно выбрать базис $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ из собственных векторов оператора $\hat{T}$, такой, что в нём матрица оператора, являющаяся одновременно и матрицей квадратичной формы, приводится к диагональному виду. Следовательно, квадратичная форма примет в этом базисе следующий вид:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \mu_i \left(\overline{x}^i\right)^2, \tag{5.2.8}$$

который называется ***каноническим видом квадратичной формы***. Коэффициенты μ_i $(i = 1, 2, \ldots, n)$ называются ***каноническими коэффициентами***, а $\overline{x}^i$ — это новые координаты в базисе собственных векторов оператора $\hat{T}$, который также называется ***каноническим базисом*** квадратичной формы. Нетрудно видеть, что канонический вид квадратичной формы определяется неединственным образом. Отметим также, что в общем случае не все канонические коэффициенты μ_i отличны от нуля — форма может быть вырожденной.

Знакоопределённые квадратичные формы. Очевидно, что ***число отличных от нуля канонических коэффициентов*** μ_k $(k = 1, 2, \ldots, m)$***, где в общем случае*** $m \neq n$ (случай вырождения формы), ***равно рангу квадратичной формы***. Эти коэффициенты могут быть как положительными, так и отрицательными. Справедлива, однако, следующая важная теорема.

Теорема 5.2.1 (закон инерции квадратичных форм). *Число положительных и число отрицательных коэффициентов в каноническом виде (5.2.8) квадратичной формы* $\varphi\left(\vec{x}, \vec{x}\right)$ *не зависит от выбора канонического базиса.*

Доказательство. Пусть в некотором базисе $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ пространства R^n задана квадратичная форма

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij} x^i x^j,$$

где

$$\vec{x} = \sum_{k=1}^{n} x^k \vec{a}_k.$$

Предположим, что в пространстве R^n зафиксированы два канонических базиса

$$\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset R^n, \left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset R^n,$$

в которых разложения произвольного вектора $\vec{x} \in R^n$ имеют соответственно вид

$$\vec{x} = \sum_{k=1}^{n} \xi^k \vec{e}_k, \ \vec{x} = \sum_{k=1}^{n} \eta^k \vec{g}_k. \qquad (5.2.9)$$

В базисе $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ форма $\varphi\left(\vec{x}, \vec{x}\right)$ имеет вид

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{k} \mu_i \left(\xi^i\right)^2 - \sum_{i=k+1}^{k+m} \mu_i \left(\xi^i\right)^2 =$$

$$= \mu_1 \left(\xi^1\right)^2 + \ldots + \mu_k \left(\xi^k\right)^2 - \mu_{k+1} \left(\xi^{k+1}\right)^2 - \ldots - \mu_{k+m} \left(\xi^{k+m}\right)^2, \qquad (5.2.10)$$

а в базисе $\left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\}$ — вид

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{j=1}^{p} \nu_j \left(\eta^j\right)^2 - \sum_{j=p+1}^{p+q} \nu_j \left(\eta^j\right)^2 =$$

$$= \nu_1 \left(\eta^1\right)^2 + \ldots + \nu_p \left(\eta^p\right)^2 - \nu_{p+1} \left(\eta^{p+1}\right)^2 - \ldots - \nu_{p+q} \left(\eta^{p+q}\right)^2. \qquad (5.2.11)$$

В выражениях (5.2.10) и (5.2.11) числа $\mu_1, \mu_2, \ldots, \mu_k$ и $\nu_1, \nu_2, \ldots, \nu_p$ считаются положительными. Покажем, что $k = p$ и $m = q$.

Допустим, что $k < p$. Рассмотрим в пространстве R^n подпространство $L^{(1)}$, порождённое векторами

$$\left\{\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_p\right\},$$

и подпространство $L^{(2)}$, порождённое векторами

$$\left\{\vec{e}_{k+1}, \vec{e}_{k+2}, \ldots, \vec{e}_n\right\}.$$

Сумма их размерностей равна

$$\dim L^{(1)} + \dim L^{(2)} = p + (n-k) = n + (p-k) > n,$$

так как $p - k > 0$. По теореме о связи размерностей подпространств имеем:

$$\dim\left(L^{(1)} \cap L^{(2)}\right) + \dim\left(L^{(1)} + L^{(2)}\right) = \dim L^{(1)} + \dim L^{(2)},$$

следовательно, $\dim\left(L^{(1)} \cap L^{(2)}\right) \neq 0$. Последнее означает, что

$$\left(\exists \vec{x} \neq \vec{0}\right): \vec{x} \in L^{(1)} \cap L^{(2)}.$$

Этот вектор можно представить в виде разложений

$$\vec{x} = \sum_{i=1}^{p} \eta^i \vec{g}_i, \ \vec{x} = \sum_{j=k+1}^{n} \xi^j \vec{e}_j.$$

По формуле (5.2.11) для вектора $\vec{x}$ имеем

$$\varphi\left(\vec{x}, \vec{x}\right) = \nu_1\left(\eta^1\right)^2 + \nu_2\left(\eta^2\right)^2 + \ldots + \nu_p\left(\eta^p\right)^2 > 0,$$

а по формуле (5.2.10)

$$\varphi\left(\vec{x}, \vec{x}\right) = -\mu_{k+1}\left(\xi^{k+1}\right)^2 - \mu_{k+2}\left(\xi^{k+2}\right)^2 - \ldots - \mu_{k+m}\left(\xi^{k+m}\right)^2 \leq 0.$$

Получили противоречие, из которого следует, что $k \geq p$.

Аналогично можно показать, что $p \geq k$. Из этих двух неравенств следует, что $p = k$. Точно также показывается, что $m = q$. ••

Полное число членов, входящих в канонический вид квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$, равное её рангу, называется ***индексом инерции*** квадратичной формы, а число положительных и число отрицательных членов называются соответственно ***положительным индексом инерции*** и ***отрицательным индексом инерции*** квадратичной формы.

Определение 5.2.2. *Квадратичная форма* $\varphi\left(\vec{x}, \vec{x}\right)$ *называется* ***положительно определённой (положительной)****, если все её канонические коэффициенты положительны, то есть если*

$$\left(\forall \vec{x} \in R^n\right): \vec{x} \neq \vec{0}, \ \varphi\left(\vec{x}, \vec{x}\right) > 0,$$

и ***отрицательно определённой (отрицательной)****, если*

$$\left(\forall \vec{x} \in R^n\right): \vec{x} \neq \vec{0}, \ \varphi\left(\vec{x}, \vec{x}\right) < 0. \bullet$$

Теорема 5.2.2. *Квадратичная форма* $\varphi\left(\vec{x}, \vec{x}\right)$ *является положительно определённой в том и только в том случае, если её положительный индекс инерции равен размерности пространства* R^n.

Д о к а з а т е л ь с т в о. Доказательство этой теоремы — прямое следствие определения 5.2.2. ••

Значения положительного и отрицательного индексов инерции квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$ определяются по знакам угловых миноров матрицы формы в каком-либо базисе, а именно: справедлива следующая теорема — ***критерий Сильвестра положительной определённости квадратичной формы***.

Теорема 5.2.3. *Квадратичная форма, определённая на векторах пространства* R^n, *является положительно определённой в том и только в том случае, если в каком-либо базисе* $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset R^n$ *все угловые миноры её матрицы* $\varphi_{ij}\left(\vec{a}_i, \vec{a}_j\right)$ *положительны, то есть*

$$M_1 = \varphi_{11} > 0,\ M_2 = \begin{vmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{vmatrix} > 0,\ \ldots,$$

$$\ldots,\ M_n = \begin{vmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{vmatrix} > 0.$$

Д о к а з а т е л ь с т в о. Применим индукцию по числу переменных. Пусть квадратичная форма зависит от одного переменного, то есть

$$\varphi\left(\vec{x}, \vec{x}\right) = \varphi_{11}\left(x^1\right)^2.$$

Очевидно, что в этом случае утверждение теоремы справедливо.

Предположим, что утверждение теоремы справедливо для квадратичной формы

$$\psi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} \varphi_{ij} x^i x^j,$$

зависящей от $n-1$ переменных, и рассмотрим квадратичную форму

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} x^i x^j .$$

Пусть квадратичная форма положительно определена. Покажем, что в этом случае все угловые миноры её матрицы положительны. Для этого представим форму в виде:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} x^i x^j =$$

$$= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \varphi_{ij} x^i x^j + 2\sum_{i=1}^{n-1} \varphi_{ij} x^i x^n + \varphi_{nn} \left(x^n\right)^2. \qquad (5.2.12)$$

Например, для квадратичной формы, зависящей от трёх переменных, имеем:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{3} \sum_{j=1}^{3} \varphi_{ij} x^i x^j = \varphi_{11} \left(x^1\right)^2 + \varphi_{12} x^1 x^2 + \varphi_{13} x^1 x^3 +$$

$$+\varphi_{21} x^2 x^1 + \varphi_{22} \left(x^2\right)^2 + \varphi_{23} x^2 x^3 + \varphi_{31} x^3 x^1 + \varphi_{32} x^3 x^2 + \varphi_{33} \left(x^3\right)^2 =$$

$$= \left[\varphi_{11} \left(x^1\right)^2 + \varphi_{12} x^1 x^2 + \varphi_{21} x^2 x^1 + \varphi_{22} \left(x^2\right)^2\right] +$$

$$+\left[\varphi_{13} x^1 x^3 + \varphi_{23} x^2 x^3 + \varphi_{31} x^3 x^1 + \varphi_{32} x^3 x^2\right] + \varphi_{33} \left(x^3\right)^2 =$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \varphi_{ij} x^i x^j + 2\left[\varphi_{13} x^1 x^3 + \varphi_{23} x^2 x^3\right] + \varphi_{33} \left(x^3\right)^2 =$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \varphi_{ij} x^i x^j + 2\sum_{i=1}^{2} \varphi_{ij} x^i x^3 + \varphi_{33} \left(x^3\right)^2.$$

В представлении (5.2.12) квадратичная форма

$$\psi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \varphi_{ij} x^i x^j ,$$

зависящая от $n-1$ переменных, будет положительно определённой. Действительно, рассмотрим форму

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} x^i x^j$$

на векторах

$$\vec{x} = \sum_{i=1}^{n-1} x^i \vec{a}_i .$$

В этом случае в представлении (5.2.12) слагаемые

$$2\sum_{i=1}^{n-1}\varphi_{ij}x^i x^n+\varphi_{nn}\left(x^n\right)^2=0,$$

так как $x^n\equiv 0$. Поэтому, если $\psi\left(\vec{x},\vec{x}\right)\le 0$, то и $\varphi\left(\vec{x},\vec{x}\right)\le 0$, что противоречит предположению о положительной определённости

$$\varphi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i x^j .$$

По предположению индукции все угловые миноры матрицы формы

$$\psi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\varphi_{ij}x^i x^j$$

положительны, то есть

$$M_1=\varphi_{11}>0,\ M_2=\begin{vmatrix}\varphi_{11} & \varphi_{12}\\ \varphi_{21} & \varphi_{22}\end{vmatrix}>0,$$

$$\ldots,\ M_{n-1}=\begin{vmatrix}\varphi_{11} & \varphi_{12} & \cdots & \varphi_{1\,n-1}\\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2\,n-1}\\ \multicolumn{4}{c}{\dotfill}\\ \varphi_{n-1\,1} & \varphi_{n-1\,2} & \cdots & \varphi_{n-1\,n-1}\end{vmatrix}>0.$$

Покажем, что положителен и минор порядка n матрицы квадратичной формы

$$\varphi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i x^j ,$$

то есть

$$M_n=\begin{vmatrix}\varphi_{11} & \varphi_{12} & \cdots & \varphi_{1\,n}\\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2\,n}\\ \multicolumn{4}{c}{\dotfill}\\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{n\,n}\end{vmatrix}>0.$$

Вспомним, что положительно определённая квадратичная форма в некотором базисе приводится к сумме квадратов

$$\varphi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i x^j=\left(x^{1'}\right)^2+\left(x^{2'}\right)^2+\ldots+\left(x^{n'}\right)^2,$$

где штрихами обозначены координаты вектора $\vec{x}$ в новом базисе. Определитель матрицы квадратичной формы в этом новом базисе равен, очевидно, единице. Матрица квадратичной формы преобразуется к новому базису по формуле

$$\Phi' = A^T \Phi A,$$

где A — матрица перехода от старого к новому базису. Поэтому

$$\det \Phi' = \det A^T \det \Phi \det A = \det \Phi \left(\det A\right)^2.$$

Так как $\det A \neq 0$ и $\det \Phi' = 1 > 0$, то

$$M_n = \det \Phi = \frac{\det \Phi'}{\left(\det A\right)^2} > 0. \qquad (5.2.13)$$

Пусть все угловые миноры матрицы квадратичной формы положительны. Покажем, что квадратичная форма

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_{ij} x^i x^j$$

является положительно определённой. По предположению индукции квадратичная форма

$$\psi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \varphi_{ij} x^i x^j,$$

зависящая от $n-1$ переменных, положительно определена. Поэтому в некотором базисе имеем:

$$\psi\left(\vec{x}, \vec{x}\right) = \sum_{i'=1}^{n-1'} \sum_{j'=1}^{n-1'} \varphi_{i'j'} x^{i'} x^{j'} = \left(x^{1'}\right)^2 + \left(x^{2'}\right)^2 + \ldots + \left(x^{n-1'}\right)^2.$$

Тогда по представлению (5.2.12) в этом базисе имеем следующий результат:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i'=1}^{n'} \sum_{j'=1}^{n'} \varphi_{i'j'} x^{i'} x^{j'} = \left(x^{1'}\right)^2 + \left(x^{2'}\right)^2 + \ldots + \left(x^{n-1'}\right)^2 +$$

$$+2\left(\xi_{1'n'} x^{1'} x^{n'} + \xi_{2'n'} x^{2'} x^{n'} + \ldots + \xi_{n-1'n'} x^{n-1'} x^{n'}\right) + \varphi_{n'n'} \left(x^{n'}\right)^2.$$

Здесь $\xi_{i\,n}$ — некоторые новые коэффициенты. Выделим полные квадраты:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i'=1}^{n'}\sum_{j'=1}^{n'} \varphi_{i'j'} x^{i'} x^{j'} = \left(x^{1'} + \xi_{1'n'} x^{n'}\right)^2 + \left(x^{2'} + \xi_{2'n'} x^{n'}\right)^2 +$$

$$+ \ldots + \left(x^{n-1'} + \xi_{n-1'n'} x^{n'}\right)^2 + \varphi \cdot \left(x^{n'}\right)^2,$$

где φ — новый коэффициент. Переходя к новому базису, положим

$$x^{i'} + \xi_{1'in'} x^{n'} = z^i$$

и $x^{n'} = z^n$. Тогда имеем:

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij} x^i x^j = \left(z^1\right)^2 + \left(z^2\right)^2 + \ldots + \left(z^{n-1}\right)^2 + \varphi \cdot \left(z^n\right)^2.$$

Определитель матрицы квадратичной формы равен φ, а его знак, как следует из формулы (5.2.13), совпадает со знаком M_n и, следовательно, $\varphi > 0$. Теперь видно, что квадратичная форма

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij} x^i x^j$$

положительно определённая. ••

Следствие из теоремы 5.2.3. *Квадратичная форма является отрицательно определённой в том и только в том случае, если угловые миноры её матрицы чётного порядка положительны, а нечётного порядка — отрицательны.*

Доказательство. Квадратичная форма

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij} x^i x^j$$

является отрицательно определённой в том и только в том случае, если квадратичная форма

$$-\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \left(-\varphi_{ij}\right) x^i x^j$$

положительно определена. По теореме 5.2.3 все угловые миноры матрицы квадратичной формы $-\varphi\left(\vec{x}, \vec{x}\right)$ в этом случае положительны, то есть

$$M_1^- = -\varphi_{11} > 0,\ M_2^- = \begin{vmatrix} -\varphi_{11} & -\varphi_{12} \\ -\varphi_{21} & -\varphi_{22} \end{vmatrix} > 0,$$

$$\ldots,\ M_n^- = \begin{vmatrix} -\varphi_{11} & -\varphi_{12} & \ldots & -\varphi_{1n} \\ -\varphi_{21} & -\varphi_{22} & \ldots & -\varphi_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -\varphi_{n1} & -\varphi_{n2} & \ldots & -\varphi_{nn} \end{vmatrix} > 0.$$

Но тогда по свойствам определителей

$$M_1 = \varphi_{11} > 0,\ M_2 = \begin{vmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{vmatrix} > 0,\ M_3 = \begin{vmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} \\ \varphi_{21} & \varphi_{22} & \varphi_{23} \\ \varphi_{31} & \varphi_{32} & \varphi_{33} \end{vmatrix} < 0,\ \ldots \bullet\bullet$$

Приведение квадратичной формы к каноническому виду ортогональным преобразованием. Выше показано, что в пространстве R^n существует базис, в котором квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$ имеет канонический вид

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \mu_i \left(x^i\right)^2$$

(черта над символом координаты опущена). Опишем последовательность действий по приведению квадратичной формы к каноническому виду методом, который назовём ***методом ассоциированного оператора***.

1. *Записываем симметричную матрицу $\Phi = \left(\varphi_{ij}\right)$ квадратичной формы (5.2.1) и ставим в соответствие квадратичной форме ассоциированный оператор $\hat{T}$ с матрицей $T = \Phi$.*

2. *Находим корни характеристического многочлена*

$$T(\mu) = \det\left(T - \mu I\right)$$

оператора $\hat{T}$.

3. *Записываем квадратичную форму (5.2.1) в каноническом виде (5.2.8) и определяем её положительный и отрицательный индексы инерции.*

4. *Находим систему собственных векторов оператора $\hat{T}$ и ортонормируем её. Получим ортонормированную систему собственных векторов оператора $\hat{T}$:*

$$\vec{x}_1 = \sum_{i=1}^{n} x_1^i \vec{e}_i,\ \vec{x}_2 = \sum_{i=1}^{n} x_2^i \vec{e}_i,\ \ldots,\ \vec{x}_n = \sum_{i=1}^{n} x_n^i \vec{e}_i. \tag{5.2.14}$$

Координаты векторов (5.2.14) позволяют написать матрицу $\left(A_{j'}^{i}\right)$ перехода от старого базиса к новому базису. Так как эта матрица ортогональная, то, транспонируя её, получаем матрицу, при помощи которой преобразуются старые координаты в новые координаты.

Теорема 5.2.4. *Для квадратичной формы* $\varphi\left(\vec{x}, \vec{x}\right)$*, определённой на векторах пространства* E^n*, существует канонический базис* $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$*, такой, что квадратичная форма в этом базисе принимает вид*

$$\varphi\left(\vec{x}, \vec{x}\right) = \left(x^1\right)^2 + \left(x^2\right)^2 + \ldots + \left(x^p\right)^2 - \left(x^{p+1}\right)^2 - \left(x^{p+2}\right)^2 - \ldots - \left(x^{p+q}\right)^2. \quad (5.2.15)$$

Д о к а з а т е л ь с т в о. Приведём квадратичную форму к каноническому виду

$$\varphi\left(\vec{x}, \vec{x}\right) = \mu_1\left(\xi^1\right)^2 + \ldots + \mu_p\left(\xi^p\right)^2 - \mu_{p+1}\left(\xi^{p+1}\right)^2 - \ldots - \mu_{p+q}\left(\xi^{p+q}\right)^2.$$

Далее совершим преобразование координат вида

$$x^1 = \sqrt{\mu_1}\xi^1,\ \ldots,\ x^p = \sqrt{\mu_p}\xi^p,\ x^{p+1} = \sqrt{\mu_{p+1}}\xi^{p+1},\ \ldots,$$
$$x^{p+q} = \sqrt{\mu_{p+q}}\xi^{p+q}.$$

Получаем для квадратичной формы представление (5.2.15). ••

Приведение квадратичной формы к каноническому виду методом выделения полного квадрата[*]. Пусть на векторах пространства R^n задана квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$. Вид этой квадратичной формы в общем случае даётся выражением

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i x^j.$$

Рассмотрим вопрос о приведении квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$ к сумме квадратов методом, отличным от метода ассоциированного оператора — ***методом выделения полного квадрата***.

Теорема 5.2.5. *Пусть* $\varphi\left(\vec{x}, \vec{x}\right)$ *— произвольная квадратичная форма, определённая на векторах пространства* E^n*. Тогда найдётся базис*

$$\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\} \subset E^n,$$

такой, что в этом базисе квадратичная форма примет вид:

$$\varphi\left(\vec{x}, \vec{x}\right) = \mu_1 \left(x^{1'}\right)^2 + \mu_2 \left(x^{2'}\right)^2 + \ldots + \mu_n \left(x^{n'}\right)^2, \tag{5.2.16}$$

где $\mu_1, \mu_2, \ldots, \mu_n$ *— некоторые фиксированные числа из поля* P.

Д о к а з а т е л ь с т в о. Пусть форма $\varphi\left(\vec{x}, \vec{x}\right)$ содержит лишь одну координату с ненулевым коэффициентом, например, x^1, то есть имеет вид

$$\varphi\left(\vec{x}, \vec{x}\right) = \varphi_{11} \left(x^1\right)^2.$$

Утверждение теоремы, очевидно, выполняется. Пусть утверждение теоремы справедливо для любых квадратичных форм, содержащих $m-1$ координат. Рассмотрим форму, содержащую m координат:

$$\begin{gathered}
\varphi\left(\vec{x}, \vec{x}\right) = \varphi_{11} \left(x^1\right)^2 + \varphi_{12} x^1 x^2 + \varphi_{13} x^1 x^3 + \ldots + \\
+\varphi_{1m} x^1 x^m + \varphi_{21} x^2 x^1 + \varphi_{22} \left(x^2\right)^2 + \varphi_{23} x^2 x^3 + \ldots + \varphi_{2m} x^2 x^m + \\
+\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots+ \\
+\varphi_{m1} x^m x^1 + \varphi_{m2} x^m x^2 + \varphi_{m3} x^m x^3 + \ldots + \varphi_{mm} \left(x^m\right)^2 = \\
= \varphi_{11} \left(x^1\right)^2 + 2\varphi_{12} x^1 x^2 + \varphi_{22} \left(x^2\right)^2 + \ldots + \varphi_{mm} \left(x^m\right)^2.
\end{gathered} \tag{5.2.17}$$

Если хотя бы при одном квадрате, например, при $\left(x^m\right)^2$, коэффициент отличен от нуля, то соберём все слагаемые, содержащие x^m:

$$2\varphi_{1m} x^1 x^m + 2\varphi_{2m} x^2 x^m + \ldots + 2\varphi_{m-1\,m} x^{m-1} x^m + \varphi_{mm} \left(x^m\right)^2.$$

Выделим в этой группе слагаемых полный квадрат:

$$\begin{gathered}
2\varphi_{1m} x^1 x^m + 2\varphi_{2m} x^2 x^m + \ldots + 2\varphi_{m-1\,m} x^{m-1} x^m + \varphi_{mm} \left(x^m\right)^2 = \\
= \frac{1}{\varphi_{mm}} \left(\varphi_{1m} x^1 + \varphi_{2m} x^2 + \ldots + \varphi_{m-1\,m} x^{m-1} + \varphi_{mm} x^m\right)^2 - \\
- \frac{1}{\varphi_{mm}} \left(\varphi_{1m} x^1 + \varphi_{2m} x^2 + \ldots + \varphi_{m-1\,m} x^{m-1}\right)^2.
\end{gathered} \tag{5.2.18}$$

Тогда квадратичная форма (5.2.17) примет вид:

$$\varphi\left(\vec{x},\vec{x}\right)=\frac{1}{\varphi_{mm}}\left(\varphi_{1m}x^1+\varphi_{2m}x^2+\ldots+\varphi_{m-1\,m}x^{m-1}+\varphi_{mm}x^m\right)^2+$$
$$+\psi\left(\vec{x},\vec{x}\right).$$

Здесь квадратичная форма $\psi\left(\vec{x},\vec{x}\right)$ зависит уже только от $m-1$ координат $x^1, x^2, \ldots, x^{m-1}$. Положим

$$\begin{aligned}
\xi^1 &= x^1,\\
\xi^2 &= x^2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots,\\
\xi^{m-1} &= x^{m-1},\\
\xi^m &= \varphi_{1m}x^1+\varphi_{2m}x^2+\ldots+\varphi_{m-1\,m}x^{m-1}+\varphi_{mm}x^m,\\
\xi^{m+1} &= x^{m+1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\\
\xi^n &= x^n.
\end{aligned}\qquad(5.2.19)$$

Нетрудно видеть, что определитель матрицы преобразования (5.2.19) $\varphi_{mm}\neq 0$. Поэтому преобразование координат (5.2.19) вызвано переходом к новому базису, причём, как известно, матрица перехода является транспонированной обратной матрицей преобразования (5.2.19). По предположению индукции квадратичную форму $\psi\left(\vec{x},\vec{x}\right)$, зависящую уже только от $m-1$ координат $x^1, x^2, \ldots, x^{m-1}$, путём перехода к новому базису можно привести к виду (5.2.16). Поэтому и форма $\varphi\left(\vec{x},\vec{x}\right)$ приведётся к виду (5.2.16).

Пусть теперь все коэффициенты при квадратах равны нулю, то есть

$$\left(\forall i=1,2,\ldots,m\right)\varphi_{ii}=0.$$

Допустим, например, что $\varphi_{12}\neq 0$. Положим

$$x^1=\xi^1+\xi^2,\ x^2=\xi^1-\xi^2,\ x^3=\xi^3,\ \ldots,\ x^n=\xi^n.$$

Это преобразование соответствует переходу к базису

$$\left\{\vec{e}_{1'}, \vec{e}_{2'}, \vec{e}_{3'}, \ldots, \vec{e}_{n'}\right\} = \left\{\vec{e}_1 + \vec{e}_2, \vec{e}_1 - \vec{e}_2, \vec{e}_3, \ldots, \vec{e}_n\right\}$$

с матрицей перехода

$$A = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 1 & -1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Определитель этой матрицы $\det A = -2$. При этом преобразовании

$$x^1 x^2 = \left(\xi^1\right)^2 - \left(\xi^2\right)^2$$

и приходим к предыдущему случаю квадрата с не равным нулю коэффициентом. ••

5.3. Кривые и поверхности второго порядка

Определение и общее уравнение поверхности второго порядка. Поверхности второго порядка играют большую роль в геометрии в первую очередь по той причине, что эти поверхности могут быть полностью исследованы алгебраическими методами. Кроме этого, теория поверхностей второго порядка имеет большие приложения в физике и технике.

Определение 5.3.1. ***Поверхностью второго порядка*** *в пространстве* R^3 *называется множество точек* $\vec{x} \in R^3$, *координаты которых удовлетворяют уравнению*

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\varphi_{ij}x^i x^j + 2\sum_{k=1}^{3}b_k x^k + c = 0. \bullet \qquad (5.3.1)$$

Это уравнение можно записать в виде

$$\varphi\left(\vec{x}, \vec{x}\right) + 2b\left(\vec{x}\right) + c = 0. \qquad (5.3.2)$$

где

$$\varphi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{3}\sum_{j=1}^{3}\varphi_{ij}x^i x^j$$

— квадратичная форма, а

$$b\left(\vec{x}\right)=\sum_{k=1}^{3}b_k x^k$$

— ***линейная форма*** от радиус-вектора $\vec{x}$, c — некоторая постоянная. Считая, что в пространстве R^3 зафиксирован канонический ортонормированный репер, покажем, что можно так выбрать новый репер, чтобы уравнение поверхности второго порядка преобразовалось к наиболее простому виду, который называется ***каноническим видом*** уравнения поверхности второго порядка.

Из теории квадратичных форм следует, что существует такое ортогональное преобразование координат

$$y^i=\sum_{k=1}^{3}A_k^i x^k, \tag{5.3.3}$$

где $i=1,2,3$, что в новых координатах квадратичная форма $\varphi\left(\vec{x},\vec{x}\right)$ приведётся к каноническому виду

$$\varphi\left(\vec{x},\vec{x}\right)=\sum_{i=1}^{3}\mu_i\left(y^i\right)^2. \tag{5.3.4}$$

Поэтому уравнение (5.3.1) в новых координатах запишется в виде

$$\sum_{i=1}^{3}\mu_i\left(y^i\right)^2+2\sum_{i=1}^{3}l_i y^i+c=0, \tag{5.3.5}$$

где l_i — новые коэффициенты линейной формы. Если в уравнении (5.3.5) некоторый коэффициент $\mu_i\neq 0\ (i=1,2,3)$, то путём переноса начала репера O можно добиться исчезновения линейного члена с тем же номером i. Действительно, пусть, например, $\mu_1\neq 0$. Тогда:

$$\mu_1\left(y^1\right)^2+2l_1y^1=\mu_1\left[\left(y^1\right)^2+2\frac{l_1}{\mu_1}y^1\right]=$$

$$=\mu_1\left[\left(y^1\right)^2+2y^1\frac{l_1}{\mu_1}+\left(\frac{l_1}{\mu_1}\right)^2-\left(\frac{l_1}{\mu_1}\right)^2\right]=\mu_1\left(y^1+\frac{l_1}{\mu_1}\right)^2-\frac{l_1^2}{\mu_1}.$$

Перенесём начало репера в точку $O'\left(-\frac{l_1}{\mu_1};0;0\right)$, полагая $\xi^1=y^1+\frac{l_1}{\mu_1}$, получим

$$\mu_1\left(y^1\right)^2+2l_1y^1=\mu_1\left(\xi^1\right)^2-\frac{l_1^2}{\mu_1}.$$

Видим, что слагаемое второго порядка имеет тот же коэффициент μ_1, слагаемое первого порядка исчезло, а свободный член уравнения получил значение $c-\frac{l_1^2}{\mu_1}$.

Пусть последовательно:

1) $\mu_1\neq 0$;

2) $\mu_1\neq 0$ и $\mu_2\neq 0$;

3) $\mu_1\neq 0$, $\mu_2\neq 0$ и $\mu_3\neq 0$.

Проводя преобразования переноса по типу преобразования, проведённого выше, получим уравнения следующих трёх видов:

$$\mu_1\left(x^1\right)^2+2l_2x^2+2l_3x^3+h_1=0; \tag{5.3.6}$$

$$\mu_1\left(x^1\right)^2+\mu_2\left(x^2\right)^2+2l_3x^3+h_2=0; \tag{5.3.7}$$

$$\mu_1\left(x^1\right)^2+\mu_2\left(x^2\right)^2+\mu_3\left(x^3\right)^2+h_3=0. \tag{5.3.8}$$

Здесь новые координаты снова обозначены $x^i\equiv\xi^i\ (i=1,2,3)$, а новый свободный член соответственно обозначен

$$h_1=c-\frac{l_1^2}{\mu_1},\ h_2=c-\frac{l_1^2}{\mu_1}-\frac{l_2^2}{\mu_2},\ h_3=c-\frac{l_1^2}{\mu_1}-\frac{l_2^2}{\mu_2}-\frac{l_3^2}{\mu_3},$$

причём координаты упорядочены по номерам так, чтобы сначала шли координаты, у которых в квадратичной форме коэффициенты $\mu_k\neq 0$, а все $\mu_k=0$ имели бы номера $k>r$, где $r=1,2,3$ — число отличных от нуля коэффициентов. При таком преобразовании можем получить следующие варианты:

1) $r=1$, то есть $\mu_1\neq 0$, а $\mu_2=\mu_3=0$ и $l_2=l_3=0$;

2) $r=2$, то есть $\mu_1\neq 0$, $\mu_2\neq 0$, а $\mu_3=0$ и $l_3=0$;

3) $r=3$, то есть $\mu_1\neq 0$, $\mu_2\neq 0$, $\mu_3\neq 0$.

Уравнения (5.3.6) — (5.3.8) записываются теперь в общем виде

$$\sum_{i=1}^{r} \mu_i \left(x^i\right)^2 + h_r = 0, \qquad (5.3.9)$$

где $r = 1, 2, 3$.

Определение 5.3.2. *Поверхность, определяемая уравнением вида (5.3.9) при* $r = 1, 2, 3$, *называется* ***центральной поверхностью***. •

При условии $h_r \neq 0$ центральная поверхность называется ***истинной центральной поверхностью***, а при условии $h_r = 0$ — ***конической поверхностью***.

Определение 5.3.3. ***Центром поверхности*** *называется точка*

$$\vec{x}_0 = x_0^1 \vec{e}_1 + x_0^2 \vec{e}_2 + x_0^3 \vec{e}_3,$$

удовлетворяющая следующему условию: если точка

$$\vec{x}_0^{\,+} = \left(x_0^1 + x^1\right)\vec{e}_1 + \left(x_0^2 + x^2\right)\vec{e}_2 + \left(x_0^3 + x^3\right)\vec{e}_3$$

принадлежит поверхности, то точка

$$\vec{x}_0^{\,-} = \left(x_0^1 - x^1\right)\vec{e}_1 + \left(x_0^2 - x^2\right)\vec{e}_2 + \left(x_0^3 - x^3\right)\vec{e}_3$$

также принадлежит поверхности. •

Если поверхность описывается каноническим уравнением (5.3.9), то у неё существуют центры. Действительно, точка с координатами $x_0^1 = x_0^2 = x_0^3 = 0$, очевидно, является центром. Поэтому эти поверхности и называются центральными.

Покажем, что у центральных поверхностей нет других центров. Пусть точка $M_0\left(x_0^1; x_0^2; x_0^3\right)$ является центром поверхности. По определению из

$$\mu_1\left(x_0^1 + x^1\right)^2 + \mu_2\left(x_0^2 + x^2\right)^2 + \mu_3\left(x_0^3 + x^3\right)^2 = 0$$

следует справедливость тождества

$$\mu_1\left(x_0^1 - x^1\right)^2 + \mu_2\left(x_0^2 - x^2\right)^2 + \mu_3\left(x_0^3 - x^3\right)^2 = 0.$$

Вычитая второе тождество из первого, получаем:

$$\mu_1 x_0^1 x^1 + \mu_2 x_0^2 x^2 + \mu_3 x_0^3 x^3 = 0.$$

Видим, что точка с $x^1 \neq 0$, но $x^2 = x^3 = 0$, удовлетворяет уравнению поверхности и $\mu_1 x_0^1 x^1 = 0$, то есть $x_0^1 = 0$.

Аналогично показывается, что $x_0^2 = x_0^3 = 0$.

Пусть $r = 1$ и $l_2 \neq 0$, $l_3 \neq 0$. Совершим ***ортогональное преобразование*** $\hat{T}$, задав его следующими формулами:

$$\begin{cases} z^1 = x^1; \\ z^2 = -l_2 x^2 \cdot \dfrac{1}{N} - l_3 x^3 \cdot \dfrac{1}{N}; \\ z^3 = -l_3 x^2 \cdot \dfrac{1}{N} + l_2 x^3 \cdot \dfrac{1}{N}. \end{cases} \Rightarrow \begin{pmatrix} z^1 \\ z^2 \\ z^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -l_2/N & -l_3/N \\ 0 & -l_3/N & l_2/N \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Здесь $N > 0$ — множитель, обеспечивающий ортогональность матрицы преобразования, а именно: так как у ортогональной матрицы сумма квадратов элементов каждой строки равна 1, то множитель N выбирается так, чтобы выполнялось условие

$$N^2 = l_2^2 + l_3^2.$$

Третья строка выбрана также из условия ортогональности матрицы. Чтобы установить ортогональность матрицы T достаточно проверить, что при выполнении условия $N^2 = l_2^2 + l_3^2$ и выборе третьей строки в указанном виде матрица преобразования T удовлетворяет условию

$$T^T = T^{-1}.$$

Действительно:

$$TT^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -l_2/N & -l_3/N \\ 0 & -l_3/N & l_2/N \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -l_2/N & -l_3/N \\ 0 & -l_3/N & l_2/N \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Таким образом, выбор множителя N в указанном виде действительно обеспечивает ортогональность преобразования.

В результате этого преобразования уравнение (5.3.6) принимает вид

$$\mu_1 \left(x^1\right)^2 = -2\left(l_2 x^2 + l_3 x^3\right) - h_1,$$

или

$$\mu_1\left(z^1\right)^2 = 2Nz^2 - h_1. \qquad (5.3.10)$$

Пусть $r = 2$ и $l_3 \neq 0$. Зададим ортогональное преобразование формулами:

$$z^1 = x^1;\ z^2 = x^2;\ z^3 = -l_3 x^3 \cdot \frac{1}{N}.$$

Тогда уравнение (5.3.7) примет вид:

$$\mu_1\left(x^1\right)^2 + \mu_2\left(x^2\right)^2 = -2l_3 x^3 - h_2$$

или

$$\mu_1\left(z^1\right)^2 + \mu_2\left(z^2\right)^2 = 2Nz^3 - h_2. \qquad (5.3.11)$$

Если $h_r \neq 0$, то осуществим ещё один параллельный перенос начала репера по формулам:

$$z^2 = z^{2'} + \frac{h_1}{2N},\ z^3 = z^{3'} + \frac{h_2}{2N}$$

соответственно в уравнениях (5.3.10) и (5.3.11). Если опустить штрихи, то преобразованные уравнения (5.3.10) и (5.3.11) примут соответственно вид

$$\mu_1\left(z^1\right)^2 = 2Nz^2, \qquad (5.3.12)$$

$$\mu_1\left(z^1\right)^2 + \mu_2\left(z^2\right)^2 = 2Nz^3. \qquad (5.3.13)$$

Поверхность, определяемая уравнениями (5.3.12) и (5.3.13), называется ***нецентральной*** поверхностью — ***вырожденной*** в случае (5.3.12) и ***невырожденной*** — в случае (5.3.13).

Эллипс и гипербола. В пространстве R^2 существуют две поверхности (***кривые***) ***второго порядка***, являющиеся ***центральными***. Эти кривые описываются каноническими уравнениями следующего вида:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1, \qquad (5.3.14)$$

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1 \qquad (5.3.15)$$

и называются соответственно ***эллипсом*** и ***гиперболой***. Общий вид этих кривых второго порядка приведён на рис. 5.3, *а* и 5.3, *б*. Изучим некоторые свойства эллипса и гиперболы.

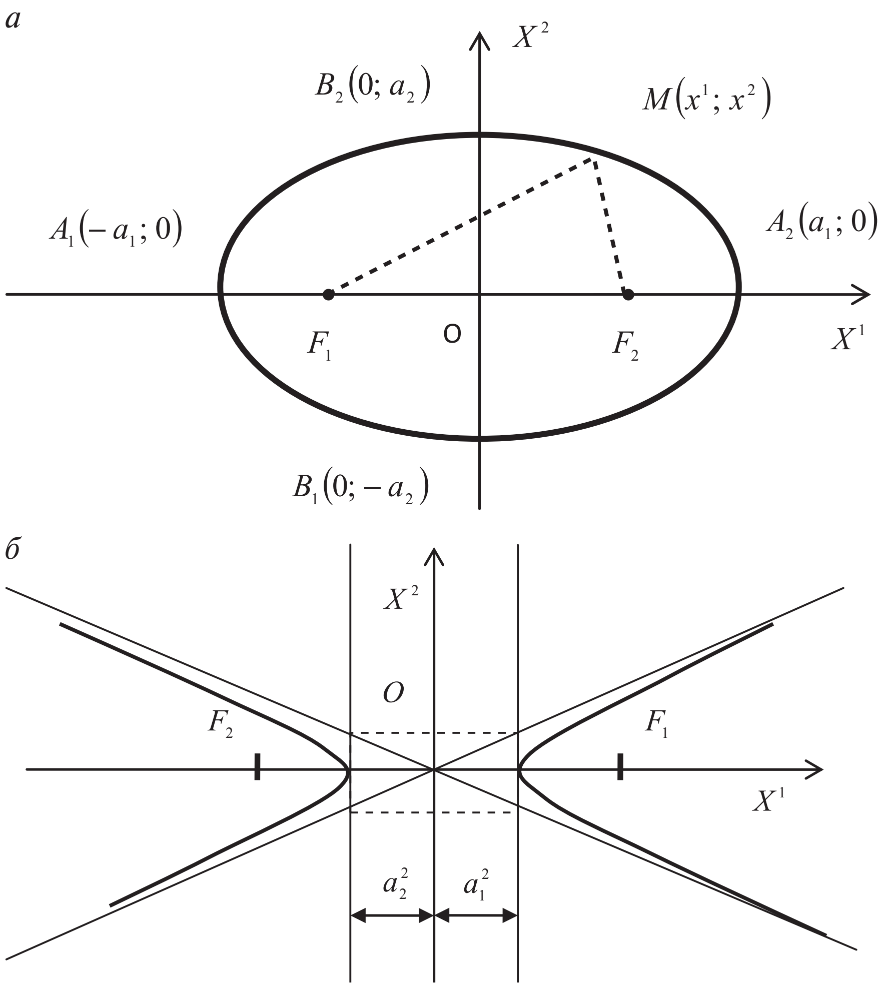


Рис. 5.3

Начнём с эллипса. Из уравнения (5.3.14) следует, что $|x_1| \le a_1$ и $|x_2| \le a_2$. Координатные оси OX_1 и OX_2 являются осями симметрии эллипса, а начало системы координат является центром симметрии эллипса. Оси симметрии называются ***главными осями эллипса***, а центр симметрии — ***центром эллипса***. Если выполняется, например, условие $a_1 > a_2$, то ось OX_1 называется ***большой*** осью эллипса, а ось OX_2 — ***малой*** осью эллипса. Точки пересечения главных осей и эллипса называются ***вершинами*** эллипса. Если $a_1 = a_2$, то эллипс вырождается в ***окружность***.

Пусть, для определённости, $a_1 > a_2$. Обозначим $c^2 = a_1^2 - a_2^2$, тогда точки $F_1(-c; 0)$ и $F_2(+c; 0)$, расположенные на оси OX_1 симметрично центру эллипса, называются ***фокусами эллипса***.

Имеют место следующие ***свойства эллипса***.

Свойство 5.3.1. *Сумма расстояний от любой точки $M(x_1; x_2)$ эллипса до его фокусов F_1 и F_2 есть постоянная величина, равная $2a_1$.*

Для всех точек эллипса выполняются следующие равенства:

$$\rho(M, F_2) = \frac{c}{a_1}\left|x_1 - \frac{a_1^2}{c}\right|, \rho(M, F_1) = \frac{c}{a_1}\left|x_1 + \frac{a_1^2}{c}\right|.$$

Прямые линии L_1 и L_2 с уравнениями

$$x_1 - \frac{a_1^2}{c} = 0, \; x_1 + \frac{a_1^2}{c} = 0, \tag{5.3.16}$$

соответственно, называются ***левой*** и ***правой директрисами*** эллипса.

Свойство 5.3.2. *Отношение расстояний от любой точки M эллипса до фокусов F_1 и F_2 и расстояний от точки M до соответствующих директрис L_1 и L_2 есть величина постоянная, то есть для $k = 1, 2$*

$$\frac{\rho(M, F_k)}{\rho(M, L_k)} = \alpha_k = \text{const}.$$

В отличие от эллипса гипербола является ***неограниченной кривой линией***, что непосредственно видно из уравнения (5.3.15). Оси координат и начало координат являются соответственно ***осями*** и ***центром симметрии*** гиперболы. Так же как и в случае эллипса, оси симметрии и центр симметрии гиперболы называются её ***главными осями*** и ***центром гиперболы*** соответственно. Точки, в которых одна из главных осей пересекается с гиперболой, называются ***вершинами*** гиперболы. На рис. 5.3, *б* это ось OX^1, которая в данном случае называется ***действительной осью*** гиперболы. Вторая ось — ось OX^2, не имеет с гиперболой общих точек и по этой причине называется ***мнимой осью*** гиперболы. Если ввести обозначение $c^2 = a_1^2 + a_2^2$, то точки $F_1(-c; 0)$ и $F_2(+c; 0)$ называются ***фокусами гиперболы***.

Имеют место следующие ***свойства гиперболы***.

Свойство 5.3.3. *Абсолютная величина разности расстояний от любой точки $M(x_1; x_2)$ гиперболы до её фокусов F_1 и F_2 есть постоянная величина $2a_1$.*

Прямые линии с уравнениями

$$x_2' = \frac{a_2}{a_1} x_1 \tag{5.3.17}$$

или прямая линия с уравнением

$$x_2' = -\frac{a_2}{a_1} x_1 \tag{5.3.18}$$

называются ***асимптотами*** гиперболы.

Для точек гиперболы, как и для точек эллипса, справедливы следующие равенства:

$$\rho(M, F_2) = \frac{c}{a_1}\left|x_1 - \frac{a_1^2}{c}\right|; \rho(M, F_1) = \frac{c}{a_1}\left|x_1 + \frac{a_1^2}{c}\right|.$$

Прямые линии L_1 и L_2 с уравнениями

$$x_1 - \frac{a_1^2}{c} = 0,\ x_1 + \frac{a_1^2}{c} = 0 \tag{5.3.19}$$

называются соответственно левой и правой ***директрисами*** гиперболы.

Свойство 5.3.4. *Для любой точки M гиперболы отношение расстояний от точки M до фокусов F_1 и F_2, и расстояний от точки M до соответствующих директрис L_1 и L_2 есть величина постоянная, то есть для* $k = 1, 2$

$$\frac{\rho(M, F_k)}{\rho(M, L_k)} = \alpha_k = \text{const}.$$

Парабола. На плоскости R^2 существует единственная нецентральная ***кривая второго порядка — парабола*** с каноническим уравнением

$$x_1^2 = 2a_1^2 x_2. \tag{5.3.20}$$

Парабола изображена на рис. 5.4, из которого видно, что парабола является ***неограниченной кривой линией***, как и рассмотренная выше гипербола. Парабола имеет одну ось симметрии, а именно ось OX^1 и не имеет центра симметрии. Точка пересечения оси параболы с самой параболой называется вершиной ***параболы***. Точка $F\left(\frac{a_1^2}{2}, 0\right)$ называется ***фокусом*** параболы. Прямая линия L с уравнением

$$x_1 = -\frac{a_1^2}{2} \tag{5.3.21}$$

называется ***директрисой*** параболы.

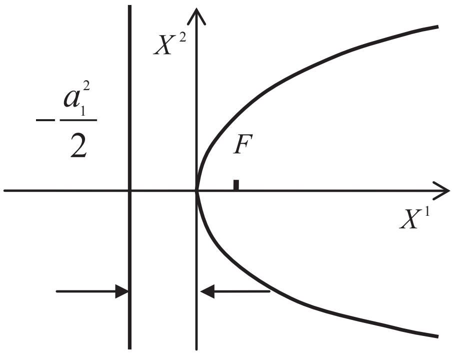


Рис. 5.4

Имеют место следующие ***свойства параболы***.

Свойство 5.3.5. *Расстояние от любой точки параболы до директрисы равно расстоянию от той же точки до фокуса параболы.*

Свойство 5.3.6. *Для любой точки M параболы отношение расстояний от точки M до фокуса F и расстояния от точки M до директрисы L есть величина постоянная, равная единице, то есть* $\dfrac{\rho(M,F)}{\rho(M,L)}=1.$

Эллипсоид и гиперболоиды. В пространстве R^3 существуют три ***невырожденные*** центральные поверхности второго порядка, определяемые каноническими уравнениями:

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\frac{x_3^2}{a_3^2}=1, \tag{5.3.22}$$

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}=1, \tag{5.3.23}$$

$$\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}=1. \tag{5.3.24}$$

Поверхности, задаваемые в пространстве R^3 уравнениями (5.3.22)–(5.3.23), называются соответственно ***эллипсоидом***, ***однополостным гиперболоидом*** и ***двуполостным гиперболоидом***.

Чтобы выяснить вид этих поверхностей, воспользуемся ***методом сечений***. Суть этого метода состоит в построении бесконечного семейства плоскостей, параллельных координатным плоскостям. Так, в случае уравнения (5.3.22) проводим семейство плоскостей, параллельных координатной плоскости X^1OX^2 с уравнениями $x_3 = c \cdot a_3$, где $-\infty < c < +\infty$. Тогда подстановка в уравнение (5.3.22) даёт уравнение

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1 - c^2,$$

которое для значений $|c| \le 1$ совпадает с уравнением эллипса (5.3.14). Следовательно, в сечениях получаем эллипсы. Сечения координатными плоскостями $x_1 = 0$ и $x_2 = 0$ также являются эллипсами. Общий вид этих поверхностей второго порядка приведён на рис. 5.5, *а*, *б*, *в*.

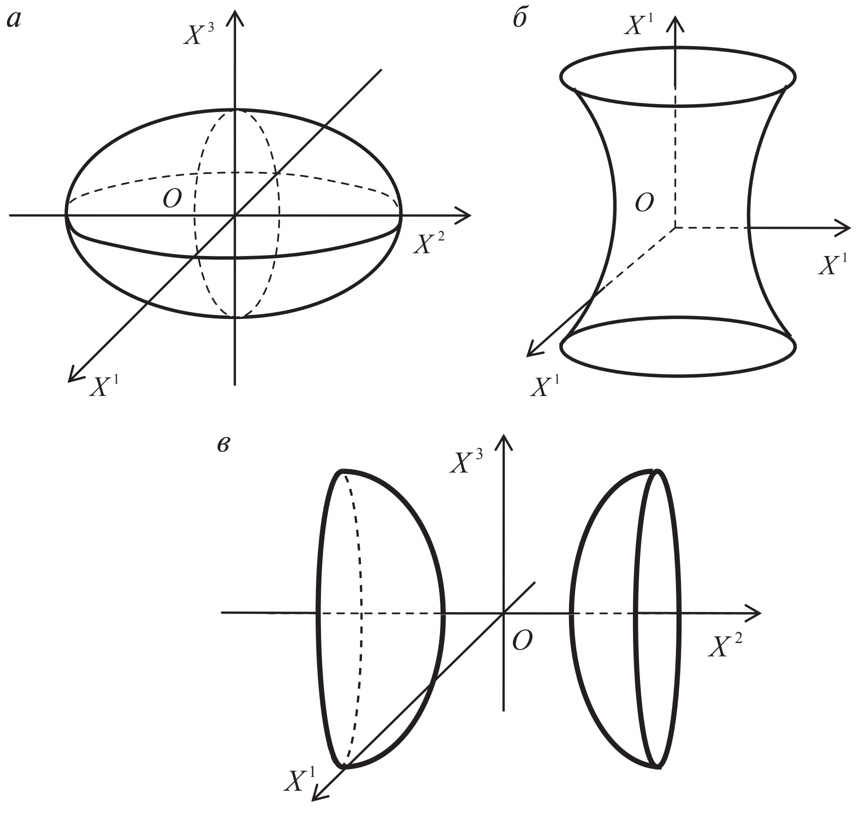


Рис. 5.5

Конус. В трёхмерном пространстве R^3 существует ***коническая поверхность*** с каноническим уравнением

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 0. \tag{5.3.25}$$

Вид поверхности, описываемой уравнением (5.3.25), приведён на рис. 5.6. Эта поверхность называется ***конусом***. При $a_1 = a_2$ конус называется ***прямым круговым конусом***.

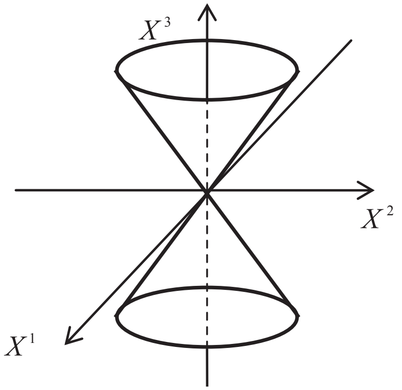


Рис. 5.6

Параболоиды. В пространстве R^3 существуют две невырожденные нецентральные поверхности с каноническими уравнениями:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 2x_3 \qquad (5.3.26)$$

— ***эллиптический параболоид***;

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 2x_3 \qquad (5.3.27)$$

— ***гиперболический параболоид***.

Исследуем поверхности, задаваемые уравнениями (5.3.26) и (5.3.27), методом сечений.

В случае эллиптического параболоида имеем: в сечении поверхности плоскостью $x_3 = C > 0$ — эллипс, а в сечениях поверхности координатными плоскостями $x_1 = 0$ и $x_2 = 0$ — параболы. Если $a_1 = a_2$, то эллиптический параболоид превращается в прямой круговой параболоид.

В случае гиперболического параболоида имеем: в сечении поверхности плоскостью $x_3 = C > 0$ — гиперболу; в сечении поверхности плоскостью $x_2 = 0$ — параболу $\frac{x_1^2}{a_1^2} = 2x^3 \Rightarrow x_1^2 = 2a_1^2 x_3$; в сечении поверхности плоскостью $x^3 = C < 0$ — гиперболу; в сечении плоскостью $x^3 = 0$ — пару прямых линий.

Общий вид эллиптического и гиперболического параболоидов приведён на рис. 5.7, *а* и 5.7, *б* соответственно.

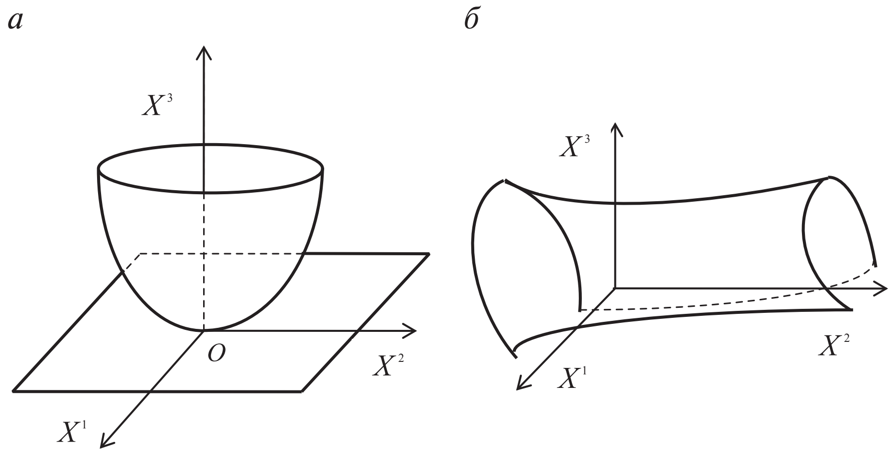


Рис. 5.7

Цилиндры. В пространстве R^3 общий вид ***вырожденной поверхности*** получается параллельным переносом вдоль оси OX_3 какой-либо кривой второго порядка на плоскости X_1OX_2.

При этом для эллипса, гиперболы и параболы получаются поверхности, носящие соответственно названия ***эллиптический***, ***гиперболический*** и ***параболический цилиндры*** (см. рис. 5.8, *а*). Пара параллельных, пересекающихся или слившихся прямых приводит, соответственно, к ***паре параллельных***, ***пересекающихся*** или ***слившихся плоскостей*** (см. рис. 5.8, *б*).

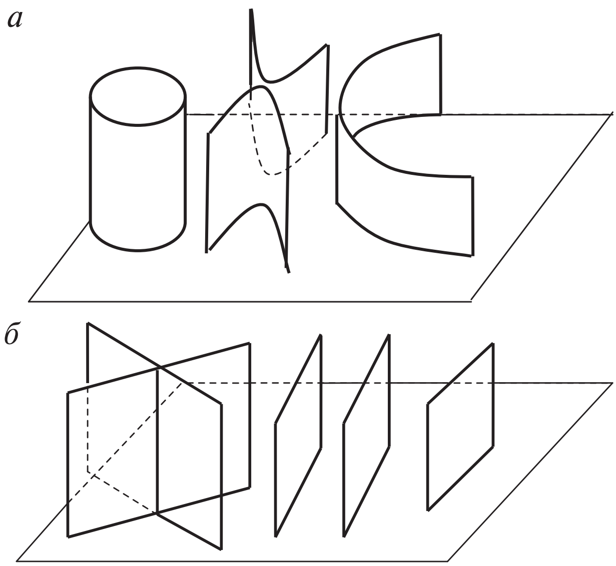


Рис. 5.8

На этом закончим краткий обзор поверхностей второго порядка, теория которых является частным случаем общей теории поверхностей. Эта теория является предметом дифференциальной геометрии.

Глава 6.
Некоторые вопросы алгебры, не вошедшие в основной курс

6.1. Кольцо многочленов от одного неизвестного

Определение многочлена. Из школьного курса математики известна задача решения уравнения второй степени вида

$$ax^2 + bx + c = 0, \tag{6.1.1}$$

где $a, b, c \in R$. Решить уравнение (6.1.1) — это значит найти такое значение неизвестного x, которое при подстановке в уравнение (***предикат***) (6.1.1) обращает его в числовое тождество (***в истинное высказывание***).

Пример 6.1.1. Найти множество истинности предиката

$$ax^2 + bx + c = 0.$$

Р е ш е н и е. Иными словами, нужно найти решение квадратного уравнения общего вида. Для этого рассмотрим тождественное преобразование квадратного трёхчлена в левой части указанного предиката:

$$ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) = a\left(x^2 + 2x\frac{b}{2a} + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} + \frac{c}{a}\right) =$$

$$= a\left(x + \frac{b}{2a}\right)^2 - a\left(\frac{b^2}{4a^2} - \frac{c}{a}\right).$$

Приравнивая последнее выражение к нулю, получаем известную формулу

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

которая позволяет найти те значения неизвестной x, которые обращают предикат в истинное высказывание. Из полученной формулы видно, что множество истинности T предиката

$$ax^2 + bx + c = 0$$

в общем случае состоит из двух элементов

$$T = \left\{ \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right\},$$

значения которых вычисляются через значения коэффициентов квадратного трёхчлена $ax^2 + bx + c$. Выражение $D = b^2 - 4ac$, стоящее под знаком квадратного корня, называется ***дискриминантом*** уравнения (6.1.1). Возможны три случая существования корней уравнения (6.1.1):

1) $D = b^2 - 4ac = 0$ — в этом случае множество истинности предиката состоит из одного действительного числа $T = \left\{ \frac{-b}{2a} \right\}$ (квадратное уравнение $ax^2 + bx + c = 0$ имеет один вещественный корень);

2) $D = b^2 - 4ac > 0$ — в этом случае множество истинности предиката состоит из двух вещественных чисел, которые вычисляются по приведенным выше формулам (квадратное уравнение $ax^2 + bx + c = 0$ имеет два вещественных корня);

3) $D = b^2 - 4ac < 0$ — в этом случае множество истинности предиката состоит из двух комплексно-сопряжённых чисел:

$$T = \left\{ \frac{-b - i\sqrt{4ac - b^2}}{2a}, \frac{-b + i\sqrt{4ac - b^2}}{2a} \right\}$$

(уравнение $ax^2 + bx + c = 0$ имеет комплексно-сопряжённые корни). ⊗

В общем случае можно поставить задачу решения ***уравнения n-й степени относительно одного неизвестного***

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0, \tag{6.1.2}$$

коэффициенты $a_0, a_1, \ldots, a_{n-1}, a_n$ которого будем считать ***произвольными комплексными числами***, причём старший коэффициент $a_n \neq (0, 0) \equiv 0$. Решить уравнение (6.1.2) — это значит найти такие значения неизвестного x, которые, будучи подставлены в уравнение (6.1.2), обращают его в числовое тождество. Задачу решения уравнения (6.1.2) заменяют более общей задачей ***изучения левой части этого уравнения***.

Определение 6.1.1. ***Многочленом,*** *или* ***полиномом степени n от одной буквы x (или от одного неизвестного x),*** *называется формальное выражение вида*

$$F(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0, \tag{6.1.3}$$

то есть формальная алгебраическая сумма целых неотрицательных степеней буквы (неизвестного) x, взятых с некоторыми, вообще говоря, комплексными коэффициентами $a_0, a_1, \ldots, a_n, a_{n-1}$. •

Обозначают многочлены различными буквами латинского и греческого алфавитов как большими $F(x), G(x), Q(x), \ldots, \Phi(x), \Omega(x), \ldots$, так и малыми $f(x), g(x), q(x), \ldots, \varphi(x), \omega(x), \ldots$.

Степенью многочлена (6.1.3) называется наивысшая степень n неизвестного x, при которой коэффициент $a_n \neq 0$. Многочлен ***нулевой степени*** — это многочлен, состоящий из одного, не равного нулю комплексного числа. Число нуль — это тоже многочлен, ***степень которого не определена***.

Степень многочлена n, если это необходимо, обозначается нижним индексом, например $F_n(x)$, или символом $n \equiv \deg F(x)$. Наряду с записью многочленов в форме (6.1.3) часто применяется форма записи по возрастающим степеням x, то есть

$$F_n(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + a_n x^n.$$

Равенство, сумма и произведение многочленов. Многочлены можно сравнивать и производить над ними действия сложения и умножения.

Определение 6.1.2. *Два многочлена* $F_n(x)$ *и* $Q_n(x)$ *считаются* ***равными*** *и пишут*

$$F_n(x) = Q_n(x)$$

в том и только в том случае, если равны их коэффициенты при одинаковых степенях неизвестного x. •

Никакой многочлен, хотя бы один коэффициент которого отличен от нуля, не может быть равным нулю. Поэтому знак равенства в записи уравнения n-й степени не имеет отношения к равенству многочленов. Смысл записи выражения (6.1.2) выяснится ниже.

В математическом анализе равенство многочленов $F_n(x)=Q_n(x)$ рассматривается как равенство двух функций, то есть

$$\left(\forall x\in R^1\right)\ F_n(x)=Q_n(x).$$

Если многочлены равны в смысле определения 6.1.2, то они равны и в смысле равенства функций. Обратное является следствием сформулированной ниже основной теоремы алгебры многочленов.

Введём две алгебраические операции над многочленами с комплексными (в общем случае) коэффициентами — ***сложение*** и ***умножение***.

Определение 6.1.3. *Пусть даны два многочлена*

$$F_n(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n,\ a_n\neq 0,$$

$$Q_m(x)=b_0+b_1x+\ldots+b_{m-1}x^{m-1}+b_mx^m,\ b_m\neq 0.$$

*Для определённости положим $n>m$. **Суммой** данных многочленов называется многочлен*

$$S_n(x)\equiv F_n(x)+Q_m(x)=s_0+s_1x+\ldots+s_mx^m+\ldots+s_{n-1}x^{n-1}+s_nx^n,$$

коэффициенты которого равны сумме коэффициентов при одинаковых степенях неизвестного x:

$$\left(\forall k=1,2,\ldots\right)\ s_k=a_k+b_k.$$

Причём, если $n>m$, полагают $b_{m+1}=b_{m+2}=\ldots=b_n=0$. •

Отметим, что степень суммы двух многочленов при $n>m$ равна n, а при $n=m$ может оказаться меньше n, например, при $b_n=-a_n$.

Определение 6.1.4. ***Произведением*** *многочленов*

$$F_n(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n,\ a_n\neq 0,$$

$$Q_m(x)=b_0+b_1x+\ldots+b_{m-1}x^{m-1}+b_mx^m,\ b_m\neq 0$$

называется многочлен

$$T_p(x)\equiv F_n(x)\cdot Q_m(x)=d_0+d_1x+\ldots+d_{p-1}x^{p-1}+d_px^p,$$

коэффициенты которого находятся по формуле

$$d_k=\sum_{i+j=k}a_ib_j,\ \left(k=0,1,\ldots,p-1,\ p\right).\ \bullet \qquad (6.1.4)$$

Из определения 6.1.4 следует, что коэффициент произведения двух многочленов с индексом $k = 0, 1, \ldots, p-1, p$ равен сумме всевозможных попарных произведений коэффициентов многочленов $F_n(x)$ и $Q_m(x)$, сумма индексов которых равна k, а именно:

$$d_0 = a_0 b_0,\ d_1 = a_0 b_1 + a_1 b_0,\ d_2 = a_0 b_2 + a_1 b_1 + a_2 b_0,\ \ldots,\ d_p = a_n b_m.$$

Учитывая, что $a_n \neq 0$ и $b_m \neq 0$, из последнего равенства имеем

$$d_p \equiv d_{n+m} = a_n b_m \neq 0.$$

Следовательно, *степень произведения двух многочленов равна сумме степеней этих многочленов*

$$\deg\left[F(x)\cdot Q(x)\right] = \deg F(x) + \deg Q(x)$$

и можно написать ($p = n + m$)

$$T_{n+m}(x) \equiv F_n(x)\cdot Q_m(x) = d_0 + d_1 x + \ldots + d_{n+m-1}x^{n+m-1} + d_{n+m}x^{n+m}.$$

Пришли к следующему утверждению.

Лемма 6.1.1. *Пусть $F_n(x) \neq 0$ и $Q_m(x) \neq 0$ — два многочлена. Тогда их произведение*

$$T_{n+m}(x) = F_n(x)\cdot Q_m(x) \neq 0.$$

Из сформулированной леммы следует, что ***множество всех многочленов является областью целостности***.

По определению полагают, что возведение многочлена в степень производится в соответствии с формулой

$$F^k(x) \overset{def}{=} \overbrace{F(x)\cdot F(x)\cdot \ldots \cdot F(x)}^{k\text{ раз}}.$$

Пример 6.1.2. Пусть даны два многочлена разной степени, например,

$$F_3(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3,\ Q_2(x) = b_0 + b_1 x + b_2 x^2.$$

Тогда их сумма и произведение есть, соответственно:

$$S_3(x) = F_3(x) + Q_2(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + a_3 x^3;$$

$$T_5(x) = F_3(x)\cdot Q_2(x) =$$

$$= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 +$$

$$+(a_1 b_2 + a_2 b_1 + a_3 b_0)x^3 + (a_2 b_2 + a_3 b_1)x^4 + a_3 b_2 x^5. \otimes$$

Таким образом, во множестве многочленов с комплексными коэффициентами введены две бинарные алгебраические операции — ***сложение*** и ***умножение***. Свойства этих операций устанавливаются следующей теоремой.

Теорема 6.1.1. *Множество всех многочленов с комплексными коэффициентами является коммутативным и ассоциативным кольцом с единицей.*

Доказательство. Для доказательства теоремы нужно проверить выполнение свойств операций сложения и умножения, которые вводятся посредством определений 6.1.3 и 6.1.4.

Рассмотрим сначала свойства операции сложения. Свойства коммутативности и ассоциативности операции сложения многочленов очевидны. Они следуют из коммутативности и ассоциативности операции сложения комплексных чисел, так как складываются коэффициенты при каждой степени неизвестного отдельно. Для операции сложения существует обратная операция вычитания, которая выполнима, так как роль нуля играет число нуль — многочлен с неопределённой степенью, а противоположным многочленом для данного многочлена

$$F(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n$$

является многочлен вида

$$-F(x)=-a_0-a_1x-\ldots-a_{n-1}x^{n-1}-a_nx^n$$

(проверка осуществляется непосредственно). Итак, множество всех многочленов является аддитивной абелевой группой по сложению.

Рассмотрим теперь свойства операции умножения. Коммутативность операции умножения следует из коммутативности операции умножения для чисел и того, что в произведение многочленов коэффициенты обоих сомножителей $F_n(x)$ и $Q_m(x)$ входят равноправным образом.

Докажем ассоциативность операции умножения. Пусть кроме многочленов $F_n(x)$ и $Q_m(x)$ степени n и m соответственно задан многочлен

$$T_l(x)=c_0+c_1x+\ldots+c_{l-1}x^{l-1}+c_lx^l,\, c_l\neq 0$$

степени l. Тогда, в соответствии с определением 6.1.4, коэффициентом при степени x^p, где $p=0,1,2,\ldots,n+m+l$, в произведении вида

$$\left[F_n(x)Q_m(x)\right]T_l(x)$$

будет служить число

$$\sum_{k+q=p}\left(\sum_{i+j=k} a_i b_j\right)c_q = \sum_{i+j+q=p} a_i b_j c_q,$$

а в произведении вида

$$F_n(x)\left[Q_m(x)T_l(x)\right]$$

равное ему число

$$\sum_{i+k=p} a_i\left(\sum_{j+q=k} b_j c_q\right) = \sum_{i+j+q=p} a_i b_j c_q.$$

Единицей при умножении служит число l, которое рассматривается как многочлен нулевой степени.

Дистрибутивные законы получаются из коммутативности операции умножения и равенства

$$\sum_{i+q=p}(a_i+b_i)c_q = \sum_{i+q=p} a_i c_q + \sum_{i+q=p} b_i c_q,$$

так как левая часть последнего равенства является коэффициентом при x^p в многочлене $\left[F_n(x)+Q_m(x)\right]T_l(x)$, правая часть — коэффициентом при той же степени неизвестного в многочлене $F_n(x)R_l(x)+Q_m(x)T_l(x)$. ••

Кольцо многочленов обозначают $K[x]$, где K — символ поля, над которым определён многочлен. Таким образом, теорема 6.1.1 утверждает: множество всех многочленов с комплексными коэффициентами является кольцом $C[x]$.

Делимость многочленов. Многочлен $F(x)$ имеет обратный многочлен $F^{-1}(x)$ в том и только в том случае, если $F(x)$ — многочлен нулевой степени. Действительно, если

$$F(x)=a\neq 0,$$

то обратный многочлен

$$F^{-1}(x)=a^{-1}.$$

Если же $\deg F(x)=n\geq 1$, то степень многочлена в левой части равенства

$$F(x)F^{-1}(x)=1$$

при условии, что $F^{-1}(x)$ существует, должна быть не меньше

$$n = \deg F(x),$$

но правая часть последнего равенства является многочленом нулевой степени.

Итак, ***в кольце многочленов*** $K[x]$ ***для операции умножения не существует обратной операции деления***. В кольце многочленов, однако, существует ***алгоритм деления с остатком***.

Теорема 6.1.2. *Для любых двух многочленов* $F(x)$ *и* $G(x)$ *существуют такие многочлены* $Q(x)$ *и* $R(x)$*, что*

$$F(x) = G(x)Q(x) + R(x), \tag{6.1.5}$$

где $\deg R(x) < \deg G(x)$*, или* $R(x) = 0$*. Представление (6.1.5) единственно.*

Д о к а з а т е л ь с т в о. Пусть $\deg F(x) = n$ и $\deg G(x) = m$. Запишем многочлены $F(x)$ и $G(x)$ по возрастающим степеням буквы, то есть в виде

$$F_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} + a_n x^n,$$

$$G_m(x) = b_0 + b_1 x + b_2 x^2 + \ldots + b_{m-1} x^{m-1} + b_m x^m.$$

Если $n < m$ или $F_n(x) = 0$, то положим в представлении (6.1.5)

$$Q(x) = 0,\ R(x) = F_n(x).$$

Тогда, очевидно, представление (6.1.5) выполняется.

Поэтому, предполагая, что $n \geq m$, положим

$$F_n(x) - \frac{a_n}{b_m} x^{n-m} G_m(x) = F_{n_1}^{(1)}(x). \tag{6.1.6}$$

Обозначим старший коэффициент многочлена $F_{n_1}^{(1)}(x)$ через $a_{n_1}^{(1)}$. Очевидно, что $n_1 < n$. Если $n_1 \geq m$, то положим

$$F_{n_1}^{(1)}(x) - \frac{a_{n_1}^{(1)}}{b_m} x^{n_1 - m} G_m(x) = F_{n_2}^{(2)}(x). \tag{6.1.7}$$

Старший коэффициент многочлена $F_{n_2}^{(2)}(x)$ обозначим $a_{n_2}^{(2)}$.

Если $n_2 \geq m$, то опять положим

$$F_{n_2}^{(2)}(x) - \frac{a_{n_2}^{(2)}}{b_m} x^{n_2 - m} G_m(x) = F_{n_3}^{(3)}(x), \tag{6.1.8}$$

и так далее...

Степени $n_1, n_2, n_3, \ldots$ многочленов $F_{n_1}^{(1)}, F_{n_2}^{(2)}, F_{n_3}^{(3)}, \ldots$, очевидно, убывают и после конечного числа шагов получим:

$$F_{n_{k-1}}^{(k-1)}(x) - \frac{a_{n_{k-1}}^{(k-1)}}{b_m} x^{n_{k-1} - m} G_m(x) = F_{n_k}^{(k)}(x), \tag{6.1.9}$$

где или $F_{n_k}^{(k)}(x) = 0$, или $n_k < m$. После этого процесс прекращается.

Складывая равенства (6.1.6)–(6.1.9), получаем

$$F_n(x) = \left(\frac{a_n}{b_m} x^{n-m} + \frac{a_{n_1}^{(1)}}{b_m} x^{n_1 - m} + \ldots + \frac{a_{n_{k-1}}^{(k-1)}}{b_m} x^{n_{k-1} - m} \right) G_m(x) + F_{n_k}^{(k)}(x). \tag{6.1.10}$$

Обозначая сумму в круглых скобках Q_{n-m}, а $F_{n_k}^{(k)}(x) = R_k(x)$, получаем представление (6.1.5), причём либо $R_k(x) = 0$, либо степень $n_k < m$.

Докажем единственность представления (6.1.5).

Пусть $(\exists Q'(x) \wedge R'(x))$:

$$F(x) = G(x)Q'(x) + R'(x), \tag{6.1.11}$$

где или $R'(x) = 0$, или $\deg R'(x) < \deg G(x)$.

Из представлений (6.1.5) и (6.1.11) имеем:

$$G(x)\left[Q(x) - Q'(x)\right] = R'(x) - R(x).$$

Степень многочлена в левой части последнего равенства не меньше степени $G(x)$, а степень многочлена в правой части или нулевая, или меньше степени $G(x)$. Поэтому последнее равенство справедливо лишь при условиях

$$Q(x) = Q'(x),\ R(x) = R'(x). \bullet\bullet$$

Многочлен $Q(x)$ в формуле (6.1.5) называется ***частным*** от деления многочлена $F(x)$ на многочлен $G(x)$, который называют ***делителем многочлена*** $F(x)$, а многочлен $R(x)$ называется ***остатком*** от этого деления. Если $R(x) = 0$, то говорят, что многочлен $F(x)$ ***делится***

на многочлен $G(x)$. Выясним, когда многочлен $F(x)$ делится на многочлен $G(x)$.

Теорема 6.1.3. *Многочлен* $F(x)$ *делится на многочлен* $G(x)$ *в том и только в том случае, если существует такой многочлен* $\Phi(x)$, *что*

$$F(x)=G(x)\Phi(x). \tag{6.1.12}$$

Д о к а з а т е л ь с т в о. Действительно, если $F(x)$ делится на $G(x)$, то в качестве $\Phi(x)$ следует взять частное от деления $F(x)$ на $G(x)$. Обратно, пусть многочлен, удовлетворяющий равенству (6.1.12), существует. Тогда из доказанной в теореме 6.1.2 единственности многочленов $Q(x)$ и $R(x)$ в представлении (6.1.5) и условия того, что степень $R(x)$ меньше степени $\Phi(x)$, следует, что частное от деления $F(x)$ на $G(x)$ равно $\Phi(x)$, а остаток $R(x)=0$. ••

Следствие из теоремы 6.1.3. *Если многочлен* $F(x)$ *и его делитель* $G(x)$ *имеют рациональные или действительные коэффициенты, то и частное* $\Phi(x)$ *также будет иметь рациональные или действительные коэффициенты.*

Пример 6.1.3. Выполнить деление с остатком многочлена

$$F(x)=2x^4-3x^3+4x^2-5x+6$$

на многочлен $G(x)=x^2-3x+1$.

Р е ш е н и е. Алгоритм деления (6.1.6) — (6.1.9) многочлена $F(x)$ на многочлен $G(x)$ реализуем в форме «***деления уголком***»:

$$\begin{array}{r|l}
2x^4-3x^3+4x^2-5x+6 & x^2-3x+1 \\ \cline{2-2}
-\left(2x^4-6x^3+2x^2\right) & 2x^2+3x+11 \\ \cline{1-1}
3x^3+2x^2-5x+6 & \\
-\left(3x^3-9x^2+3x\right) & \\ \cline{1-1}
11x^2-8x+6 & \\
-\left(11x^2-33x+11\right) & \\ \cline{1-1}
25x-5 &
\end{array}$$

Итак, частное $Q(x)=2x^2+3x+11$, остаток $R(x)=25x-5$. Поэтому имеет место представление следующего вида:

$$2x^4-3x^3+4x^2-5x+6=\left(2x^2+3x+11\right)\left(x^2-3x+1\right)+25x-5,$$

которое можно проверить непосредственным умножением. ⊗

Определение 6.1.5. *Пусть $F(x)\neq 0$ и $G(x)\neq 0$ — два многочлена. Многочлен $D(x)\neq 0$ называется* ***наибольшим общим делителем*** (*НОД*) *этих многочленов, если он является их общим делителем и сам делится на любой другой общий делитель этих многочленов.* •

НОД многочленов $F(x)$ и $G(x)$ обозначается

$$D(x)\equiv\left(F(x),G(x)\right).$$

Сформулируем и докажем теорему, дающую конструктивный алгоритм нахождения НОД для любых двух многочленов.

Теорема 6.1.4 (алгоритм Евклида). *Для любых двух многочленов $F(x)\neq 0$ и $G(x)\neq 0$ существует наибольший общий делитель*

$$D(x)=\left(F(x),G(x)\right).$$

Доказательство. Сначала сформулируем ***алгоритм Евклида*** для нахождения $\left(F(x),G(x)\right)$, а потом докажем, что полученный в процессе реализации этого алгоритма многочлен является наибольшим общим делителем двух данных многочленов.

Сначала делим многочлен $F(x)$ на многочлен $G(x)$ и получаем в общем случае некоторый остаток $R_1(x)$. Далее делим $G(x)$ на $R_1(x)$ и получаем остаток $R_2(x)$, делим $R_1(x)$ на $R_2(x)$ и получаем остаток $R_3(x)$, и так далее. В результате таких последовательных делений мы придём к остатку $R_k(x)$, на который делится предыдущий остаток $R_{k-1}(x)$. Этот остаток и будет наибольшим общим делителем данных многочленов.

Для доказательства выпишем последовательно цепочку делений:

$$F(x)=G(x)Q_1(x)+R_1(x),$$

$$G(x)=R_1(x)Q_2(x)+R_2(x),$$

$$R_1(x)=R_2(x)Q_3(x)+R_3(x),$$

$$R_2(x) = R_3(x)Q_4(x) + R_4(x),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \tag{6.1.13}$$

$$R_{k-3}(x) = R_{k-2}(x)Q_{k-1}(x) + R_{k-1}(x),$$

$$R_{k-2}(x) = R_{k-1}(x)Q_k(x) + R_k(x),$$

$$R_{k-1}(x) = R_k(x)Q_{k+1}(x).$$

Последнее равенство показывает, что $R_k(x)$ является делителем для $R_{k-1}(x)$. Поэтому оба слагаемых в правой части предпоследнего равенства делятся на $R_k(x)$ и, следовательно, на $R_k(x)$ делится и $R_{k-2}(x)$. Поднимаясь по цепочке делений вверх, получим, что $R_k(x)$ является делителем и для $R_{k-3}(x)$, …, $R_2(x)$, $R_1(x)$. Из второго равенства цепочки видим, что $R_k(x)$ является делителем и для $G(x)$ и, следовательно, на основании первого равенства — для $F(x)$. Итак, $R_k(x)$ является общим делителем для $F(x)$ и $G(x)$.

Пусть $\Phi(x)$ — произвольный общий делитель $F(x)$ и $G(x)$. Так как левая часть и первое слагаемое правой части первого равенства из соотношений (6.1.13) делятся на многочлен $\Phi(x)$, то на многочлен $\Phi(x)$ делится и $R_1(x)$. Спускаясь по цепочке (6.1.13) вниз, получаем, что многочлены

$$R_2(x), R_3(x), \ldots, R_{k-2}(x), R_{k-1}(x), R_k(x)$$

также делятся на многочлен $\Phi(x)$. Таким образом, в силу произвольности многочлена $\Phi(x)$ получаем, что $R_k(x)$ действительно является наибольшим общим делителем многочленов $F(x)$ и $G(x)$, то есть

$$R_k(x) = (F(x), G(x)). \bullet\bullet$$

Следствие из теоремы 6.1.4. *Если $F(x)$ и $G(x)$ — многочлены с рациональными или действительными коэффициентами, то*

$$D(x) = (F(x), G(x))$$

также будет иметь, соответственно, рациональные или действительные коэффициенты.

У этих многочленов могут быть, однако, и делители, не все коэффициенты которых являются рациональными или действительными числами.

Нетрудно показать, что если

$$D(x)=\big(F(x),G(x)\big),$$

то в качестве НОД этих многочленов можно взять многочлен $c\cdot D(x)$, где $c\in R^1$. Поэтому ***наибольший общий делитель двух многочленов определяется неоднозначно, с точностью до множителя нулевой степени***. Вследствие неоднозначности нахождения наибольшего общего делителя принято соглашение, в соответствии с которым ***старший коэффициент*** $D(x)$ ***всегда полагают равным единице***.

Пример 6.1.4. Найти $D(x)=\big(F(x),G(x)\big)$ для многочленов

$$F(x)=x^4+x^3-3x^2-4x-1,\ G(x)=x^3+x^2-x-1.$$

Р е ш е н и е. Производя деление уголком, последовательно находим:

1) $F(x)=G(x)Q_1(x)+R_1(x)=\left(x^3+x^2-x-1\right)x+\left(-2x^2-3x-1\right);$

2) $G(x)=R_1(x)Q_2(x)+R_2(x)=$

$$=\left(-2x^2-3x-1\right)\left(-\frac{1}{2}x+\frac{1}{4}\right)+\left(-\frac{3}{4}x-\frac{3}{4}\right);$$

3) $R_1(x)=R_2(x)Q_3(x)+R_3(x)=\left(-\frac{3}{4}x-\frac{3}{4}\right)\left(\frac{8}{3}x+\frac{4}{3}\right)+0.$

Откуда имеем:

$$D(x)=\big(F(x),G(x)\big)=-\frac{3}{4}(x+1).$$

По неоднозначности определения НОД можем взять

$$D(x)=\big(F(x),G(x)\big)=x+1.\ \otimes$$

Определение 6.1.6. *Два многочлена* $F(x)$ *и* $G(x)$ *называются* ***взаимно простыми****, если они имеют в качестве общих делителей только многочлены нулевой степени.* •

Из определения 6.1.6 и принятого соглашения о старшем коэффициенте наибольшего общего делителя двух многочленов непосредственно следует лемма.

Лемма 6.1.2. *Многочлены $F(x)$ и $G(x)$ взаимно просты в том и только в том случае, если $D(x)=(F(x),G(x))=1$.*

Корни многочленов. В начале параграфа мы уже касались решения уравнений второй степени. Сейчас мы рассмотрим вопрос о решении уравнений произвольной конечной степени. Итак, пусть

$$F_n(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n \tag{6.1.14}$$

— некоторый многочлен степени n. Будем предполагать, что $F_n(x)\in C[x]$ и коэффициент при старшей степени неизвестного $a_n\neq 0$.

Определение 6.1.7. *Пусть $c\in C$ — некоторое число. Тогда число*

$$F_n(c)=a_0+a_1c+\ldots+a_{n-1}c^{n-1}+a_nc^n, \tag{6.1.15}$$

полученное заменой неизвестного x числом c и последующим выполнением всех операций, обозначенных в сумме (6.1.14), называется ***значением многочлена*** $F_n(x)$ ***при*** $x=c$. •

Пусть $F_n(x)$ и $G_m(x)$ — два многочлена, причём для определённости предположим, что степень $n\geq m$. Очевидно, что если степени многочленов равны $(n=m)$ и $F_n(x)=G_n(x)$ в смысле определения 6.1.2, то и значения многочленов равны, то есть

$$F_n(c)=G_n(c).$$

Далее, обозначая

$$S_n(x)=F_n(x)+G_m(x),\ T_{n+m}(x)=F_n(x)G_m(x),$$

нетрудно видеть, что

$$S_n(c)=F_n(c)+G_m(c),\ T_{n+m}(c)=F_n(c)G_m(c),$$

откуда следует, что операции сложения и умножения многочленов, вообще говоря, можно рассматривать как соответствующие операции над функциями.

Определение 6.1.8. *Число $c\in C$ называется* ***корнем многочлена*** $F(x)$, *если при подстановке этого числа в многочлен $F(x)$ вместо неизвестного x многочлен принимает значение нуль, то есть $F(c)=0$.* •

Рассмотрим вопрос о делении многочлена $F(x)$ на ***линейный двучлен*** вида $x-c$.

Теорема 6.1.5 (Безу). *Значение $F(c)$ многочлена $F(x)$ равно остатку от деления многочлена на линейный двучлен $x-c$.*

Доказательство. Деля многочлен $F(x)$ на линейный двучлен $x-c$, получаем

$$F(x)=(x-c)Q(x)+R,$$

где $\deg R<\deg(x-c)\Rightarrow \deg R=0$, то есть остаток R — число. Подставляя в обе части $x=c$, получаем:

$$F(c)=(c-c)Q(c)+R=R. \bullet\bullet$$

Следствие из теоремы 6.1.5. *Число c является корнем многочлена $F(x)$ в том и только в том случае, если $F(x)$ делится на $x-c$, то есть существует многочлен $Q(x)$, такой, что*

$$F(x)=(x-c)Q(x).$$

Доказательство. Действительно, подставляя в это равенство $x=c$, получаем $F(x)=0$. ••

Введём понятие кратности корня многочлена.

Определение 6.1.9. *Число $c\in C$ называется* ***корнем многочлена*** $F_n(x)$ ***кратности*** $k\le n$, *если* $(\exists k\in N)$:

$$F_n(x)=(x-c)^k Q_{n-k}(x), \tag{6.1.16}$$

где многочлен $Q_{n-k}(x)$ на линейный двучлен $x-c$ не делится. •

Итак, из определения следует, что если число *c* является k-кратным корнем многочлена $F_n(x)$, то этот многочлен делится (без остатка) на $(x-c)^k$ и не делится на $(x-c)^{k+1}$.

В математическом анализе даётся определение производной функции одного действительного переменного. Для дальнейшего изложения потребуется определение производной многочлена $F_n(x)\in P[x]$, где P — в общем случае произвольное поле. Дадим это определение для поля C.

Определение 6.1.10. *Пусть дан многочлен*

$$F_n(x) = a_0 + a_1x + a_2x^2 + \ldots + a_{n-1}x^{n-1} + a_nx^n \qquad (6.1.17)$$

с произвольными комплексными коэффициентами. ***Первой производной многочлена*** *(6.1.17) называется многочлен степени $n-1$ вида*

$$F_n^{(1)}(x) = a_1 + 2a_2x + \ldots + (n-1)a_{n-1}x^{n-2} + na_nx^{n-1}, \qquad (6.1.18)$$

причём производная от многочлена нулевой степени и от нуля по определению считается равной нулю. •

Операция нахождения производной от многочлена называется ***дифференцированием многочлена***. Производная от первой производной называется ***второй производной*** и обозначается $F_n^{(2)}(x)$, и так далее. Производные обозначаются также штрихами, например $F_n'(x), F_n''(x), \ldots$. Нетрудно видеть, что

$$F_n^{(n)}(x) = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 \cdot a_n = n! \cdot a_n.$$

Поэтому очевидно, что $(n+1)$*-я производная от многочлена n-й степени $F_n(x)$ равна нулю:* $F_n^{(n+1)}(x) = 0$.

Лемма 6.1.3. *Дифференцирование многочленов производится по следующим правилам:*

$$\left(F(x) + G(x)\right)' = F'(x) + G'(x); \qquad (6.1.19)$$

$$\left(F(x)G(x)\right)' = F'(x)G(x) + F(x)G'(x); \qquad (6.1.20)$$

$$\left(F^k(x)\right)' = kF^{k-1}(x)F'(x). \qquad (6.1.21)$$

Доказательство. Доказательство леммы проводится непосредственной проверкой для многочленов произвольной степени с применением определения 6.1.10. ••

Теорема 6.1.6. *Если число c является корнем кратности k многочлена $F(x)$, то при $k > 1$ это число является корнем кратности $k-1$ первой производной многочлена $F'(x)$. Если же $k = 1$, то c не будет корнем производной многочлена $F'(x)$.*

Доказательство. Пусть c — корень кратности k многочлена $F(x)$, то есть

$$F(x) = (x-c)^k Q(x), \tag{6.1.22}$$

где $k > 1$ и $Q(x)$ на $x-c$ уже не делится. Продифференцируем обе части равенства (6.1.22), получим

$$F'(x) = (x-c)^k Q'(x) + k(x-c)^{k-1} Q(x) =$$
$$= (x-c)^{k-1}\left[(x-c)Q'(x) + kQ(x)\right]. \tag{6.1.23}$$

В формуле (6.1.23) первое слагаемое в квадратных скобках делится на $x-c$, а второе не делится. Поэтому их сумма также на двучлен $x-c$ не делится. Так как частное от деления $F'(x)$ на $(x-c)^{k-1}$ определено однозначно, получаем, что $(x-c)^{k-1}$ является наивысшей степенью линейного двучлена, на которую делится $F'(x)$. ••

Если теорему 6.1.6 применить несколько раз, то можем получить такое следствие.

Следствие из теоремы 6.1.6. *Корень кратности k многочлена $F(x)$ будет корнем кратности $k-p$ производной порядка p этого многочлена $F^{(p)}(x)$ $(k \geq p)$, причём он не будет служить корнем для производной порядка k этого многочлена.*

Основная теорема алгебры многочленов и следствия из неё. Заметим, что многочлен

$$F(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n ,$$

определённый над произвольным полем P, в общем случае может не иметь в этом поле корней. В частности, если все коэффициенты многочлена являются действительными числами, то, как показывает приведённый выше пример многочлена x^2+1, многочлен, тем не менее, может не иметь вещественных корней.

Определение 6.1.11. *Числовое поле P называется **алгебраически замкнутым**, если любой многочлен с коэффициентами из поля P имеет хотя бы один корень из этого поля.* •

Как следует из примера многочлена x^2+1, поле вещественных чисел R^1 не является ***алгебраически замкнутым***. Однако справедлива так называемая ***основная теорема алгебры многочленов,*** которая фактически утверждает, что ***поле комплексных чисел алгебраически замкнуто***.

Теорема 6.1.7 (основная теорема алгебры многочленов). *Каждый многочлен степени $n \geq 1$, определённый над полем C комплексных (в частности, над полем R^1 действительных) чисел, имеет хотя бы один, в общем случае комплексный, корень.*

Эта теорема является крупнейшим достижением алгебры. Все современные методы её доказательства используют топологические свойства множества R^1 действительных и множества C комплексных чисел. По этой причине доказательство основной теоремы обычно проводится в курсе теории функций комплексного переменного.

Приведём некоторые следствия из основной теоремы.

Следствие 1 из теоремы 6.1.7. *Каждый многочлен*

$$F_n(x) = a_0 + a_1 x + \ldots + a_n x^n \in C[x]$$

при $n \geq 1$ единственным образом представляется в виде разложения

$$F_n(x) = a_n (x - c_1)(x - c_2)\ldots(x - c_n), \qquad (6.1.24)$$

где c_k $(k = 1, 2, \ldots, n)$ — корни многочлена $F_n(x)$, среди которых могут быть равные.

Д о к а з а т е л ь с т в о. Первая часть теоремы доказывается непосредственным применением к многочлену $F_n(x)$ теоремы 6.1.7 и следствия из теоремы Безу. Действительно, по теореме 6.1.7 у многочлена $F_n(x)$ имеется хотя бы один, вообще говоря, комплексный корень. Поэтому по следствию из теоремы Безу этот многочлен представляется в виде

$$F_n(x) = (x - c_1) Q_{n-1}(x).$$

Так как коэффициенты многочлена $Q_{n-1}(x)$ снова являются комплексными или действительными числами, то имеем представление

$$F_n(x) = (x - c_1)(x - c_2) Q_{n-2}(x),$$

и так далее, пока не получим разложение (6.1.24)

$$F_n(x) = a_n (x - c_1)(x - c_2)\ldots(x - c_n).$$

Причём коэффициент a_n в правой части получается по следующей причине. Коэффициент при старшей степени x в многочлене

$$F_n(x) = a_0 + a_1 x + \ldots + a_n x^n$$

равен a_n. Предположим, что в правой части разложения (6.1.24) появился коэффициент $b \neq a_n$. Тогда при раскрытии скобок в представлении (6.1.24) старший член у $F_n(x)$ имел бы вид bx^n, но так как

$$F_n(x) = a_0 + a_1 x + \ldots + a_n x^n,$$

им является $a_n x^n$. Поэтому $b \equiv a_n$.

Докажем единственность разложения (6.1.24). Пусть имеется ещё разложение вида

$$F_n(x) = a_n(x - b_1)(x - b_2)\ldots(x - b_n). \tag{6.1.25}$$

В силу предыдущего получаем

$$(x - c_1)(x - c_2)\ldots(x - c_n) = (x - b_1)(x - b_2)\ldots(x - b_n). \tag{6.1.26}$$

Если $(\forall j = 1, 2, \ldots, n)\ c_i \neq b_j$, то, подставляя корень c_i в левую часть, получили бы слева нуль, а справа не равное нулю число. Поэтому каждый корень c_i равен некоторому из корней b_j, и, естественно, наоборот. Однако совпадение разложений (6.1.24) и (6.1.25) ещё не доказано, так как среди корней $c_i\ (i = 1, 2, \ldots, n)$ могут быть равные (кратные) корни. Предположим, что среди корней c_i имеется p корней, равных c_1, а среди корней b_j имеется q корней, равных c_1. Если мы покажем, что $p = q$, то единственность будет доказана.

По лемме 6.1.1 произведение не равных нулю многочленов не равно нулю. Пусть имеем два равных произведения многочленов

$$F(x)Q(x) = G(x)Q(x),$$

причём $Q(x) \neq 0$. Имеем $\left[F(x) - G(x)\right]Q(x) = 0$, откуда следует, что $F(x) - G(x) = 0$ или $F(x) = G(x)$. Таким образом, обе части равенства $F(x)Q(x) = G(x)Q(x)$ мы сократили на равный общий множитель. Применяя этот результат к обеим частям равенства (6.1.26) при $p > q$, то есть сокращая разложение (6.1.26) на общий множитель $(x - c_1)^q$, получаем, что левая часть разложения (6.1.26) ещё содержит множитель $x - c_1$, а правая уже его не содержит. Это, однако, как показано выше, приводит к противоречию. Таким образом, единственность разложения (6.1.24) доказана. ••

Следствие 2 из теоремы 6.1.7. *Каждый многочлен*

$$F_n(x) = a_0 + a_1x + \ldots + a_nx^n \in C[x]$$

при $n \geq 1$ *имеет ровно n корней, если каждый корень считать столько раз, какова его кратность.*

Доказательство. Объединяя в разложении (6.1.24) одинаковые сомножители, получаем

$$F_n(x) = a_n(x - c_1)^{k_1}(x - c_2)^{k_2}\ldots(x - c_p)^{k_p}, \qquad (6.1.27)$$

где $k_1 + k_2 + \ldots + k_p = n$, и среди корней $c_1, c_2, \ldots, c_p$ уже нет равных. Выясним смысл чисел k_i $(i = 1, 2, \ldots, p)$ из представления (6.1.27). Пусть кратность соответствующего корня равна s_i, тогда $k_i \leq s_i$. Полагая $k_i < s_i$, в силу определения кратности корня получаем: $F(x) = (x - c_i)^{k_i} Q(x)$. Заменяя в этом разложении многочлен $Q(x)$ его разложением на линейные множители, мы пришли бы к разложению, отличному от разложения (6.1.24), что приводит к противоречию с доказанной выше единственностью этого разложения. Поэтому $k_i = s_i$. ••

Следствие 2 справедливо и при степени многочлена $n = 0$, так как многочлен нулевой степени корней не имеет. Это следствие не применимо только к многочлену $F(x) = 0$.

Следствие 3 из теоремы 6.1.7. *Пусть* $F(x)$ *и* $G(x)$ *— некоторые многочлены, причём* $\deg F(x) \leq n$ *и* $\deg G(x) \leq n$*. Тогда если значения этих многочленов совпадают более чем при n различных значениях неизвестного x, то* $F(x) \equiv G(x)$.

Доказательство. Действительно, в указанных предположениях многочлен $F(x) - G(x)$ имеет более чем *n* корней, но его степень не превосходит *n*. Поэтому $F(x) = G(x)$. ••

Так как различных чисел $c \in C$ бесконечно много, получаем такое следствие.

Следствие 4 из теоремы 6.1.7. *Пусть* $F(x) \neq G(x)$ *— два многочлена. Тогда* $(\exists c \in C): F(c) \neq G(c)$.

Отметим, что такие числа можно найти не только в поле комплексных чисел, но и в поле действительных чисел, в поле рациональных чисел и даже во множестве целых чисел.

Из основной теоремы алгебры многочленов вытекает такая теорема.

Теорема 6.1.8 (Виета). *Пусть многочлен $F_n(x)$ имеет коэффициентом при старшей степени неизвестного 1, то есть*

$$F_n(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + x^n, \tag{6.1.28}$$

и пусть $c_1, c_2, \ldots, c_n$ — его корни. Тогда имеют место следующие ***формулы Виета****:*

$$a_0 = (-1)^n c_1 c_2 \ldots c_n;$$

$$a_1 = (-1)^{n-1} (c_1 c_2 \ldots c_{n-1} + c_1 c_2 \ldots c_{n-2} c_n + \ldots + c_2 c_3 \ldots c_n);$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;$$

$$a_{n-2} = c_1 c_2 + c_1 c_3 + \ldots + c_1 c_n + c_2 c_3 + \ldots + c_{n-1} c_n;$$

$$a_{n-1} = -(c_1 + c_2 + \ldots + c_n).$$

Доказательство. В указанных предположениях имеет место следующее разложение

$$F_n(x) = (x - c_1)(x - c_2)\ldots(x - c_n).$$

Перемножая в правой части этого разложения скобки, приводя подобные и приравнивая коэффициенты при одинаковых степенях неизвестного и сравнивая с представлением (6.1.28), получим формулы Виета. ••

Например, для многочлена третьей степени

$$F_3(x) = a_0 + a_1 x + a_2 x^2 + x^3$$

имеем:

$$a_0 = -c_1 c_2 c_3;\ a_1 = c_1 c_2 + c_1 c_3 + c_2 c_3;\ a_2 = -(c_1 + c_2 + c_3).$$

Следствие из теоремы 6.1.8. *Если p — целый корень многочлена*

$$F_n(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + x^n$$

с целыми коэффициентами, то p является делителем свободного члена a_0.

Пример 6.1.5. Найти многочлен четвёртой степени

$$F_4(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + x^4,$$

если известно, что он имеет простыми корнями числа 5 и –2 и двукратным корнем число 3.

Р е ш е н и е. По формулам Виета имеем:

$$a_3 = -(5-2+3+3) = -9;$$

$$a_2 = 5\cdot(-2)+5\cdot 3+5\cdot 3+(-2)\cdot 3+(-2)\cdot 3+3\cdot 3 = 17;$$

$$a_1 = -\left[5\cdot(-2)\cdot 3+5\cdot(-2)\cdot 3+5\cdot 3\cdot 3+(-2)\cdot 3\cdot 3\right] = 33;$$

$$a_0 = 5\cdot(-2)\cdot 3\cdot 3 = -90.$$

Искомый многочлен равен:

$$F_4(x) = -90+33x+17x^2-9x^3+x^4. \otimes$$

Если старший коэффициент a_n многочлена $F_n(x)$ отличен от 1, то для применения формул Виета следует сначала все коэффициенты многочлена разделить на a_n, что не повлияет на его корни.

Разложение многочлена с действительными коэффициентами на множители. Сформулируем и докажем важное следствие основной теоремы алгебры многочленов, относящееся к многочленам, коэффициенты которых являются действительными числами.

Теорема 6.1.9. *Пусть*

$$F_n(x) = a_0 + a_1x + \ldots + a_{n-1}x^{n-1} + a_nx^n \tag{6.1.29}$$

— некоторый многочлен с действительными коэффициентами. Тогда если комплексное число $z_0 = \alpha + \beta i$ $(\beta \neq 0)$ *является корнем многочлена* $F_n(x)$, *то число* $\overline{z_0} = \alpha - \beta i$ *также является корнем этого многочлена*.

Д о к а з а т е л ь с т в о. По условию теоремы

$$F_n(z_0) = a_0 + a_1z_0 + \ldots + a_{n-1}z_0^{\,n-1} + a_nz_0^{\,n} = 0.$$

Откуда имеем

$$\overline{F_n(z_0)} = \overline{a_0 + a_1z_0 + \ldots + a_{n-1}z_0^{\,n-1} + a_nz_0^{\,n}} = 0,$$

что приводит к равенству

$$\overline{F_n(z_0)} = a_0 + a_1\overline{z_0} + \ldots + a_{n-1}\overline{z_0^{\,n-1}} + a_n\overline{z_0^{\,n}} = 0 \Rightarrow$$

$$\Rightarrow \overline{F_n(z_0)} = a_0 + a_1\overline{z_0} + \ldots + a_{n-1}\left(\overline{z_0}\right)^{n-1} + a_n\left(\overline{z_0}\right)^n = 0.$$

Последнее и означает, что число $\overline{z_0} = \alpha - \beta i$ является корнем многочлена (6.1.29). ••

Из доказанной теоремы вытекает следствие.

Следствие из теоремы 6.1.9. *Если многочлен*

$$F_n(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n \in R[x]$$

имеет комплексный корень $z_0 = \alpha + \beta i \ (\beta \neq 0)$*, то он делится на квадратный трёхчлен с действительными коэффициентами, имеющий корни* z_0 *и* $\overline{z_0}$.

Доказательство. Действительно, так как $z_0 = \alpha + \beta i$ — корень многочлена $F_n(x)$, то по теореме $\overline{z_0} = \alpha - \beta i$ также является его корнем. Тогда по следствию 1 из основной теоремы получаем

$$F_n(x) = a_n(x - z_0)\left(x - \overline{z_0}\right)Q_{n-2}(x) =$$
$$= a_n\left[x^2 - \left(z_0 + \overline{z_0}\right)x + z_0 \cdot \overline{z_0}\right]Q_{n-2}(x).$$

Откуда и следует доказательство утверждения. ••

Утверждение справедливо и в том случае, когда число z_0 является корнем многочлена $F_n(x) \in R[x]$ кратности $1 < k \leq n$. В этом случае для многочлена имеем

$$F_n(x) = a_n\left[x^2 - \left(z_0 + \overline{z_0}\right)x + z_0 \cdot \overline{z_0}\right]^k Q_{n-2k}(x).$$

Теперь для многочленов с действительными коэффициентами получаем такой окончательный результат.

Теорема 6.1.10. *Всякий многочлен* $F_n(x) \in R[x]$ *единственным образом представляется в виде произведения своего старшего коэффициента* a_n *и нескольких многочленов с действительными коэффициентами, линейных вида* $x - c$*, соответствующих его действительным корням, и квадратных вида*

$$x^2 - \left(z_0 + \overline{z_0}\right)x + z_0 \cdot \overline{z_0},$$

соответствующих парам комплексно-сопряжённых корней z_s *и* $\overline{z_s}$*, то есть*

$$F_n(x) = a_n (x - c_1)^{\alpha_1} (x - c_2)^{\alpha_2} \ldots (x - c_k)^{\alpha_k} \cdot$$
$$\cdot \left(x^2 + p_1 x + q_1\right)^{\beta_1} \ldots \left(x^2 + p_s x + q_s\right)^{\beta_s}, \tag{6.1.30}$$

где $\sum_{i=1}^{k} \alpha_i + \sum_{j=1}^{s} \beta_j = n$.

Таким образом, зная все вещественные и комплексные корни многочлена с действительными коэффициентами $F_n(x)$, можно этот многочлен единственным образом разложить на множители, то есть представить в виде (6.1.30), где числа $a_n, c_1, c_2, \ldots, c_k, p_1, \ldots, p_s, q_1, \ldots, q_s$ являются вещественными.

Рациональные дроби. В этом пункте рассмотрим выражения, представленные формальным отношением двух многочленов.

Определение 6.1.12. ***Рациональной дробью*** *называется выражение вида* $\dfrac{F(x)}{G(x)}$, *где* $F(x), G(x) \in P[x]$, причём $G(x) \neq 0$. •

Подчеркнём, что в определении 6.1.12 многочлены рассматриваются как формальные суммы степеней неизвестного x с коэффициентами из поля P, то есть в соответствии с определением 6.1.1.

Определение 6.1.13. *Две рациональные дроби* $\dfrac{F_{(1)}(x)}{G_{(1)}(x)}$ *и* $\dfrac{F_{(2)}(x)}{G_{(2)}(x)}$ *считаются* ***равными*** и пишут

$$\frac{F_{(1)}(x)}{G_{(1)}(x)} = \frac{F_{(2)}(x)}{G_{(2)}(x)},$$

если выполняется условие

$$F_{(1)}(x) G_{(2)}(x) - F_{(2)}(x) G_{(1)}(x) = 0. \bullet$$

Равенство рациональных дробей в определении 6.1.13 является не тождественной записью, а ***соглашением***. Однако справедливо следующее утверждение, обосновывающее правомерность определения равенства по определению 6.1.13.

Лемма 6.1.4. *Равенство двух рациональных дробей является отношением эквивалентности.*

На доказательстве этого утверждения мы не останавливаемся.

Из определения 6.1.13 следует, что

$$\left(\forall\, W(x) \in P[x], W(x) \neq 0\right) \frac{F(x)}{G(x)} = \frac{F(x)W(x)}{G(x)W(x)},$$

то есть числитель и знаменатель рациональной дроби можно сократить или умножить на один и тот же множитель (многочлен, не равный нулю). Теперь очевидно, ***что две рациональные дроби равны, если одну из них можно получить из другой посредством умножения числителя и знаменателя на один и тот же множитель или путём сокращения***.

Операции сложения и умножения рациональных дробей вводятся аналогично соответствующим операциям над обычными дробями.

Определение 6.1.14. ***Суммой*** *и* ***произведением*** *двух рациональных дробей называются рациональные дроби, которые получаются, соответственно, по правилам:*

$$\frac{F_{(1)}(x)}{G_{(1)}(x)}+\frac{F_{(2)}(x)}{G_{(2)}(x)}\overset{def}{=}\frac{F_{(1)}(x)G_{(2)}(x)+F_{(2)}(x)G_{(1)}(x)}{G_{(1)}(x)G_{(2)}(x)}; \tag{6.1.31}$$

$$\frac{F_{(1)}(x)}{G_{(1)}(x)}\cdot\frac{F_{(2)}(x)}{G_{(2)}(x)}\overset{def}{=}\frac{F_{(1)}(x)F_{(2)}(x)}{G_{(1)}(x)G_{(2)}(x)}.\bullet \tag{6.1.32}$$

Очевидно, что введённые операции сложения и умножения рациональных дробей являются бинарными алгебраическими операциями.

Нетрудно показать, что операция сложения рациональных дробей коммутативна и ассоциативна и что в соответствии с определением операции сложения во множестве всех рациональных дробей существуют ***нулевой*** и ***противоположный элементы***, которые соответственно равны $\frac{0}{1}$ и $\frac{-F(x)}{G(x)}$.

Теперь можем заключить, что *множество всех рациональных дробей с операцией сложения дробей, введённой формулой (6.1.31), является абелевой группой*.

Аналогично устанавливается коммутативность и ассоциативность операции умножения рациональных дробей, а также дистрибутивность операции умножения относительно операции сложения. Элемент $\frac{1}{1}$ во множестве рациональных дробей является единичным, так как

$$\frac{F(x)}{G(x)}\cdot\frac{1}{1}=\frac{F(x)}{G(x)}.$$

Каждый отличный от нуля элемент во множестве рациональных дробей имеет обратный. Действительно, если

$$\frac{F(x)}{G(x)} \neq \frac{0}{1},$$

то $F(x) \neq 0$ и имеет смысл элемент $\dfrac{G(x)}{F(x)}$, откуда получаем

$$\frac{F(x)}{G(x)} \cdot \frac{G(x)}{F(x)} = \frac{1}{1}.$$

Из проведённого рассмотрения следует, что *множество всех рациональных дробей является **полем***, для которого принято обозначение $K(x)$. Кольцо $K[x]$ является подмножеством $K(x)$, что следует из отождествления:

$$(\forall F(x) \in K[x])\; F(x) \equiv \frac{F(x)}{1}.$$

Корректность этого отождествления можно проверить непосредственно, доказав его непротиворечивость с определениями отношения равенства, а также операций сложения и умножения в поле $K(x)$.

Определение 6.1.15. *Рациональная дробь $\dfrac{F(x)}{G(x)}$ называется **несократимой**, если многочлены $F(x)$ и $G(x)$ взаимно просты.* •

От дроби $\dfrac{F(x)}{G(x)}$ можно перейти к несократимой дроби, найдя НОД числителя и знаменателя и сократив дробь на него. Далее можно вынести коэффициент при старшей степени в знаменателе и отнести его к числителю, тогда получим так называемую ***нормализованную форму записи*** рациональной дроби.

Определение 6.1.16. *Рациональная дробь $\dfrac{F(x)}{G(x)}$ называется **правильной**, если степень её числителя меньше степени знаменателя.* •

Очевидно, что если дробь является правильной в некоторой записи, то она остаётся правильной и в несократимой записи, так как при сокращении степени числителя и знаменателя уменьшаются на одно и то же число.

Лемма 6.1.5. *Любая рациональная дробь есть сумма многочлена и правильной рациональной дроби.*

Доказательство. Пусть $\dfrac{F(x)}{G(x)}$ — некоторая дробь. Деля числитель на знаменатель с остатком, получим

$$F(x)=G(x)Q(x)+R(x),$$

где $\deg R(x)<\deg G(x)$. Откуда

$$\frac{F(x)}{G(x)}=\frac{G(x)Q(x)}{G(x)}+\frac{R(x)}{G(x)}=Q(x)+\frac{R(x)}{G(x)}.$$

Здесь $Q(x)$ — многочлен, а $\dfrac{R(x)}{G(x)}$ — правильная рациональная дробь. ••

Лемма 6.1.6. *Сумма, разность и произведение правильных рациональных дробей — снова есть правильная рациональная дробь.*

Доказательство. Пусть дроби $\dfrac{F_{(1)}(x)}{G_{(1)}(x)}$ и $\dfrac{F_{(2)}(x)}{G_{(2)}(x)}$ — правильные рациональные дроби. Для суммы имеем:

$$\frac{F_{(1)}(x)}{G_{(1)}(x)}+\frac{F_{(2)}(x)}{G_{(2)}(x)}=\frac{F_{(1)}(x)G_{(2)}(x)+F_{(2)}(x)G_{(1)}(x)}{G_{(1)}(x)G_{(2)}(x)}.$$

Степени обоих слагаемых в числителе меньше степени знаменателя, так как дроби $\dfrac{F_{(1)}(x)G_{(2)}(x)}{G_{(1)}(x)G_{(2)}(x)}$ и $\dfrac{F_{(2)}(x)G_{(1)}(x)}{G_{(1)}(x)G_{(2)}(x)}$ — правильные. Поэтому степень числителя меньше степени знаменателя. Для произведения имеем

$$\frac{F_{(1)}(x)}{G_{(1)}(x)}\cdot\frac{F_{(2)}(x)}{G_{(2)}(x)}=\frac{F_{(1)}(x)F_{(2)}(x)}{G_{(1)}(x)G_{(2)}(x)}.$$

Очевидно, что

$$\deg\left(F_{(1)}(x)F_{(2)}(x)\right)=\deg F_{(1)}(x)+\deg F_{(2)}(x)<$$
$$<\deg G_{(1)}(x)+\deg G_{(2)}(x)=\deg\left(G_{(1)}(x)G_{(2)}(x)\right). \bullet\bullet$$

Теорема 6.1.11. *Пусть* $\dfrac{F(x)}{G(x)}$ *— правильная рациональная дробь, знаменатель которой имеет вид*

$$G(x)=(x-a)^m\left(x^2+bx+c\right)^n,$$

где $b^2-4c<0$. *Тогда дробь* $\dfrac{F(x)}{G(x)}$ *представляется в виде суммы простейших рациональных дробей по следующей формуле:*

$$\frac{F(x)}{G(x)}=\frac{A_1}{x-a}+\frac{A_2}{(x-a)^2}+\dots+\frac{A_m}{(x-a)^m}+$$

$$+\frac{B_1x+C_1}{x^2+bx+c}+\frac{B_2x+C_2}{(x^2+bx+c)^2}+\dots+\frac{B_nx+C_n}{(x^2+bx+c)^n}, \qquad (6.1.33)$$

где $A_1, A_2, \dots, A_m, B_1, C_1, \dots, B_n, C_n$ *— неопределённые коэффициенты.*

Пример 6.1.6. Разложить рациональную дробь $\dfrac{F(x)}{G(x)}=\dfrac{2x^3-40x-8}{x^3+2x^2-8x}$ на простейшие дроби.

Р е ш е н и е. Так как $m=n=3$, делим числитель на знаменатель:

$$2x^3-40x-8=2\left(x^3+2x-8x\right)+\left(-4x^2-24x-8\right);$$

$$\frac{F(x)}{G(x)}=\frac{2x^3-40x-8}{x^3+2x^2-8x}=2-\frac{4x^2+24x+8}{x^3+2x^2-8x};$$

$$\frac{F(x)}{G(x)}=\frac{2x^3-40x-8}{x^3+2x^2-8x}=2-\frac{4x^2+24x+8}{x(x+4)(x-2)}.$$

Знаменатель дроби имеет простые корни $x_1=0, x_2=-4, x_3=2$, и дробь представляется в виде разложения

$$\frac{4x^2+24x+8}{x(x+4)(x-2)}=\frac{A}{x}+\frac{B}{x+4}+\frac{C}{x-2}.$$

Приводя к общему знаменателю и приравнивая коэффициенты при одинаковых степенях неизвестного в числителе, получаем систему линейных алгебраических уравнений

$$\begin{cases} A+B+C=4, \\ 2A-2B+4C=24, \\ -8A \qquad\quad =8, \end{cases}$$

имеющую решение $A=-1, B=-1, C=6$. Получаем:

$$\frac{4x^2+24x+8}{x(x+4)(x-2)}=2+\frac{1}{x}+\frac{1}{x+4}-\frac{6}{x-2}. \otimes$$

6.2. Общая теория определителей

Понятие определителя. Выше (параграф 2.10) была рассмотрена эвристическая теория определителей. В этом дополнении изложена общая теория ***определителей***, ориентированная на применение к исследованию совместности СЛАУ с квадратной матрицей (число неизвестных равно числу уравнений). Рассматривается случай квадратных матриц общего вида порядка n и при этом напоминаются некоторые понятия, введённые ранее.

Запишем СЛАУ с квадратной матрицей, имеющей n строк и n столбцов

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_n^1 x^n = b^1, \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_n^2 x^n = b^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_1^n x^1 + a_2^n x^2 + \ldots + a_n^n x^n = b^n. \end{cases} \qquad (6.2.1)$$

Определение 6.2.1. *Квадратная матрица размера $n \times n$, составленная из n^2 коэффициентов a_j^i, где $i, j = 1, 2, \ldots, n$, СЛАУ (6.2.1)*

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_n^n \end{pmatrix}, \qquad (6.2.2)$$

называется ***основной матрицей СЛАУ*** (6.2.1). *Число n столбцов (строк) матрицы A называется её* ***порядком****, числа a_j^i называются* ***элементами матрицы****. Элементы*

$$\{a_1^1, a_2^2, \ldots, a_n^n\}$$

образуют ***главную диагональ*** *матрицы A.* •

Напомним, что в теории квадратных матриц большое значение имеет величина

$$SpA \stackrel{def}{=} \sum_{i=1}^{n} a_i^i$$

— сумма элементов, стоящих на главной диагонали, которая называется ***след матрицы***.

Определим на множестве матриц вида (6.2.2) числовую функцию — ***определитель***. Для этого рассмотрим любое произведение n элементов матрицы (6.2.2), расположенных ***по одному в каждом столбце и в каждой строке***. Запишем это произведение следующим образом:

$$a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \ldots \cdot a_n^{\alpha_n}. \tag{6.2.3}$$

Здесь $\alpha_1, \alpha_2, \ldots, \alpha_n$ — номера строк, а $1, 2, \ldots, n$ — номера столбцов матрицы. Таким образом, произведение (6.2.3) ***упорядочено по номерам столбцов***.

По условию элементы произведений $a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \ldots \cdot a_n^{\alpha_n}$ расположены в разных строках матрицы (6.2.2), поэтому числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ различны и образуют ***перестановку*** чисел $1, 2, \ldots, n$, которую обозначают $P(n)$.

Лемма 6.2.1. *Число всех возможных перестановок из n чисел* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *находится по формуле*

$$P\{\alpha_1, \alpha_2, \ldots, \alpha_n\} = n! \equiv 1 \cdot 2 \cdot \ldots \cdot n.$$

Доказательство. 1) Для $n = 1$ имеем $P(1) = 1$.

2) Пусть $P(n-1) = (n-1)!$.

3) Все перестановки из n чисел разобьём на n классов, помещая в каждый класс только те перестановки, которые имеют на первом месте одно и то же число. Тогда в каждом классе будет $(n-1)!$ перестановок. Так как число классов равно n, то общее число перестановок из n чисел будет равно $n!$. ••

Определение 6.2.2. ***Беспорядком (инверсией)*** *во множестве индексов* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *называется такая их перестановка, в которой старший индекс предшествует младшему.* •

Число всех беспорядков обозначим $N\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Пример 6.2.1. Найти число беспорядков в перестановках:

$$P\{2,1,4,3\},\ P\{4,3,1,2\}.$$

Решение. 1) $N\{2,1,4,3\} = 2$; 2) $N\{4,3,1,2\} = 5$. $\otimes$

Если в перестановке индексов $P(n)$ число беспорядков $N\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ ***чётное***, поставим перед произведением (6.2.3) знак «плюс», а если число беспорядков $N\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ ***нечётное***, то поставим перед произведением (6.2.3) знак «минус».

Определение 6.2.3. ***Определителем матрицы*** (6.2.2) *(порядка n) называется отображение D множества квадратных* $(n \times n)$ *матриц вида* (6.2.2) *во множество действительных чисел, если его* ***значение*** $\det A$ *на каждой матрице вычисляется как алгебраическая сумма* $n!$ *всевозможных произведений вида* (6.2.3) *по правилу:*

$$\det A \overset{def}{=} \sum_{P\{\alpha_1, \alpha_2, \ldots, \alpha_n\}} (-1)^{N\{\alpha_1, \alpha_2, \ldots, \alpha_n\}} a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \ldots \cdot a_n^{\alpha_n}. \qquad (6.2.4)$$

Произведения вида $a_1^{\pm_1} \cdot a_2^{\pm_2} \cdot \ldots \cdot a_n^{\pm_n}$ *называются* ***членами определителя,*** *а элементы* a_j^i *матрицы (6.2.2) называются его* ***элементами.*** •

Напомним, что для определителя матрицы (6.2.2) используются различные обозначения:

$$D(A) = \det A = \det\left(a_j^i\right) = \det\left\|a_j^i\right\| = \det\left[a_j^i\right] =$$

$$= \det\left\|a_j^i\right\|_{j=1, 2, \ldots, n}^{i=1, 2, \ldots, n} = \begin{vmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_n^n \end{vmatrix}. \qquad (6.2.5)$$

Пример 6.2.2. Вычислить определитель $\begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix}$.

Р е ш е н и е. По формуле (6.2.4) имеем

$$\begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix} = a_1^1 \cdot a_2^2 - a_1^2 \cdot a_2^1. \otimes$$

Определение 6.2.4. *Если в матрице n-го порядка A вычеркнуть i-ю строку и j-й столбец* $(i, j = 1, 2, \ldots, n)$, *то определитель получившейся матрицы* $(n-1)$*-го порядка называется* ***минором элемента*** a_j^i ***матрицы A*** *и обозначается*

$$M_j^i = M_j^i\left(a_j^i\right). \bullet$$

Определим одно полезное для теории определителей понятие.

Определение 6.2.5. *Говорят, что отрезок, соединяющий элемент* a_j^i *матрицы* $A = \left(a_m^l\right)_{m=1, 2, \ldots, n}^{l=1, 2, \ldots, n}$ *с элементом* a_k^p *той же матрицы, имеет* ***положительный наклон,*** *если его правый конец расположен ниже левого, и* ***отрицательный наклон,*** *если его правый конец расположен выше левого.* •

Пример 6.2.3. Рассмотрим матрицу четвёртого порядка:

$$A = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 & a_4^1 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \\ a_1^4 & a_2^4 & a_3^4 & a_4^4 \end{pmatrix}.$$

Отрезок, соединяющий элемент a_3^1 с элементом a_4^3, очевидно имеет положительный наклон, а отрезок, соединяющий элемент a_2^1 с элементом a_1^2, имеет отрицательный наклон. $\otimes$

Лемма 6.2.2. *Число беспорядков* $P\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *в перестановке индексов члена* $a_1^{\pm_1} \cdot a_2^{\pm_2} \cdot \ldots \cdot a_n^{\pm_n}$ *определителя* $\det A$ *матрицы (6.2.2) равно числу отрезков отрицательного наклона, соединяющих элементы, входящие в данный член определителя.*

Доказательство. Пусть $a_i^{\pm_i}$ и $a_j^{\pm_j}$ — некоторые элементы определителя. Тогда наличие отрезка отрицательного наклона при $i < j$ означает, что $\alpha_i > \alpha_j$, то есть наличие беспорядка в расположении индексов α_i, α_j. ••

Таким образом, знак члена определителя определяется числом отрезков отрицательного наклона, соединяющих входящие в него элементы.

Определение 6.2.6. *Матрица*

$$A^T \overset{def}{=} \begin{pmatrix} a_1^1 & a_1^2 & \ldots & a_1^n \\ a_2^1 & a_2^2 & \ldots & a_2^n \\ \ldots & \ldots & \ldots & \ldots \\ a_n^1 & a_n^2 & \ldots & a_n^n \end{pmatrix}, \tag{6.2.6}$$

полученная из матрицы (6.2.2) путём замены строк на столбцы с теми же номерами, называется ***матрицей, транспонированной*** *по отношению к исходной.* •

Свойства определителей. Определители обладают рядом свойств, облегчающих их вычисление.

Свойство 1. *Определитель транспонированной матрицы по величине равен определителю исходной матрицы, то есть для любой квадратной матрицы A имеет место равенство* $\det A = \det A^T$.

Доказательство. Оба определителя состоят, очевидно, из одних и тех же членов. Поэтому достаточно показать, что одинаковые члены в обоих определителях имеют один знак. Но это очевидно, так как при транспонировании матрицы (для квадратной матрицы транспонирование равносильно повороту исходной матрицы на угол 180° вокруг главной диагонали) отрезки с отрицательным наклоном переходят снова в отрезки отрицательного наклона. ••

Благодаря свойству 1 все приведённые ниже утверждения справедливы как в отношении столбцов, так и в отношении строк определителя.

Свойство 2. *При перестановке двух столбцов в определителе его знак изменяется на противоположный.*

Доказательство. Предположим, что переставляются j-й и $(j+1)$-й столбцы определителя. Каждый член определителя имеет в своём составе как элемент из j-го столбца, так и элемент из $(j+1)$-го столбца. Отрезок, соединяющий эти элементы, при перестановке столбцов изменяет свой наклон на противоположный. Следовательно, каждый член определителя при перестановке указанных столбцов меняет свой знак.

Пусть теперь переставляются j-й и k-й столбцы определителя, причём между ними имеется m других столбцов. Тогда для перестановки j-го и k-го столбцов необходимо совершить $2m+1$ перестановок. Опять имеем смену знака члена определителя. ••

Следствие из свойства 2. *Определитель, имеющий два одинаковых столбца, равен нулю.*

Доказательство. Если переставить одинаковые столбцы, то согласно свойству 2 определитель изменит знак на противоположный. Но, с другой стороны, перестановка одинаковых столбцов не может изменить значение определителя. Следовательно, определитель действительно равен нулю. ••

Свойство 3. *Если все элементы j-го столбца определителя D представлены в виде линейной комбинации двух слагаемых*

$$a_j^i = \lambda b_j^i + \mu c_j^i \left(i = 1, 2, \ldots, n\right),$$

где λ и μ — фиксированные числа, то определитель представляется в виде

$$D = \lambda D_1 + \mu D_2, \tag{6.2.7}$$

причём, у каждого из определителей D_1 и D_2 все столбцы, кроме j-го, такие же, как и у определителя D, а j-й столбец у определителя D_1 состоит из чисел b_j^i, а у определителя D_2 — из чисел c_j^i.

Доказательство. Действительно, каждый член определителя представляется в виде:

$$a_1^{\alpha_1}a_2^{\alpha_2}\ldots a_j^{\alpha_j}\ldots a_n^{\alpha_n}=a_1^{\alpha_1}a_2^{\alpha_2}\ldots\left(\lambda b_j^{\alpha_j}+\mu c_j^{\alpha_j}\right)\ldots a_n^{\alpha_n}=$$
$$=\lambda a_1^{\alpha_1}a_2^{\alpha_2}\ldots b_j^{\alpha_j}\ldots a_n^{\alpha_n}+\mu\, a_1^{\alpha_1}a_2^{\alpha_2}\ldots c_j^{\alpha_j}\ldots a_n^{\alpha_n}.$$

Собирая вместе первые и, соответственно, вторые слагаемые, вынося в первой сумме за скобки множитель λ, а во второй сумме множитель μ, получаем представление (6.2.7). ••

Доказанное свойство непосредственно обобщается на случай, когда некоторый столбец (строка) определителя представлен в виде линейной комбинации произвольного конечного числа элементов вида

$$a_j^i=\alpha\cdot b_j^i+\beta\cdot c_j^i+\ldots+\lambda\cdot d_j^i.$$

Следствие из свойства 3. *Если некоторый столбец определителя состоит только из нулей, то такой определитель равен нулю.*

Доказательство. Так как 0 есть общий множитель элементов данного столбца, то согласно свойству 3 его можно вынести за знак определителя в качестве множителя. ••

Свойство 4. *Определитель не изменится, если к элементам одного из его столбцов прибавить соответствующие (по номерам) элементы любого другого его столбца, умноженные на фиксированное число:*

$$D\left(a_j^i+\lambda a_k^i\right)=D\left(a_j^i\right).$$

Доказательство. Прибавляя, например, к элементам j-го столбца умноженные на λ элементы k-го столбца, получаем:

$$D\left(a_j^i+\lambda a_k^i\right)=D_1\left(a_j^i\right)+\lambda D_2\left(a_k^i\right).$$

В определителе $D_2\left(a_k^i\right)$ столбец с номером j состоит из элементов a_k^i, то есть совпадает с k-м столбцом. Поэтому $D_2\left(a_k^i\right)=0$, откуда и следует доказательство свойства 4.••

Определение 6.2.7. *В формуле (6.2.4) соберём лишь те члены определителя, которые содержат элемент a_j^i, и вынесем элемент a_j^i за скобки.*

Сумма A^i_j, оставшаяся в скобках, называется ***алгебраическим дополнением элемента*** a^i_j *определителя.* ••

Свойство 5. *Сумма всех произведений элементов какого-либо столбца (строки) определителя D на соответствующие алгебраические дополнения равна значению определителя:*

$$D = a^1_j A^1_j + a^2_j A^2_j + \ldots + a^n_j A^n_j; \tag{6.2.8}$$

$$D = a^i_1 A^i_1 + a^i_2 A^i_2 + \ldots + a^i_n A^i_n. \tag{6.2.9}$$

Доказательство следует из того, что каждый член определителя содержит элемент j-го столбца (i-й строки). ••

Свойство 6. *Справедливо равенство*

$$A^i_j = (-1)^{i+j} \cdot M^i_j. \tag{6.2.10}$$

Доказательство. 1) Рассмотрим сначала случай $i = j = 1$. Тогда, используя определение определителя (6.2.4), получаем:

$$\det A = \sum_{P\{\alpha_2,\ldots,\alpha_n\}} (-1)^{N\{1,\alpha_2,\ldots,\alpha_n\}} a^1_1 a^{\alpha_2}_2 \ldots a^{\alpha_n}_n +$$

$$+ \sum_{\substack{P\{\alpha_1,\ldots,\alpha_n\} \\ \alpha_1 \neq 1}} (-1)^{N\{\alpha_1,\alpha_2,\ldots,\alpha_n\}} a^{\alpha_1}_1 a^{\alpha_2}_2 \ldots a^{\alpha_n}_n =$$

$$= a^1_1 \sum_{P\{\alpha_2,\ldots,\alpha_n\}} (-1)^{N\{\alpha_2,\ldots,\alpha_n\}} a^{\alpha_2}_2 \ldots a^{\alpha_n}_n + \sum_{\substack{P\{\alpha_1,\ldots,\alpha_n\} \\ \alpha_1 \neq 1}} (-1)^{N\{\alpha_1,\alpha_2,\ldots,\alpha_n\}} a^{\alpha_1}_1 a^{\alpha_2}_2 \ldots a^{\alpha_n}_n .$$

Снова используя определение (6.2.4), видим, что

$$A^1_1 = \sum_{P\{\alpha_2,\ldots,\alpha_n\}} (-1)^{N\{\alpha_2,\alpha_3,\ldots,\alpha_n\}} a^{\alpha_2}_2 \cdot a^{\alpha_3}_3 \cdot \ldots \cdot a^{\alpha_n}_n = M^1_1.$$

2) Пусть теперь $i, j \neq 1$, то есть $i, j = 2, \ldots, n$. Тогда элемент a^i_j путём последовательных $i-1$ перестановок соседних строк и $j-1$ соседних столбцов переместим на место элемента a^1_1. Полученную матрицу обозначим через B. Минор M^i_j, соответствующий элементу a^i_j, одинаков как в матрице A, так и в матрице B, так как относительное положение строк не изменилось. Поэтому, используя первый пункт доказательства, получаем:

$$\det B = a^i_j M^i_j + \left(\text{члены, не содержащие } a^i_j\right).$$

По свойству 2 определителей имеем:

$$\det B = (-1)^{i-1} \cdot (-1)^{j-1} \cdot \det A = (-1)^{i+j} \det A.$$

Из двух последних равенств имеем:

$$\det A = (-1)^{i+j} \cdot a_j^i M_j^i + \left(\text{члены, не содержащие } a_j^i\right).$$

Сравнивая с разложением (6.2.8) (или (6.2.9)), получаем доказательство свойства. ••

Следствие из свойства 6. *Сумма всех произведений элементов какого-либо столбца (строки) определителя D на алгебраические дополнения соответствующих элементов любого другого столбца (строки) равна нулю.*

Д о к а з а т е л ь с т в о. Если в разложении определителя

$$D = (-1)^{1+j} a_j^1 M_j^1 + (-1)^{2+j} a_j^2 M_j^2 + \ldots + (-1)^{n+j} a_j^n M_j^n \qquad (6.2.11)$$

заменить a_j^i $(i = 1, 2, \ldots, n)$ на любые другие величины, то алгебраические дополнения $A_j^1, A_j^2, \ldots, A_j^n$ (определение 6.2.4) не изменятся, так как от a_j^i $(i = 1, 2, \ldots, n)$ они не зависят. Поэтому, заменив в разложении (6.2.11) справа и слева элементы $a_j^1, a_j^2, \ldots, a_j^n$ на элементы, например, k-го столбца, получим, что определитель слева в разложении (6.2.11) имеет два одинаковых столбца и, следовательно, будет равен нулю. Получаем требуемый результат:

$$a_k^1 A_j^1 + a_k^2 A_j^2 + a_k^i A_j^i + \ldots + a_k^n A_j^n = 0. \;••$$

Свойство 7. *Определитель произведения двух матриц равен произведению определителей этих матриц.*

Д о к а з а т е л ь с т в о. Для доказательства заметим, что определитель единичной матрицы $\det\left(I_i^j\right) = 1$. Это свойство легко проверяется, если определитель единичной матрицы (произвольного порядка) раскрыть по элементам какого-либо столбца (или строки). Кроме этого, свойства 2 и 3 определителей являются основными (из них можно получить все остальные свойства). Поэтому, если мы для двух матриц A и B в соответствии с утверждением леммы положим

$$\det(AB) = \det A \cdot \det B$$

и докажем, что определённая таким образом величина обладает всеми свойствами определителей, то мы докажем лемму.

Итак, пусть матрицы A и B невырождены, в противном случае равенство

$$\det(AB)=\det A\cdot\det B$$

сводится к тождеству $0=0$. Рассмотрим величину

$$D(A)=\frac{\det(AB)}{\det B}.$$

Пусть матрица A — единичная. Тогда имеем:

$$D(I)=\frac{\det(IB)}{\det B}=\frac{\det(B)}{\det B}=1.$$

Таким образом, если матрица $A=I$, то величина

$$D(I)=1.$$

Легко проверить, что при перестановке двух столбцов матрицы A то же самое произойдёт и с соответствующими столбцами матрицы AB и, следовательно, знак величины $D(A)$ изменится на противоположный.

Наконец, нетрудно проверить, что если некоторый столбец матрицы A представлен в виде линейной комбинации конечного числа слагаемых, то же самое произойдёт и со столбцами матрицы AB.

Итак, получили, что величина $D(A)$ обладает всеми свойствами определителей и, следовательно,

$$D(A)\equiv\det A=\frac{\det(AB)}{\det B},$$

откуда

$$\det(AB)=\det A\cdot\det B. \bullet\bullet$$

6.3. Билинейные и квадратичные формы

Определение билинейных и квадратичных форм. Выше (параграф 3.2) мы достаточно подробно изучили теорию квадратичных форм в частном случае евклидова пространства R^n. Однако квадратичные формы могут появиться и в более общем случае абстрактного векторного пространства. Основой теории квадратичных форм в этом случае служит теория ***билинейных форм*** (***функций***). В данном дополнении устанавливается связь теории квадратичных форм с теорией билинейных форм в общем случае абстрактного векторного пространства X^n.

С теорией билинейных форм мы, по существу, уже встречались при изучении теории (собственно) евклидова пространства. Так, в параграфе 2.5 (формула (2.5.4)) установлено, что в евклидовом пространстве E^n скалярное произведение, являясь симметрической, билинейной функцией, удовлетворяющей аксиомам симметрии и невырожденности, выражается через координаты векторов пространства E^n так:

$$\left(\forall \vec{x}, \vec{y} \in E^n\right)\left(\vec{x}, \vec{y}\right) = \left(\sum_{i=1}^{n} x^i \vec{e}_i, \sum_{j=1}^{n} y^j \vec{e}_j\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j \left(\vec{e}_i, \vec{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i y^j . \qquad (6.3.1)$$

Правая часть формулы (6.3.1) может рассматриваться как многочлен, составленный из подобных одночленов одинаковой степени от $2n$ неизвестных и, следовательно, является однородной функцией (формой) второй степени. Действительно, произведём в формуле (6.3.1) замены

$$x^i \to tx^i,\ y^j \to ty^j.$$

Подстановка в формулу (6.3.1) даёт

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} tx^i ty^j = t^2 \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i y^j .$$

Таким образом, свойство однородности проявляется по отношению к каждому аргументу формы. Более того, эта форма, имея вторую степень по переменным, является билинейной, то есть линейной по обоим векторным аргументам. Если в формуле (6.3.1) положить $\vec{x} = \vec{y}$, то получим скалярный квадрат вектора $\vec{x}$ в виде

$$\left(\vec{x}, \vec{x}\right) = \left(\sum_{i=1}^{n} x^i \vec{e}_i, \sum_{j=1}^{n} x^j \vec{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i x^j \left(\vec{e}_i, \vec{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i x^j . \qquad (6.3.2)$$

Форма (6.3.2) также имеет вторую степень, но уже является ***нелинейной***. Напомним, что матрица $\left(g_{ij}\right)$ в параграфе 2.5 была названа матрицей метрических коэффициентов.

Значение скалярного произведения для геометрии очевидно из всего предыдущего изложения. Однако, значение билинейных и квадратичных форм, не связанных напрямую с понятием скалярного произведения векторов в евклидовом пространстве, не ограничивается только применениями в теории измерений в векторных пространствах. Теория билинейных и квадратичных форм находит иные, весьма широкие применения как в собственно геометрии, так и в других разделах математики и математической физики. Перейдём к общим определениям.

Определение 6.3.1. *Скалярная функция* $\varphi = \varphi\left(\vec{x}, \vec{y}\right)$ *(отображение* $\varphi : X^n \times X^n \to P$*) двух векторных переменных* $\vec{x}, \vec{y} \in X^n$ *из абстрактного векторного пространства* X^n *называется* ***билинейной функцией*** *или* ***билинейной формой****, если она является линейной по обоим аргументам, то есть если*

$$\left(\forall \vec{x}, \vec{y}, \vec{z} \in X^n\right) \wedge \left(\forall \alpha \in P\right)$$

$$f\left(\vec{x}+\vec{y}, \vec{z}\right) = f\left(\vec{x}, \vec{z}\right) + f\left(\vec{y}, \vec{z}\right),$$

$$f\left(\alpha \vec{x}, \vec{y}\right) = \alpha f\left(\vec{x}, \vec{y}\right),$$

$$f\left(\vec{x}, \vec{y}+\vec{z}\right) = f\left(\vec{x}, \vec{y}\right) + f\left(\vec{x}, \vec{z}\right),$$

$$f\left(\vec{x}, \alpha \vec{y}\right) = \alpha f\left(\vec{x}, \vec{y}\right). \bullet \tag{6.3.3}$$

Если $\left\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\right\}$ и $\left\{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_m\right\}$ — две системы произвольных векторов из пространства X^n, а $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ и $\{\beta_1, \beta_2, \ldots, \beta_m\}$ — две системы произвольных чисел из поля P, то из свойств (6.3.3) получается ***общее выражение для билинейной формы***:

$$f\left(\sum_{i=1}^{k} \alpha_i \vec{x}_i, \sum_{j=1}^{m} \beta_j \vec{y}_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{m} \alpha_i \beta_j f\left(\vec{x}_i, \vec{y}_j\right). \tag{6.3.4}$$

Отметим, что билинейную форму часто называют ещё ***билинейным функционалом*** (от двух переменных).

Теперь становится очевидным, что квадратичные формы (параграф 3.2) являются специальным случаем билинейных форм.

Определение 6.3.2. ***Квадратичной формой*** *в пространстве* X^n *называется функция* $\varphi\left(\vec{x}, \vec{x}\right)$ *от одного векторного аргумента* $\vec{x} \in X^n$*, которая получается из произвольной билинейной формы* $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ *путём замены* $\vec{y}$ *на* $\vec{x}$. •

Из определения видно, что функция $\varphi\left(\vec{x}, \vec{x}\right)$, как было отмечено выше, уже не является линейной функцией.

Матрицы билинейных и квадратичных форм. Пусть в пространстве X^n задана билинейная форма $\mathrm{f}\left(\vec{x}, \vec{y}\right)$. Зафиксируем базис $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$. Тогда $\left(\forall \vec{x}, \vec{y} \in X^n\right)$ справедливы разложения:

$$\vec{x} = x^1 \vec{a}_1 + x^2 \vec{a}_2 + \ldots + x^n \vec{a}_n = \sum_{i=1}^{n} x^i \vec{a}_i,$$

$$\vec{y} = y^1 \vec{a}_1 + y^2 \vec{a}_2 + \ldots + y^n \vec{a}_n = \sum_{j=1}^{n} y^j \vec{a}_j.$$

Если обозначить $\mathrm{f}\left(\vec{a}_i, \vec{a}_j\right) = \mathrm{f}_{ij}$, то по формуле (6.3.4) получим:

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \mathrm{f}\left(\sum_{i=1}^{n} x^i \vec{a}_i, \sum_{j=1}^{n} y^j \vec{a}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j \mathrm{f}\left(\vec{a}_i, \vec{a}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i y^j. \quad (6.3.5)$$

В формуле (6.3.5) коэффициенты f_{ij} образуют матрицу

$$\Phi = \left(\mathrm{f}_{ij}\right) = \begin{pmatrix} \mathrm{f}_{11} & \mathrm{f}_{12} & \ldots & \mathrm{f}_{1n} \\ \mathrm{f}_{21} & \mathrm{f}_{22} & \ldots & \mathrm{f}_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \mathrm{f}_{n1} & \mathrm{f}_{n2} & \ldots & \mathrm{f}_{nn} \end{pmatrix}, \quad (6.3.6)$$

которая называется ***матрицей билинейной формы в базисе*** $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$.

При переходе к новому базису матрица билинейной формы изменяется. Получим закон преобразования матрицы билинейной формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ при переходе к новому базису.

Итак, пусть в пространстве X^n зафиксированы два базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ и $\left\{\vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'}\right\}$. Переход от старого к новому базису осуществляется по формулам

$$\vec{a}_{i'} = \sum_{j=1}^{n} A_{i'}^{j} \vec{a}_j. \tag{6.3.7}$$

В старом базисе $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ выражение для билинейной формы (6.3.5) имеет вид

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i y^j \mathrm{f}\left(\vec{a}_i, \vec{a}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i y^j.$$

В новом базисе $\left\{\vec{a}_{1'}, \vec{a}_{2'}, \ldots, \vec{a}_{n'}\right\}$ выражение для билинейной формы имеет аналогичный вид

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{k'=1}^{n}\sum_{p'=1}^{n} x^{k'} y^{p'} \mathrm{f}\left(\vec{a}_{k'}, \vec{a}_{p'}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{k'p'} x^{k'} y^{p'}.$$

Учитывая формулу (6.3.7) перехода от старого базиса к новому базису, получаем

$$\mathrm{f}_{k'p'} = \mathrm{f}\left(\vec{a}_{k'}, \vec{a}_{p'}\right) = \mathrm{f}\left(\sum_{i=1}^{n} A_{k'}^{i} \vec{a}_i, \sum_{j=1}^{n} A_{p'}^{j} \vec{a}_j\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} A_{k'}^{i} A_{p'}^{j} \mathrm{f}\left(\vec{a}_i, \vec{a}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{k'}^{i} A_{p'}^{j} \mathrm{f}_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{i}^{k'} \mathrm{f}_{ij} A_{p'}^{j}. \tag{6.3.8}$$

Здесь $A_i^{k'}$ — элемент матрицы, транспонированной к матрице перехода от старого базиса к новому базису. В матричных обозначениях формула (6.3.8) имеет вид:

$$\Phi' = A^T \Phi A. \tag{6.3.8'}$$

Ранг матрицы Φ' равен рангу матрицы Φ в силу невырожденности матрицы A, а следовательно, и матрицы A^T. Таким образом, *ранг матрицы Φ билинейной формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ не зависит от выбора базиса и по этой причине называется* ***рангом билинейной формы*** $\mathrm{f}\left(\vec{x}, \vec{y}\right)$.

Теперь очевидно, что ***матрица квадратичной формы*** $\varphi\left(\vec{x}, \vec{y}\right)$**,** порождённой билинейной формой $\mathrm{f}\left(\vec{x}, \vec{y}\right)$, это просто матрица порождающей билинейной формы.

В пространстве X^n с базисом $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ каждая квадратичная форма имеет в силу формулы (6.3.5) следующий вид:

$$\varphi\left(\vec{x}, \vec{x}\right) = \varphi\left(\sum_{i=1}^{n} x^i \vec{a}_i, \sum_{j=1}^{n} x^j \vec{a}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i x^j \varphi\left(\vec{a}_i, \vec{a}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij} x^i x^j. \quad (6.3.9)$$

Нетрудно показать, что имеет место и обратное, а именно, справедливо следующее утверждение, которое оформим в виде вспомогательной теоремы.

Лемма 6.3.1. *Если функция* $\varphi\left(\vec{x}, \vec{x}\right)$ *в базисе* $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$ *определена формулой (6.3.9), то эта функция является квадратичной формой от вектора* $\vec{x} \in X^n$.

Д о к а з а т е л ь с т в о. Для доказательства достаточно ввести билинейную форму

$$\psi\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i y^j.$$

Тогда, очевидно, что квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$, полученная путём такой замены из билинейной формы $\psi\left(\vec{x}, \vec{y}\right)$, совпадёт с функцией $\varphi\left(\vec{x}, \vec{x}\right)$. ••

Симметрические билинейные формы. Для приложений основное значение имеют ***симметрические билинейные формы***.

Определение 6.3.3. *Билинейная форма* $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ *называется* ***симметричной****, если*

$$\left(\forall \vec{x}, \vec{y} \in X^n\right) \mathrm{f}\left(\vec{x}, \vec{y}\right) = \mathrm{f}\left(\vec{y}, \vec{x}\right). \bullet \quad (6.3.10)$$

Таким образом, по определению для симметричной билинейной формы, определённой на пространстве X^n, имеем:

$$\mathrm{f}_{ij} = \mathrm{f}\left(\vec{a}_i, \vec{a}_j\right) = \mathrm{f}\left(\vec{a}_j, \vec{a}_i\right) = \mathrm{f}_{ji}.$$

Теперь видно, что матрица симметричной билинейной формы в любом базисе пространства X^n удовлетворяет условию $\Phi = \Phi^T$, то есть является симметрической матрицей.

Лемма 6.3.2. *Если в некотором базисе*

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$$

матрица $\Phi = \left(\mathrm{f}_{ij}\right)$ *симметрична, то билинейная форма* $\mathrm{f}\left(\vec{x}, \vec{y}\right)$*, которой она соответствует, — симметрическая.*

Д о к а з а т е л ь с т в о. Действительно, в этом случае имеем:

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i y^j = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ji} x^i y^j =$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n} \mathrm{f}_{ji} y^j x^i = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} y^i x^j = \mathrm{f}\left(\vec{y}, \vec{x}\right). \bullet\bullet$$

В двойной сумме (6.3.9) можно привести подобные члены. Действительно, положим для $i \neq j$

$$\mathrm{f}_{ij} x^i x^j + \mathrm{f}_{ji} x^j x^i = \left(\mathrm{f}_{ij} + \mathrm{f}_{ji}\right) x^i x^j = \psi_{ij} x^i x^j,$$

а для $i = j$ положим $\psi_{ii} = \mathrm{f}_{ii}$. Тогда вместо двойной суммы (6.3.9) имеем:

$$\mathrm{f}\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i x^j = \sum_{j=1}^{n}\sum_{i<j} \varphi_{ij} x^i x^j. \qquad (6.3.11)$$

Из этого следует, что две различные билинейные формы

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i y^j, \ \psi\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \psi_{ij} x^i y^j$$

иногда возможно привести после замены аргумента $\vec{y}$ на аргумент $\vec{x}$ к одной и той же квадратичной форме. Для этого нужно, чтобы выполнялось равенство

$$\left(\forall i, j = 1, 2, \ldots, n\right) \mathrm{f}_{ij} + \mathrm{f}_{ji} = \psi_{ij} + \psi_{ji}.$$

Следовательно, *восстановить единственным образом по квадратичной форме породившую её билинейную форму в общем случае невозможно*. Однако справедливо следующее утверждение.

Лемма 6.3.3. *Пусть* $\varphi\left(\vec{x}, \vec{x}\right)$ *— квадратичная форма, порождённая билинейной формой* $f\left(\vec{x}, \vec{y}\right)$. *Тогда, если порождающая билинейная форма* $f\left(\vec{x}, \vec{y}\right)$ *симметрическая, то её можно единственным образом восстановить по квадратичной форме* $\varphi\left(\vec{x}, \vec{x}\right)$.

Д о к а з а т е л ь с т в о. Пусть $f\left(\vec{x}, \vec{y}\right)$ — симметрическая билинейная форма. Тогда для любых векторов $\vec{x}$ и $\vec{y}$ из пространства X^n

$$f\left(\vec{x}+\vec{y}, \vec{x}+\vec{y}\right) = \varphi\left(\vec{x}, \vec{x}\right) + f\left(\vec{x}, \vec{y}\right) + f\left(\vec{y}, \vec{x}\right) + \varphi\left(\vec{y}, \vec{y}\right),$$

откуда

$$f\left(\vec{x}, \vec{y}\right) = \frac{1}{2}\left[f\left(\vec{x}, \vec{y}\right) + f\left(\vec{y}, \vec{x}\right)\right] =$$

$$= \frac{1}{2}\left[f\left(\vec{x}+\vec{y}, \vec{x}+\vec{y}\right) - \varphi\left(\vec{x}, \vec{x}\right) - \varphi\left(\vec{y}, \vec{y}\right)\right].$$

Следовательно, значение симметричной билинейной формы $f\left(\vec{x}, \vec{y}\right)$ единственным образом определяется по значениям квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$ на векторах $\vec{x}$, $\vec{y}$ и $\vec{x}+\vec{y}$. ●●

Пусть $f\left(\vec{x}, \vec{y}\right)$ — произвольная билинейная форма, тогда билинейная форма

$$\psi\left(\vec{x}, \vec{y}\right) = \frac{1}{2}\left[f\left(\vec{x}, \vec{y}\right) + f\left(\vec{y}, \vec{x}\right)\right],$$

очевидно, симметрическая. Поэтому

$$\psi\left(\vec{x}, \vec{x}\right) = \frac{1}{2}\left[f\left(\vec{x}, \vec{x}\right) + f\left(\vec{x}, \vec{x}\right)\right] = \varphi\left(\vec{x}, \vec{x}\right).$$

Итак, билинейная форма $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ и полученная из неё симметрическая билинейная форма $\psi\left(\vec{x}, \vec{y}\right)$ порождают одну и ту же квадратичную форму.

В силу приведённых соображений ***при изучении квадратичных форм достаточно ограничиться симметрическими порождающими билинейными формами***.

По этой причине ***матрицей квадратичной формы*** $\varphi\left(\vec{x}, \vec{x}\right)$ естественно назвать симметричную матрицу $\Phi = \left(\mathrm{f}_{ij}\right)$ породившей её симметричной билинейной формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$.

При изменении базиса матрица квадратичной формы преобразуется по формуле (5.2.7). В частности, *ранг матрицы* Φ *квадратичной формы* $\varphi\left(\vec{x}, \vec{x}\right)$ *не зависит от выбора базиса и по этой причине называется* ***рангом квадратичной формы*** $\varphi\left(\vec{x}, \vec{x}\right)$. *Квадратичная форма, имеющая ранг матрицы, совпадающий с размерностью пространства n, называется* ***невырожденной***.

Приведение квадратичной формы к каноническому виду. Пусть на векторах абстрактного векторного пространства X^n задана квадратичная форма. Вид этой квадратичной формы в общем случае даётся выражением (6.3.9):

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{ij} x^i x^j .$$

Рассмотрим вопрос о приведении квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$ к сумме квадратов, то есть к такому виду, в котором все коэффициенты при попарных произведениях разноимённых координат равны нулю.

Теорема 6.3.1. *Пусть* $\varphi\left(\vec{x}, \vec{x}\right)$ *— произвольная квадратичная форма, определённая на векторах абстрактного векторного пространства* X^n*. Тогда найдётся базис* $\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\} \subset X^n$*, такой, что в этом базисе квадратичная форма примет вид*

$$\varphi\left(\vec{x}, \vec{x}\right) = \mu_1 \left(x^{1'}\right)^2 + \mu_2 \left(x^{2'}\right)^2 + \ldots + \mu_n \left(x^{n'}\right)^2, \qquad (6.3.12)$$

где $\mu_1, \mu_2, \ldots, \mu_n$ *— некоторые фиксированные числа из поля P.*

Д о к а з а т е л ь с т в о. Пусть квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$ содержит лишь одну координату с ненулевым коэффициентом, например, x^1, то есть имеет вид

$$\varphi\left(\vec{x}, \vec{x}\right) = \varphi_{11}\left(x^1\right)^2.$$

Утверждение теоремы, очевидно, выполняется.

Пусть утверждение теоремы справедливо для любых квадратичных форм, содержащих $m-1$ координат. Рассмотрим форму, содержащую m координат:

$$\varphi\left(\vec{x}, \vec{x}\right) = \varphi_{11}\left(x^1\right)^2 + \varphi_{12}x^1x^2 + \varphi_{13}x^1x^3 + \ldots + \varphi_{1m}x^1x^m +$$
$$+\varphi_{21}x^2x^1 + \varphi_{22}\left(x^2\right)^2 + \varphi_{23}x^2x^3 + \ldots + \varphi_{2m}x^2x^m +$$
$$+\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots +$$
$$+\varphi_{m1}x^mx^1 + \varphi_{m2}x^mx^2 + \varphi_{m3}x^mx^3 + \ldots + \varphi_{mm}\left(x^m\right)^2 =$$
$$= \varphi_{11}\left(x^1\right)^2 + 2\varphi_{12}x^1x^2 + \varphi_{22}\left(x^2\right)^2 + \ldots + \varphi_{mm}\left(x^m\right)^2. \qquad (6.3.13)$$

Если хотя бы при одном квадрате, например при $\left(x^m\right)^2$, имеется отличный от нуля коэффициент, то соберём все слагаемые, содержащие x^m:

$$2\varphi_{1m}x^1x^m + 2\varphi_{2m}x^2x^m + \ldots + 2\varphi_{m-1,m}x^{m-1}x^m + \varphi_{mm}\left(x^m\right)^2.$$

Выделим в этой группе слагаемых полный квадрат:

$$2\varphi_{1m}x^1x^m + 2\varphi_{2m}x^2x^m + \ldots + 2\varphi_{m-1,m}x^{m-1}x^m + \varphi_{mm}\left(x^m\right)^2 =$$
$$= \frac{1}{\varphi_{mm}}\left(\varphi_{1m}x^1 + \varphi_{2m}x^2 + \ldots + \varphi_{m-1\,m}x^{m-1} + \varphi_{mm}x^m\right)^2 -$$
$$-\frac{1}{\varphi_{mm}}\left(\varphi_{1m}x^1 + \varphi_{2m}x^2 + \ldots + \varphi_{m-1\,m}x^{m-1}\right)^2. \qquad (6.3.14)$$

Тогда квадратичная форма (6.3.13) примет вид:

$$\varphi\left(\vec{x}, \vec{x}\right) = \frac{1}{\varphi_{mm}}\left(\varphi_{1m}x^1 + \varphi_{2m}x^2 + \ldots + \varphi_{m-1,\,m}x^{m-1} + \varphi_{mm}x^m\right)^2 + \psi\left(\vec{x}, \vec{x}\right).$$

Здесь квадратичная форма $\psi\left(\vec{x}, \vec{x}\right)$ зависит уже только от $m-1$ координат $x^1, x^2, \ldots, x^{m-1}$.

Положим

$$
\begin{aligned}
\xi^1 &= x^1,\\
\xi^2 &= x^2,\\
&\dots\dots\dots\dots,\\
\xi^{m-1} &= x^{m-1},\\
\xi^m &= \varphi_{1m}x^1 + \varphi_{2m}x^2 + \ldots + \varphi_{m-1\,m}x^{m-1} + \varphi_{mm}x^m,\\
\xi^{m+1} &= x^{m+1},\\
&\dots\dots\dots\dots\dots\dots\dots\dots,\\
\xi^n &= x^n.
\end{aligned} \tag{6.3.15}
$$

Нетрудно видеть, что определитель матрицы преобразования (6.3.15) равен $\varphi_{mm} \neq 0$. Поэтому преобразование координат (6.3.15) вызвано переходом к новому базису, причём, как известно, матрица перехода является транспонированной обратной матрицей преобразования (6.3.15). По предположению индукции квадратичную форму $\psi\left(\vec{x}, \vec{x}\right)$, зависящую уже только от $m-1$ координат $x^1, x^2, \ldots, x^{m-1}$, путём перехода к новому базису можно привести к виду (6.3.12). Поэтому и квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$ приведётся к виду (6.3.12).

Пусть теперь все коэффициенты при квадратах равны нулю, то есть $\left(\forall i = 1, 2, \ldots, m\right)$ $\varphi_{ii} = 0$. Допустим, например, что $\varphi_{12} \neq 0$. Положим

$$x^1 = \xi^1 + \xi^2,\, x^2 = \xi^1 - \xi^2,\, x^3 = \xi^3, \ldots, x^n = \xi^n.$$

Это преобразование соответствует переходу к базису

$$\vec{e}_{1'} = \vec{e}_1 + \vec{e}_2,\, \vec{e}_{2'} = \vec{e}_1 - \vec{e}_2,\, \vec{e}_{3'} = \vec{e}_3, \ldots, \vec{e}_{n'} = \vec{e}_n$$

с матрицей перехода

$$A = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 1 & -1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Определитель этой матрицы равен $\det A = -2$. При этом преобразовании $x^1x^2 = \left(\xi^1\right)^2 - \left(\xi^2\right)^2$ и приходим к предыдущему случаю, когда имеется квадрат с отличным от нуля коэффициентом. ••

Итак, доказано, что если в пространстве X^n задана произвольная квадратичная форма, то найдётся базис $\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\} \subset X^n$, такой, что в этом базисе квадратичная форма примет вид (6.3.12), где $x^{1'}, x^{2'}, \ldots, x^{n'}$ — координаты вектора $\vec{x}$ в новом базисе.

Определение 6.3.4. *Говорят, что квадратичная форма* $\varphi\left(\vec{x}, \vec{x}\right)$*, определённая на векторах пространства* X^n*, приведена к* ***каноническому виду****, если в некотором базисе* $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset X^n$ *она может быть записана в виде (6.3.12). Базис, в котором квадратичная форма принимает вид (6.3.12), называется* ***каноническим базисом формы*** $\varphi\left(\vec{x}, \vec{x}\right)$. •

Отметим, что канонический вид (6.3.12) квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$ не является единственным, что непосредственно следует из его получения. Естественно, что и канонический базис билинейной формы не является единственным. Действительно, выделение полного квадрата можно производить относительно другой координаты, нежели x^m. Тем не менее, ниже мы увидим, что у всех канонических видов записи квадратичной формы есть важная общая характеристика, которая сохраняется при переходе от одного канонического базиса к другому.

Канонический базис билинейной формы, метод Якоби. Для билинейной формы, так же как и для квадратичной формы, можно построить базис, в котором форма $\phi\left(\vec{x}, \vec{y}\right)$ имеет наиболее простой вид. Начнём с определения.

Определение 6.3.5. *Вектор* $\vec{x} \in X^n$ *называется* ***сопряжённым с вектором*** $\vec{y} \in X^n$ ***относительно билинейной формы*** $\mathrm{f}\left(\vec{x}, \vec{y}\right)$*, если*

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = 0. \bullet$$

Пусть $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$ — некоторый фиксированный базис и

$$\Phi = \left(\mathrm{f}_{ij}\right) \left(i, j = 1, 2, \ldots, n\right)$$

— матрица билинейной формы $\phi\left(\vec{x}, \vec{y}\right)$ в этом базисе. Так как

$$\vec{x}=\sum_{i=1}^{n}x^{i}\vec{e}_{i},\ \vec{y}=\sum_{j=1}^{n}y^{j}\vec{e}_{j},$$

то условие сопряжённости векторов $\vec{x}$ и $\vec{y}$ принимает вид

$$\mathrm{f}\left(\vec{x},\vec{y}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{f}_{ij}x^{i}y^{j}=0. \qquad (6.3.16)$$

Лемма 6.3.4. *Пусть* $\mathrm{f}\left(\vec{x},\vec{y}\right)$ *— билинейная форма, определённая на векторах абстрактного векторного пространства* X^{n}*. Тогда, если векторы системы*

$$\left\{\vec{x}_{1},\vec{x}_{2},\ldots,\vec{x}_{m}\right\}\subset X^{n}$$

сопряжены с вектором $\vec{y}\subset X^{n}$*, то и любая линейная комбинация*

$$\vec{x}=\alpha_{1}\vec{x}_{1}+\alpha_{2}\vec{x}_{2}+\ldots+\alpha_{m}\vec{x}_{m},$$

где коэффициенты $\alpha_{1},\alpha_{2},\ldots,\alpha_{m}$ *— произвольные, не равные нулю одновременно числа из поля P, также сопряжена с вектором* $\vec{y}$.

Доказательство. Действительно, из свойств билинейной формы получаем:

$$\mathrm{f}\left(\vec{x},\vec{y}\right)=\mathrm{f}\left(\alpha_{1}\vec{x}_{1}+\alpha_{2}\vec{x}_{2}+\ldots+\alpha_{m}\vec{x}_{m},\vec{y}\right)=$$

$$=\alpha_{1}\mathrm{f}\left(\vec{x}_{1},\vec{y}\right)+\alpha_{2}\mathrm{f}\left(\vec{x}_{2},\vec{y}\right)+\ldots+\alpha_{m}\mathrm{f}\left(\vec{x}_{m},\vec{y}\right)=0.\ \bullet\bullet$$

Итак, из леммы следует, что если векторы $\vec{x}_{1},\vec{x}_{2},\ldots,\vec{x}_{m}$ сопряжены с вектором $\vec{y}$, то и любой вектор $\vec{x}\in L\left\{\vec{x}_{1},\vec{x}_{2},\ldots,\vec{x}_{m}\right\}$, где $L\left\{\vec{x}_{1},\vec{x}_{2},\ldots,\vec{x}_{m}\right\}$ — линейная оболочка системы векторов $\vec{x}_{1},\vec{x}_{2},\ldots,\vec{x}_{m}$ (подпространство пространства X^{n}), также сопряжён с вектором $\vec{y}$. Вектор $\vec{y}$, сопряжённый со всеми векторами подпространства $L\left\{\vec{x}_{1},\vec{x}_{2},\ldots,\vec{x}_{m}\right\}$, называется ***сопряжённым к подпространству L***.

Из леммы 6.3.4 имеем такие следствия.

Следствие 1. *Вектор* $\vec{y} \in X^n$ *является сопряжённым к подпространству* $L^m \subset X^n$ *в том и только в том случае, если он сопряжён к векторам некоторого базиса* $\left\{ \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m \right\} \subset L^m$.

Следствие 2. *Совокупность* L' *всех векторов* $\vec{y} \in X^n$, *сопряжённых к подпространству* $L^m \subset X^n$, *является подпространством пространства* X^n.

Подпространство L', все векторы которого сопряжены к подпространству $L \subset X^n$, называется подпространством, ***сопряжённым к*** $L \subset X^n$.

Определение 6.3.6. *Пусть* $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ *— билинейная форма, определённая на векторах пространства* X^n. *Базис* $\left\{ \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n \right\} \subset X^n$ *называется* ***каноническим базисом билинейной формы*** $\mathrm{f}\left(\vec{x}, \vec{y}\right)$, *если базисные векторы взаимно сопряжены, то есть* $\mathrm{f}\left(\vec{e}_i, \vec{e}_j\right) = 0$ *при* $i \neq j$. •

Пример 6.3.1. Пусть $\vec{x}, \vec{y} \in R^n$. В качестве билинейной формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ возьмём скалярное произведение $\left(\vec{x}, \vec{y}\right)$ векторов. Сопряжённость векторов относительно билинейной формы

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\vec{e}_i, \vec{e}_j\right) x^i y^j,$$

очевидно, эквивалентна их ортогональности. Следовательно, каноническим базисом для формы $\left(\vec{x}, \vec{y}\right)$ является любой ортогональный базис в пространстве R^n. ⊗

Матрица билинейной формы в каноническом базисе является диагональной, так как $\mathrm{f}_{ij} = \mathrm{f}\left(\vec{e}_i, \vec{e}_j\right) = 0$ при $i \neq j$, следовательно, ***матрица билинейной формы в каноническом базисе симметрическая***.

Теорема 6.3.2. *Для каждой симметричной билинейной формы* $f\left(\vec{x}, \vec{y}\right)$, *определённой на векторах пространства* X^n, *существует канонический базис.*

Д о к а з а т е л ь с т в о. Рассмотрим квадратичную форму $\varphi\left(\vec{x}, \vec{x}\right)$, порождённую билинейной формой $\varphi\left(\vec{x}, \vec{x}\right)$. По теореме 6.3.1 в пространстве X^n существует базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, в котором квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$ имеет канонический вид: $\varphi\left(\vec{x}, \vec{x}\right) = \sum\limits_{i=1}^{n} \mu_i \left(x^i\right)^2$. Порождающая симметричная билинейная форма $\varphi\left(\vec{x}, \vec{x}\right)$ по лемме 6.3.3 восстанавливается по квадратичной форме $\varphi\left(\vec{x}, \vec{x}\right)$ единственным образом. Для пояснения достаточно заметить, что из равенства $\varphi_{ij} + \varphi_{ji} = \psi_{ij}$ для $i \neq j$ в формуле (6.3.11) получаем $\varphi_{ij} = \frac{1}{2}\psi_{ij}\ (i \neq j)$ и $\varphi_{ii} = \psi_{ii}$. Поэтому для билинейной формы имеем $\varphi\left(\vec{x}, \vec{y}\right) = \sum\limits_{i=1}^{n} \mu_i x^i y^i$, где $\vec{x} = \sum\limits_{j=1}^{n} x^j \vec{e}_j$, $\vec{y} = \sum\limits_{j=1}^{n} y^j \vec{e}_j$. Следовательно, матрица билинейной формы $\varphi\left(\vec{x}, \vec{x}\right)$ является диагональной. Последнее и означает, что базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ является каноническим для формы $\varphi\left(\vec{x}, \vec{x}\right)$. ●●

Пусть $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m\right\}$ — канонический базис билинейной формы $f\left(\vec{x}, \vec{y}\right)$ в m-мерном подпространстве $L^m \subset X^n$. Пусть, далее, $\mu_1, \mu_2, \ldots, \mu_m$ — соответствующие канонические коэффициенты. Выразим числа

$$f\left(\vec{x}, \vec{e}_j\right)\ (j = 1, 2, \ldots, m)$$

через координаты вектора $\vec{x} \in L^m$:

$$f\left(\vec{x}, \vec{e}_j\right) = f\left(\sum_{i=1}^{m} x^i \vec{e}_i, \vec{e}_j\right) = \sum_{i=1}^{m} x^i f\left(\vec{e}_i, \vec{e}_j\right) = x^j f\left(\vec{e}_j, \vec{e}_j\right) = \mu_j x^j.$$

Видим, что числа $f\left(\vec{x}, \vec{e}_j\right)$ единственным образом определяются координатами вектора $\vec{x}$. Если билинейная форма $f\left(\vec{x}, \vec{e}_j\right)$ невырождена в подпространстве L^m, то все числа μ_j отличны от нуля. В этом случае справедливо и обратное, а именно: *значения формы* $f\left(\vec{x}, \vec{e}_j\right)$ *единственным образом определяют координаты вектора* $\vec{x}$.

Вышеизложенное является обобщением результатов из параграфа 2.5 относительно двух типов координат в евклидовом пространстве.

Рассмотрим метод Якоби построения канонического базиса симметричной билинейной формы $f\left(\vec{x}, \vec{y}\right)$, определённой на векторах абстрактного векторного пространства X^n.

Итак, пусть $f\left(\vec{x}, \vec{y}\right)$ — симметричная билинейная форма с матрицей $\Phi = \left(f_{ij}\right)$ в фиксированном базисе $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\}$ пространства X^n. Наложим на матрицу Φ дополнительное условие, заключающееся в том, все главные миноры этой матрицы до порядка $n-1$ включительно отличны от нуля, то есть:

$$M_1 = f_{11} \neq 0;\ M_2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} \neq 0;\ M_3 = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} \neq 0;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ M_{n-1} = \begin{vmatrix} f_{11} & f_{12} & \ldots & f_{1,\,n-1} \\ f_{21} & f_{22} & \ldots & f_{2,\,n-1} \\ \ldots & \ldots & \ldots & \ldots \\ f_{n-1,\,1} & f_{n-1,\,2} & \ldots & f_{n,\,n-1} \end{vmatrix} \neq 0.$$

Определим векторы $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ искомого канонического базиса следующими формулами:

$$\begin{aligned} \vec{e}_1 &= \vec{a}_1, \\ \vec{e}_2 &= \alpha_1^1 \vec{a}_1 + \vec{a}_2, \\ \vec{e}_3 &= \alpha_2^1 \vec{a}_1 + \alpha_2^2 \vec{a}_2 + \vec{a}_3, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \vec{e}_{k+1} &= \alpha_k^1 \vec{a}_1 + \alpha_k^2 \vec{a}_2 + \ldots + \alpha_k^k \vec{a}_k + \vec{a}_{k+1}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \vec{e}_n &= \alpha_{n-1}^1 \vec{a}_1 + \alpha_{n-1}^2 \vec{a}_2 + \ldots + \alpha_{n-1}^{n-1} \vec{a}_{n-1} + \vec{a}_n. \end{aligned} \qquad (6.3.17)$$

В формулах (6.3.17) коэффициенты α_j^i $(i, j = 1, 2, \ldots, n-1)$ подлежат определению.

Матрица перехода от векторов $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ к векторам $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k$ имеет вид

$$A = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ \alpha_1^1 & 1 & 0 & \ldots & 0 & 0 \\ \alpha_2^1 & \alpha_2^2 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n-1}^1 & \alpha_{n-1}^2 & \alpha_{n-1}^3 & \ldots & \alpha_{n-1}^{n-1} & 1 \end{pmatrix}.$$

Легко видеть, что определитель этой матрицы $\det A = 1$. Поэтому векторы

$$(\forall k = 1, 2, \ldots, n)\ \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$$

линейно выражаются через векторы $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k$. Следовательно, линейные оболочки $L\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k\right\}$ и $L\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right\}$ совпадают.

Подчиним коэффициенты α_k^i $(i = 1, 2, \ldots, k)$ условию, состоящему в том, что вектор $\vec{e}_{k+1}$ должен быть сопряжён к подпространству $L\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right\}$.

Вектор $\vec{e}_{k+1}$ сопряжён к подпространству $L\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k\right\}$ в том и только в том случае, если выполняются условия:

$$\mathrm{f}\left(\vec{e}_{k+1}, \vec{a}_1\right) = 0; \mathrm{f}\left(\vec{e}_{k+1}, \vec{a}_2\right) = 0; \ldots; \mathrm{f}\left(\vec{e}_{k+1}, \vec{a}_k\right) = 0. \qquad (6.3.18)$$

Действительно, из выражения (6.3.18) следует, что вектор $\vec{e}_{k+1}$ сопряжён с линейной оболочкой $L\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right\}$ (следствие из леммы 6.3.4), которая, как показано выше, совпадает с линейной оболочкой $L\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k\right\}$.

Обратно, если вектор $\vec{e}_{k+1}$ сопряжён с линейной оболочкой $L\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k\right\}$, то он сопряжён и с линейной оболочкой (подпространством) $L\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right\}$, то есть он (следствие леммы 6.3.4) сопряжён с каждым из векторов $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$. Поэтому выполняются равенства (6.3.18).

Подставим выражение для вектора $\vec{e}_{k+1}$ из формул (6.3.17) в каждую из формул (6.3.18). Учитывая определение билинейной формы, получаем систему уравнений относительно величин $\alpha_k^i \ (i=1, 2, \ldots, k)$:

$$\begin{aligned}
&\mathrm{f}\left(\vec{e}_{k+1}, \vec{a}_1\right) = \alpha_k^1 \mathrm{f}\left(\vec{a}_1, \vec{a}_1\right) + \alpha_k^2 \mathrm{f}\left(\vec{a}_2, \vec{a}_1\right) + \ldots + \\
&\qquad + \alpha_k^k \mathrm{f}\left(\vec{a}_k, \vec{a}_1\right) + \mathrm{f}\left(\vec{a}_{k+1}, \vec{a}_1\right) = 0, \\
&\mathrm{f}\left(\vec{e}_{k+1}, \vec{a}_2\right) = \alpha_k^1 \mathrm{f}\left(\vec{a}_1, \vec{a}_2\right) + \alpha_k^2 \mathrm{f}\left(\vec{a}_2, \vec{a}_2\right) + \ldots + \\
&\qquad + \alpha_k^k \mathrm{f}\left(\vec{a}_k, \vec{a}_2\right) + \mathrm{f}\left(\vec{a}_{k+1}, \vec{a}_2\right) = 0, \\
&\mathrm{f}\left(\vec{e}_{k+1}, \vec{a}_k\right) = \alpha_k^1 \mathrm{f}\left(\vec{a}_1, \vec{a}_k\right) + \alpha_k^2 \mathrm{f}\left(\vec{a}_2, \vec{a}_k\right) + \ldots + \\
&\qquad + \alpha_k^k \mathrm{f}\left(\vec{a}_k, \vec{a}_k\right) + \mathrm{f}\left(\vec{a}_{k+1}, \vec{a}_k\right) = 0.
\end{aligned} \tag{6.3.19}$$

Система (6.3.19) — это неоднородная СЛАУ относительно α_k^i с коэффициентами

$$\mathrm{f}\left(\vec{a}_i, \vec{a}_j\right) = \mathrm{f}_{ij} \left(i, j = 1, 2, \ldots, k\right),$$

определитель которой по условию не равен нулю. Поэтому СЛАУ (6.3.19) совместна и определённа. Решая СЛАУ (6.3.19), можем найти коэффициенты $\alpha_k^i \ (i=1, 2, \ldots, k)$ и, следовательно, построить вектор $\vec{e}_{k+1}$ по формулам (6.3.17).

Для определения всех коэффициентов $\alpha_k^i \ (i=1, 2, \ldots, k)$ и всех векторов базиса $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ нужно при каждом $k=1, 2, \ldots, n$ решить СЛАУ порядка $n-1$.

Рассмотрим теперь вычисление коэффициентов μ_i в выражении билинейной формы

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n} \mu_i x^i y^i.$$

Пусть в построенном базисе $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset X^n$

$$\vec{x} = \xi^1 \vec{e}_1 + \xi^2 \vec{e}_2 + \ldots + \xi^n \vec{e}_n,$$

$$\vec{y} = \eta^1 \vec{e}_1 + \eta^2 \vec{e}_2 + \ldots + \eta^n \vec{e}_n$$

билинейная форма в базисе $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ принимает вид

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n} \mu_{i'} \xi^i \eta^i. \tag{6.3.20}$$

Чтобы определить коэффициенты $\mu_{i'}$, рассмотрим билинейную форму в подпространстве $L^m\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ при $m \le n$.

Матрица формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ в базисе $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\} \subset L_m$ имеет вид

$$\Phi = \begin{pmatrix} \mathrm{f}_{11} & \mathrm{f}_{12} & \ldots & \mathrm{f}_{1m} \\ \mathrm{f}_{21} & \mathrm{f}_{22} & \ldots & \mathrm{f}_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \mathrm{f}_{m1} & \mathrm{f}_{m2} & \ldots & \mathrm{f}_{mm} \end{pmatrix},$$

а в базисе $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ (для простоты опустим штрихи у μ_i)

$$\Phi' = \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu_m \end{pmatrix}.$$

Переход от базиса $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ к базису $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ осуществляется по формулам (6.3.17) с матрицей, определитель которой равен 1. Поэтому по формуле (6.3.8) получаем:

$$\det \Phi' = \det \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu_m \end{pmatrix} = \det\left(A^T \Phi A\right) =$$

$$= \det A^T \det \Phi \det A = \det \Phi = \det \begin{pmatrix} \mathrm{f}_{11} & \mathrm{f}_{12} & \ldots & \mathrm{f}_{1m} \\ \mathrm{f}_{21} & \mathrm{f}_{22} & \ldots & \mathrm{f}_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \mathrm{f}_{m1} & \mathrm{f}_{m2} & \ldots & \mathrm{f}_{mm} \end{pmatrix}.$$

Используя обозначения главных миноров, перепишем эту формулу так:

$$M_m = \mu_1\mu_2\ldots\mu_m\,(m=1,2,\ldots,n). \tag{6.3.21}$$

Из формулы (6.3.21) следуют рекуррентные формулы для коэффициентов μ_i:

$$\mu_1 = M_1 = \mathrm{f}_{11};\ \mu_2 = \frac{M_2}{M_1},\ \mu_3 = \frac{M_3}{M_2},\ldots,\mu_n = \frac{M_n}{M_{n-1}}. \tag{6.3.22}$$

Формулы (6.3.22) позволяют найти канонические коэффициенты билинейной формы, не находя канонического базиса.

Из формул (6.3.17) выберем формулу с некоторым номером k и переписав её в виде

$$\vec{a}_{k+1} = -\alpha_k^1\vec{a}_1 - \alpha_k^2\vec{a}_2 - \ldots - \alpha_k^k\vec{a}_k + \vec{e}_{k+1} = \vec{g}_k + \vec{e}_{k+1}.$$

Здесь $\vec{g}_k \in L\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_k\right\}$, а вектор $\vec{e}_{k+1}$ сопряжён к этому подпространству. Коэффициенты $\alpha_k^i\ (i=1,2,\ldots,k)$ единственным образом определяются из решения СЛАУ (6.3.19) при условии, что

$$\det\left(\mathrm{f}_{ij}\right) \neq 0\,(i,j=1,2,\ldots,k),$$

то есть при условии невырожденности формы $\phi\left(\vec{x},\vec{y}\right)$ на подпространстве $L\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_k\right\}$. В силу произвольности вектора $\vec{e}_{k+1}$, обозначая

$$\vec{e}_{k+1} = \vec{h},\ \vec{g}_k = \vec{g},\ \vec{a}_{k+1} = \vec{a},\ L\left\{\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_k\right\} = L \subset X^n,$$

приходим к следующей теореме.

Теорема 6.3.3. *Пусть билинейная форма $\phi\left(\vec{x},\vec{y}\right)$, определённая на векторах пространства X^n, является невырожденной на подпространстве $L\subset X^n$ и вектор $\vec{a}\notin L$. Тогда существует единственное разложение вида*

$$\vec{a} = \vec{g} + \vec{h}, \tag{6.3.23}$$

где $\vec{g}\in L$, а вектор $\vec{h}$ сопряжён с подпространством L.

Теорема 6.3.3 и формула (6.3.23) являются широким обобщением теоремы 3.1.2 и формулы (3.1.5) из параграфа 3.1.

Билинейные и квадратичные формы в вещественном пространстве. Поле вещественных чисел упорядочено — для вещественных чисел определены знаки (+) или (–). Поэтому в случае билинейных и квадратичных форм, определённых в вещественном пространстве, теория может быть значительно развита по сравнению со случаем пространств, определённых над произвольным полем *P*.

По теореме 6.3.1 в пространстве X^n существует базис, в котором квадратичная форма $\varphi\left(\vec{x}, \vec{x}\right)$ имеет канонический вид

$$\varphi\left(\vec{x}, \vec{x}\right) = \mu_1 \left(x^1\right)^2 + \mu_2 \left(x^2\right)^2 + \ldots + \mu_n \left(x^n\right)^2,$$

где штрихи при координатах опущены для простоты записи. Из предыдущего изложения следует, что число отличных от нуля канонических коэффициентов μ_k $(k = 1, 2, \ldots, n)$ равно рангу формы. Эти коэффициенты могут быть как положительными, так и отрицательными.

Теорема 6.3.4 (закон инерции квадратичных форм). *Число положительных и число отрицательных коэффициентов в каноническом виде квадратичной формы* $\varphi\left(\vec{x}, \vec{x}\right)$ *не зависят от выбора канонического базиса.*

Доказательство повторяет доказательство теоремы 3.2.1.

Полное число членов, входящих в канонический вид квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$, равное её рангу, называется ***индексом инерции*** квадратичной формы, а числа положительных и отрицательных членов называются, соответственно, ***положительным индексом инерции*** и ***отрицательным индексом инерции*** квадратичной формы. Выше результат, аналогичный теореме 6.3.4, был получен в параграфе 3.2 (теорема 3.2.6).

Определение 6.3.7. *Квадратичная форма* $\varphi\left(\vec{x}, \vec{x}\right)$ *называется* ***положительно определённой****, если все n её канонических коэффициентов положительны, то есть если выполняется условие*

$$\left(\forall \vec{x} \in X^n\right) : \vec{x} \neq \vec{0} \;\; \varphi\left(\vec{x}, \vec{x}\right) > 0,$$

и ***отрицательно определённой****, если все n её канонических коэффициентов отрицательны, то есть если выполняется условие*

$$\left(\forall \vec{x} \in X^n\right) : \vec{x} \neq \vec{0}\ \varphi\left(\vec{x}, \vec{x}\right) < 0. \bullet$$

Теорема 6.3.5. *Квадратичная форма* $\varphi\left(\vec{x}, \vec{x}\right)$ *является положительно определённой в том и только в том случае, если её положительный индекс инерции равен размерности пространства* X^n.

Д о к а з а т е л ь с т в о. Доказательство этой теоремы является прямым следствием определения 6.3.7. ••

Закон инерции (теорема 6.3.4) непосредственно переносится на симметричные билинейные формы.

Теорема 6.3.6. *Число положительных и число отрицательных коэффициентов в каноническом виде*

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n} \mu_i x^i y^i$$

симметричной билинейной формы не зависят от выбора базиса.

По этой причине для симметричных билинейных форм, так же как для квадратичных форм, вводятся понятия положительного и отрицательного индексов инерции. Значения положительного и отрицательного индексов инерции симметричной билинейной формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ и, следовательно, квадратичной формы $\varphi\left(\vec{x}, \vec{x}\right)$, определяются по знакам угловых миноров матрицы формы в каком-либо базисе. Из формул (6.3.22) вытекает ***критерий Сильвестра положительной определённости квадратичной формы***.

Теорема 6.3.7. *Квадратичная форма, определённая на векторах вещественного пространства* X^n, *является положительно определённой в том и только в том случае, если в каком-либо базисе* $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset X^n$ *все угловые миноры её матрицы* $\varphi\left(\vec{x}, \vec{x}\right)$ *положительны, то есть*

$$M_1 = \varphi_{11} > 0,\ M_2 = \begin{vmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{vmatrix} > 0, \ldots,\ M_n = \begin{vmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{vmatrix} > 0.$$

Доказательство дословно повторяет доказательство теоремы 3.2.3, если в последнем заменить R^n на X^n.

Следствие из теоремы 6.3.7. *Квадратичная форма является отрицательно определённой в том и только в том случае, если угловые миноры её матрицы чётного порядка положительны, а нечётного порядка — отрицательны.*

Билинейные и квадратичные формы в евклидовом пространстве. Пусть E^n — вещественное евклидово пространство и $\hat{T}: E^n \to \hat{T}\left(E^n\right)$ — некоторый ортогональный оператор, действующий в пространстве E^n. Как известно, матрица ортогонального оператора удовлетворяет условию $T^T = T^{-1}$. Докажем сначала вспомогательное утверждение.

Лемма 6.3.5. *В вещественном евклидовом пространстве E^n матрица*

$$A = \left(A_{i'}^{j}\right)$$

перехода от ортонормированного базиса $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ к ортонормированному базису $\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\}$ является ортогональной.

Д о к а з а т е л ь с т в о. По условию леммы имеем:

$$\begin{aligned} \vec{e}_{1'} &= A_{1'}^{1}\vec{e}_1 + A_{1'}^{2}\vec{e}_2 + \ldots + A_{1'}^{n}\vec{e}_n, \\ \vec{e}_{2'} &= A_{2'}^{1}\vec{e}_1 + A_{2'}^{2}\vec{e}_2 + \ldots + A_{2'}^{n}\vec{e}_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \vec{e}_{n'} &= A_{n'}^{1}\vec{e}_1 + A_{n'}^{2}\vec{e}_2 + \ldots + A_{n'}^{n}\vec{e}_n. \end{aligned} \tag{6.3.24}$$

Пусть оператор $\hat{A}$ в базисе $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ имеет матрицу A преобразования (6.3.24). Тогда

$$\begin{aligned} \hat{A}\vec{e}_1 &= A_{1'}^{1}\vec{e}_1 + A_{1'}^{2}\vec{e}_2 + \ldots + A_{1'}^{n}\vec{e}_n = \vec{e}_{1'}, \\ \hat{A}\vec{e}_2 &= A_{2'}^{1}\vec{e}_1 + A_{2'}^{2}\vec{e}_2 + \ldots + A_{2'}^{n}\vec{e}_n = \vec{e}_{2'}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \hat{A}\vec{e}_n &= A_{n'}^{1}\vec{e}_1 + A_{n'}^{2}\vec{e}_2 + \ldots + A_{n'}^{n}\vec{e}_n = \vec{e}_{n'}. \end{aligned}$$

Оператор $\hat{A}$ является ортогональным в том и только в том случае, если он переводит хотя бы один ортонормированный базис в ортонормированный базис. Поэтому матрица $A = \left(A_{i'}^{j}\right)$ ортогональная. ••

Пусть в вещественном евклидовом пространстве E^n зафиксирован ортонормированный базис $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, в котором задана билинейная форма

$$\mathrm{f}\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{f}_{ij} x^i y^j,$$

где

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{e}_i, \ \vec{y} = \sum_{j=1}^{n} y^j \vec{e}_j.$$

Определим линейный оператор $\hat{T} : E^n \to \hat{T}\left(E^n\right)$ с матрицей, равной матрице формы $T = \Phi$, который будем называть далее ***оператором, ассоциированным с формой*** $\mathrm{f}\left(\vec{x}, \vec{y}\right)$. При переходе к новому базису $\left\{\vec{e}_{1'}, \vec{e}_{2'}, \ldots, \vec{e}_{n'}\right\}$ с матрицей перехода $A = \left(A_{i'}^{j}\right)$ матрица билинейной формы $\mathrm{f}\left(\vec{x}, \vec{y}\right)$ преобразуется по формуле

$$\Phi' = A^T \Phi A,$$

а матрица оператора по формуле

$$T' = \left(A^{-1}\right)^T T A^T.$$

Таким образом, матрицы билинейной формы и оператора преобразуются, вообще говоря, неодинаково. Если новый базис также ортонормированный, то матрица перехода ортогональная, то есть $A^T = A^{-1}$. Учитывая симметричность матрицы билинейной формы и $(T = \Phi)$ оператора $\hat{T}$, получаем

$$T' = \left(A^{-1}\right)^T T A^T = \left(A^T\right)^T T A^T = A T A^T = A^T T A$$

и, следовательно, матрицы билинейной формы и оператора преобразуются по одному закону.

Итак, получаем такой результат.

Лемма 6.3.6. *В вещественном евклидовом пространстве каждой билинейной форме (билинейному функционалу) соответствует вполне определённый линейный оператор, имеющий в каждом ортонормированном базисе матрицу, совпадающую с матрицей билинейной формы.*

Если $f\left(\vec{x}, \vec{y}\right)$ — симметричная билинейная форма, то ассоциированный оператор $\hat{T}$ будет самосопряжённым. Известно, что матрица самосопряжённого оператора в ортонормированном базисе, составленном из собственных векторов оператора $\hat{T}$, приводится к диагональному виду, причём на главной диагонали расположены собственные значения оператора $\mu_1, \mu_2, \ldots, \mu_n$. Если

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{e}_i, \ \vec{y} = \sum_{j=1}^{n} y^j \vec{e}_j,$$

то билинейная форма приводится к каноническому виду

$$f\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n} \mu_i x^i y^i,$$

а соответствующая квадратичная форма к каноническому виду

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \mu_i \left(x^i\right)^2.$$

Из общего выражения квадратичной формы

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \phi_{ij} x^i x^j$$

следует, что её можно записать в виде

$$\varphi\left(\vec{x}, \vec{x}\right) = \left(\hat{T}\vec{x}, \vec{x}\right),$$

где $\hat{T}$ — самосопряжённый оператор с матрицей, равной матрице квадратичной формы (в параграфе 3.2 это представление было положено в основу определения квадратичной формы). Напомним, что самосопряжённый оператор называется положительно определённым, если все его собственные значения положительны.

Имеет место почти очевидная теорема.

Теорема 6.3.8. *Самосопряжённый оператор* $\hat{T}: E^n \to \hat{T}\left(E^n\right)$ *является положительно определённым в том и только в том случае, если соответствующая квадратичная форма*

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i x^j$$

является положительно определённой, то есть

$$\left(\forall \vec{x} \in E^n\right)\varphi\left(\vec{x}, \vec{x}\right) = \left(\hat{T}\vec{x}, \vec{x}\right) > 0.$$

Последовательность действий, которые нужно произвести для приведения квадратичной формы

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}x^i x^j$$

к каноническому виду

$$\varphi\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\mu_i\left(x^i\right)^2,$$

описана в параграфе 5.2.

6.4. Жорданова форма матрицы линейного оператора*)

В этом параграфе рассмотрим приведение матрицы линейного оператора общего вида к ***жордановой канонической форме***, которая является в определённом смысле слова простейшей. Пространство, на котором действует оператор, считается комплексным векторным пространством.

Корневые векторы и корневые подпространства линейного оператора. Рассмотрим оператор $\hat{T}: X^n \to X^n$.

Определение 6.4.1. *Вектор* $\vec{q} \in X^n$ *называется* ***корневым вектором оператора*** $\hat{T}: X^n \to X^n$, *соответствующим числу* $\mu \in C$, *если для некоторого неотрицательного целого числа m справедливо условие*

$$\left(\hat{T} - \mu\hat{I}\right)^m \vec{q} = \vec{0}. \qquad (6.4.1)$$

Наименьшее из чисел

$$m = ht\vec{q},$$

*для которого выполняется условие (6.4.1), называется **высотой** корневого вектора* $\vec{q}$. •

Например, собственные векторы оператора $\hat{T}: X^n \to X^n$, удовлетворяющие уравнению

$$\left(\hat{T} - \mu\hat{I}\right)\vec{q} = \vec{0},$$

являются корневыми векторами высоты 1. По определению ***нулевой вектор*** считают ***корневым вектором высоты 0***, соответствующим любому числу μ.

Характеристический многочлен оператора $\hat{T}: X^n \to X^n$ можно представить в виде разложения

$$T(\mu) = (\mu - \mu_1)^{k_1}(\mu - \mu_2)^{k_2}\dots(\mu - \mu_p)^{k_p}, \tag{6.4.2}$$

где $\mu_1, \mu_2, \dots, \mu_p$ — попарно различные собственные значения, а кратности собственных значений удовлетворяют очевидному условию

$$k_1 + k_2 + \dots + k_p = n.$$

Рассмотрим многочлены

$$(\mu - \mu_1)^{k_1}, (\mu - \mu_2)^{k_2}, \dots, (\mu - \mu_p)^{k_p}.$$

Эти многочлены, в соответствии с разложением (6.4.2), являются делителями характеристического многочлена $T(\mu)$, причём никакая пара из них не имеет общих корней. По следствию из теоремы 4.1.7 существуют инвариантные подпространства

$$L_{\mu_1}\left(\hat{T}\right), L_{\mu_2}\left(\hat{T}\right), \dots, L_{\mu_p}\left(\hat{T}\right)$$

оператора $\hat{T}$, такие, что справедливо разложение

$$X^n = L_{\mu_1}\left(\hat{T}\right) \oplus L_{\mu_2}\left(\hat{T}\right) \oplus \dots \oplus L_{\mu_p}\left(\hat{T}\right).$$

Определение 6.4.2. *Подпространства* $L_{\mu_1}\left(\hat{T}\right), L_{\mu_2}\left(\hat{T}\right), \ldots, L_{\mu_p}\left(\hat{T}\right)$, *удовлетворяющие приведённому выше условию, называются* ***корневыми подпространствами*** *линейного оператора* $\hat{T}: X^n \to X^n$, *соответствующими корням* $\mu_1, \mu_2, \ldots, \mu_p$ *характеристического многочлена.* •

Лемма 6.4.1. *Размерность корневого подпространства равна кратности корня характеристического многочлена, которому оно соответствует.*

Д о к а з а т е л ь с т в о полностью аналогично доказательству следствия из теоремы 4.1.7. ••

Очевидно, что векторы $\vec{q}_{k_j}$ $(j=1, 2, \ldots, p)$, соответствующие корням μ_j $(j=1, 2, \ldots, p)$ характеристического многочлена (6.4.2) оператора $\hat{T}: X^n \to X^n$ (с учётом их кратности) и удовлетворяющие уравнениям

$$\left(\hat{T}-\mu_j \hat{I}\right)^{k_j} \vec{q}_{k_j} = \vec{0}, \tag{6.4.3}$$

являются корневыми векторами оператора $\hat{T}$.

Множество корневых векторов $\vec{q}$ высоты $ht\,\vec{q} \le m$ является ядром оператора $\left(\hat{T}-\mu \hat{I}\right)^m$. Действительно, если

$$\left(\hat{T}-\mu \hat{I}\right)\vec{q} = \vec{0},$$

то

$$\left(\hat{T}-\mu \hat{I}\right)^2 \vec{q} = \left(\hat{T}-\mu \hat{I}\right)\left[\left(\hat{T}-\mu \hat{I}\right)\vec{q}\right] = \vec{0},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\left(\hat{T}-\mu \hat{I}\right)^m \vec{q} = \left(\hat{T}-\mu \hat{I}\right)^{m-1}\left[\left(\hat{T}-\mu \hat{I}\right)\vec{q}\right] = \vec{0}.$$

Таким образом, получаем флаг инвариантных подпространств оператора $\hat{T}$

$$K\left(\hat{T}-\mu \hat{I}\right) \subset K\left(\hat{T}-\mu \hat{I}\right)^2 \subset K\left(\hat{T}-\mu \hat{I}\right)^3 \subset \ldots,$$

так как ядро операторного многочлена $\left(\hat{T}-\mu\hat{I}\right)^m$ является инвариантным подпространством оператора $\hat{T}$.

Из приведённого выше рассмотрения и теорем вытекает следующее утверждение.

Теорема 6.4.1. *Любой оператор* $\hat{T}: X^n \to X^n$ *можно разложить в прямую сумму операторов, индуцированных на корневых подпространствах.*

Составляя базис пространства X^n как объединение базисов корневых подпространств, запишем матрицу оператора $\hat{T}$ в ***квазидиагональном*** виде

$$T=\begin{pmatrix} T_1^1 & O & \dots & O \\ O & T_2^2 & \dots & O \\ \dots & \dots & \dots & \dots \\ O & O & \dots & T_p^p \end{pmatrix}. \tag{6.4.4}$$

Каждая клетка T_j^j имеет порядок k_j и является матрицей оператора, индуцированного на корневом подпространстве L_{μ_j}. Далее следует выбрать базис в каждом корневом подпространстве.

Нильпотентные операторы и циклические подпространства. Рассмотрим действие оператора $\hat{T}$ на каждом из корневых подпространств более подробно. Начнём с оператора специального вида.

Определение 6.4.3. *Линейный оператор* $\hat{N}$ *называется* ***нильпотентным****, если*

$$\left(\exists m\in Z^+\right):\hat{N}^m=\hat{O}.$$

Наименьшее из чисел m, для которого выполняется указанное условие, называется высотой оператора $\hat{N}$. •

Например, оператор дифференцирования $\hat{D}=\dfrac{d}{dx}$ в пространстве многочленов $a_0+a_1x+\ldots+a_nx^n$ степени не выше *n* является нильпотентным оператором высоты $n+1$.

Так как для некоторого $m\in Z^+$ корневое подпространство

$$L_\mu\left(\hat{T}\right)=K\left(\left(\hat{T}-\mu\hat{I}\right)^m\right),$$

то

$$\left(\forall \vec{q} \in L_{\mu}\right)\left(\hat{T}-\mu \hat{I}\right)^{m} \vec{q}=\vec{0}$$

и сужение

$$\hat{N}=\left(\hat{T}-\mu \hat{I}\right)^{m} \Big| L_{\mu}$$

является нильпотентным оператором. Поэтому сначала будем изучать нильпотентные операторы.

Определение 6.4.4. ***Высотой*** *вектора* $\vec{q} \in X^{n}$ *относительно нильпотентного оператора* $\hat{N}$ *называется наименьшее число* $m \in Z^{+}$*, для которого*

$$\hat{N}^{m} \vec{q}=\vec{0},$$

то есть высота вектора $\vec{q}$ *как корневого вектора оператора* $\hat{N}$*, соответствующего корню* $\mu=0$. •

Лемма 6.4.2. *Если* $\vec{q} \in X^{n}$ *— вектор высоты m относительно нильпотентного оператора* $\hat{N}$*, то система векторов*

$$\left\{\vec{q}, \hat{N} \vec{q}, \hat{N}^{2} \vec{q}, \ldots, \hat{N}^{m-1} \vec{q}\right\}$$

линейно независима.

Доказательство. Предположим, что система векторов

$$\left\{\vec{q}, \hat{N} \vec{q}, \hat{N}^{2} \vec{q}, \ldots, \hat{N}^{m-1} \vec{q}\right\}$$

линейно зависима. Тогда найдутся такие числа $\alpha_{0}, \alpha_{1}, \alpha_{2}, \ldots, \alpha_{m-1}$, не равные нулю одновременно, что выполняется основное равенство

$$\alpha_{0} \vec{q}+\alpha_{1} \hat{N} \vec{q}+\alpha_{2} \hat{N}^{2} \vec{q}+\ldots+\alpha_{m-1} \hat{N}^{m-1} \vec{q}=\vec{0}.$$

Пусть, например, первым отличным от нуля коэффициентом является $\alpha_{k} \neq 0$. Действуя на обе части этого равенства оператором $\hat{N}^{m-(k+1)}$, получаем (первые $k-1$ слагаемых равны $\vec{0}$):

$$\alpha_k \hat{N}^{m-(k+1)} \hat{N}^{k} \vec{q} + \alpha_{k+1} \hat{N}^{m-(k+1)} \hat{N}^{k+1} \vec{q} +$$
$$+ \ldots + \alpha_{m-1} \hat{N}^{m-(k+1)} \hat{N}^{m-1} \vec{q} = \vec{0},$$

или

$$\alpha_k \hat{N}^{m-1} \vec{q} + \alpha_{k+1} \hat{N}^{m} \vec{q} + \ldots + \alpha_{m-1} \hat{N}^{2m-(k+2)} \vec{q} = \vec{0}.$$

Так как все слагаемые с коэффициентами $\alpha_{k+1}, \alpha_{k+2}, \alpha_{k+3}, \ldots, \alpha_{m\text{-}1}$ равны $\vec{0}$ по определению 6.4.4 высоты вектора, то

$$\alpha_k \hat{N}^{m-1} \vec{q} = \vec{0},$$

откуда в силу $\alpha_k \neq 0$ следует, что

$$\hat{N}^{m-1} \vec{q} = \vec{0}.$$

Но вектор $\vec{q}$ по предположению имеет высоту m. Получили противоречие. ••

Определение 6.4.5. *Линейная оболочка (где m — высота вектора $\vec{q}$ относительно $\hat{N}$)*

$$L^m \left\{ \vec{q}, \hat{N} \vec{q}, \hat{N}^2 \vec{q}, \ldots, \hat{N}^{m-1} \vec{q} \right\}$$

называется ***циклическим подпространством нильпотентного оператора*** $\hat{N}$*, порождённым вектором* $\vec{q}$. •

Циклическое подпространство инвариантно относительно оператора $\hat{N}$, что проверяется непосредственной выкладкой:

$$\hat{N} \left(\alpha_0 \vec{q} + \alpha_1 \hat{N} \vec{q} + \ldots + \alpha_{m-1} \hat{N}^{m-1} \vec{q} \right) =$$
$$= \alpha_0 \hat{N} \vec{q} + \alpha_1 \hat{N}^2 \vec{q} + \ldots + \alpha_{m-2} \hat{N}^{m-1} \vec{q}.$$

Название «циклическое подпространство» объясняется следующим фактом: любой вектор, принадлежащий циклическому подпространству

$$L^m\left\{\vec{q}, \hat{N}\vec{q}, \hat{N}^2\vec{q}, \ldots, \hat{N}^{m-1}\vec{q}\right\},$$

но не принадлежащий подпространству

$$L^{m-1}\left\{\hat{N}\vec{q}, \hat{N}^2\vec{q}, \ldots, \hat{N}^{m-1}\vec{q}\right\},$$

имеет высоту m и, следовательно, порождает то же самое циклическое подпространство.

Заметим, что базис циклического подпространства принято записывать в обратном порядке, то есть совершив полную круговую перестановку:

$$\left\{\vec{a}_0, \vec{a}_1, \ldots, \vec{a}_{m-3}, \vec{a}_{m-2}, \vec{a}_{m-1}\right\} \stackrel{def}{=}$$

$$= \left\{\hat{N}^{m-1}\vec{q}, \hat{N}^{m-2}\vec{q}, \ldots, \hat{N}^2\vec{q}, \hat{N}\vec{q}, \vec{q}\right\}.$$

Жорданов базис и жордановы клетки. Сужение оператора $\hat{N}$ на циклическое подпространство имеет высоту m и в базисе

$$\left\{\vec{a}_0, \vec{a}_1, \ldots, \vec{a}_{m-3}, \vec{a}_{m-2}, \vec{a}_{m-1}\right\} =$$

$$= \left\{\hat{N}^{m-1}\vec{q}, \hat{N}^{m-2}\vec{q}, \ldots, \hat{N}^2\vec{q}, \hat{N}\vec{q}, \vec{q}\right\}$$

циклического подпространства представляется матрицей

$$N = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}, \qquad (6.4.5)$$

которая называется ***нильпотентной жордановой клеткой***. Действительно, действуя нильпотентным оператором на векторы базиса циклического подпространства, получаем:

$$\hat{N}\vec{a}_0 = \hat{N}\left(\hat{N}^{m-1}\vec{q}\right) = \hat{N}^m\vec{q} = \vec{0},$$

$$\hat{N}\vec{a}_1 = \hat{N}\left(\hat{N}^{m-2}\vec{q}\right) = \hat{N}^{m-1}\vec{q} = \vec{a}_0,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\hat{N}\vec{a}_{m-2} = \hat{N}\left(\hat{N}\vec{q}\right) = \hat{N}^{2}\vec{q} = \vec{a}_{m-3},$$

$$\hat{N}\vec{a}_{m-1} = \hat{N}\vec{q} = \vec{a}_{m-2}.$$

Так как оператор $\hat{N} = \hat{T} - \mu_j \hat{I}$ на каждом корневом подпространстве является нильпотентным оператором, то имеем:

$$\left(\hat{T} - \mu_j \hat{I}\right)\vec{a}_0 = \vec{0};$$

$$\left(\hat{T} - \mu_j \hat{I}\right)\vec{a}_1 = \vec{a}_0;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;$$

$$\left(\hat{T} - \mu_j \hat{I}\right)\vec{a}_{m-2} = \vec{a}_{m-3};$$

$$\left(\hat{T} - \mu_j \hat{I}\right)\vec{a}_{m-1} = \vec{a}_{m-2}.$$

Откуда получаем:

$$\hat{T}\vec{a}_0 = \mu_j \vec{a}_0;$$

$$\hat{T}\vec{a}_1 = \vec{a}_0 + \mu_j \vec{a}_1;$$

$$\hat{T}\vec{a}_2 = \vec{a}_1 + \mu_j \vec{a}_2;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;$$

$$\hat{T}\vec{a}_{m-2} = \vec{a}_{m-3} + \mu_j \vec{a}_{m-2};$$

$$\hat{T}\vec{a}_{m-1} = \vec{a}_{m-2} + \mu_j \vec{a}_{m-1}.$$

Теперь на каждом корневом подпространстве матрица индуцированного оператора имеет вид

$$J_k^k \equiv T_k^k = \begin{pmatrix} \mu_j & 1 & 0 & \dots & 0 & 0 \\ 0 & \mu_j & 1 & \dots & 0 & 0 \\ 0 & 0 & \mu_j & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mu_j & 1 \\ 0 & 0 & 0 & \dots & 0 & \mu_j \end{pmatrix}. \qquad (6.4.6)$$

Матрица (6.4.6) называется ***жордановой клеткой*** или ***жордановым ящиком***. Учитывая вид матрицы (6.4.4), получаем следующий ***канонический жорданов вид*** матрицы линейного оператора $\hat{T}$ в ***жордановом базисе***, составленном как объединение базисов циклических подпространств оператора $\hat{N} = \hat{T} - \mu_j \hat{I}$:

$$T_J = \begin{pmatrix} J_1^1 & O & \dots & O & O \\ O & J_2^2 & \dots & O & O \\ \dots & \dots & \dots & \dots & \dots \\ O & O & \dots & O & J_p^p \end{pmatrix}. \qquad (6.4.7)$$

Пример нахождения жорданова базиса пространства X^n и приведения матрицы оператора $\hat{T}: X^n \to X^n$ в этом базисе к каноническому виду приведён в параграфе, посвящённом решению типовых задач.

6.5. Метрика в евклидовых пространствах

Плоскости в n-мерном евклидовом пространстве. Под m-мерной плоскостью в n-мерном евклидовом пространстве мы понимаем тот же геометрический объект, что и в n-мерном аффинном пространстве, ***порождающем*** данное евклидово пространство [21, 38]. В порождающем аффинном пространстве плоскость является также аффинным пространством. На плоскости в евклидовом пространстве, однако, как и во всём вмещающем пространстве, определено скалярное произведение. Тем не менее, можно представить ситуацию, когда m-мерная плоскость в n-мерном евклидовом пространстве сама может не быть евклидовым пространством меньшего числа измерений [21, 38], а именно: ***скалярное произведение на этой плоскости может не удов-***

летворять условию невырожденности, в то время как во всём пространстве это условие удовлетворяется. Последнее означает, что *на данной плоскости, возможно, найдётся вектор* $\vec{x} \neq \vec{0}$*, такой, что для любого вектора*

$$\left(\forall \vec{y} \in H^m\right)\left(\vec{x}, \vec{y}\right) = 0.$$

Таким образом, вектор $\vec{x}$ *лежит в плоскости* H^m *и одновременно ортогонален ко всем её векторам. Плоскости, на которых реализуется описанная ситуация, называются* ***изотропными*** [21]. Расстояние или метрика

$$\rho\left(\vec{x}, \vec{y}\right) \overset{def}{=} \left\|\vec{y} - \vec{x}\right\| = \sqrt{g_{ij}\left(y^i - x^i\right)\left(y^j - x^j\right)}$$

на изотропной плоскости называется ***вырожденной***.

Изотропная плоскость не может считаться евклидовым пространством в силу невыполнения на ней входящего в определение последнего условия ***невырожденности***. Как правило, плоскости являются ***неизотропными***, а изотропные плоскости являются исключением. В частности, в обычном трёхмерном пространстве геометрических векторов изотропных плоскостей нет вовсе. По этой причине наглядное представление об изотропных плоскостях затруднено.

Определение 6.5.1. *Вещественное евклидово пространство с положительно определённым, то есть удовлетворяющим условию* $\left(\forall \vec{x} \neq \vec{0}\right)$ $(\vec{x}, \vec{x}) > 0$ *скалярным произведением, называется* ***собственно евклидовым пространством***. •

Теорема 6.5.1. *В собственно евклидовом пространстве не существует изотропных плоскостей.*

Доказательство. В собственно евклидовом пространстве на любой плоскости $H^m \subset E^n$ соблюдается условие $(\vec{x}, \vec{x}) > 0$ для любого $\vec{x} \neq \vec{0}$. Но это означает, что на плоскости невозможен вектор $\vec{x} \neq \vec{0}$, ортогональный ко всем векторам плоскости. Действительно, если бы такой вектор существовал, то он, будучи ортогонален ко всем векторам плоскости, в частности, был бы ортогонален к самому себе, то есть $(\vec{x}, \vec{x}) = 0$. Но это противоречит определению собственно евклидовых пространств. ••

Вектор $\vec{x} \neq \vec{0}$, для которого $(\vec{x}, \vec{x}) = 0$, то есть ортогональный самому себе, называется ***изотропным*** вектором.

Пусть плоскость $H^m \subset E^n$, заданная своим репером $\left\{O^*, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$, изотропная (метрика на плоскости вырождена), то есть существует вектор $\vec{x} \neq \vec{0}$, такой, что для любого вектора $\vec{y}$ всегда $\left(\vec{x}, \vec{y}\right) = 0$. Вспоминая выражение скалярного произведения

$$(\vec{x}, \vec{y}) = \sum_{i=1}^{m}\sum_{j=1}^{m} g_{ij}\, x^i y^j,$$

где $g_{ij} \equiv (\vec{a}_i, \vec{a}_j)$, видим, что из условия

$$(\vec{x}, \vec{y}) = \sum_{i=1}^{m}\sum_{j=1}^{m} g_{ij}\, x^i y^j = 0$$

следует, что коэффициенты при y^j должны обращаться в нуль, то есть

$$\sum_{i=1}^{m} g_{ij}\, x^i = 0.$$

Так как все x^i не могут быть равными нулю одновременно, то однородная СЛАУ

$$\sum_{i=1}^{m} g_{ij}\, x^i = 0$$

нетривиально совместна, а значит, $\det\left(g_{ij}\right) = 0$. Очевидно, что условие невырожденности эквивалентно условию $\det\left(g_{ij}\right) \neq 0$, в противном случае плоскость будет изотропной.

Пусть в пространстве E^n задан некоторый репер $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ и прямая линия с направляющим вектором

$$\vec{a} = a^1 \vec{e}_1 + a^2 \vec{e}_2 + \ldots + a^n \vec{e}_n \neq \vec{0}.$$

Рассмотрим множество векторов

$$\vec{x} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + \ldots + x^n \vec{e}_n,$$

ортогональных данной прямой линии, то есть таких, что

$$\left(\vec{a}, \vec{x}\right) = 0. \tag{6.5.1}$$

Вспоминая, что

$$\left(\vec{a}, \vec{x}\right) = \sum_{j=1}^{n}\sum_{i=1}^{n} g_{ij}\, a^i x^j$$

и вводя обозначение

$$\sum_{i=1}^{n} g_{ij}\, a^i \overset{def}{=} a_j,$$

получаем уравнение

$$\sum_{j=1}^{n} a_j x^j = a_1 x^1 + a_2 x^2 + \ldots + a_n x^n = 0. \qquad (6.5.2)$$

Если все векторы $\vec{x}$ отложить от начала репера $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$, то нетрудно видеть, что *концы векторов $\vec{x}$ образуют гиперплоскость с уравнением (6.5.2), проходящую через начало O, причём все векторы этой гиперплоскости ортогональны векторам данной прямой. Эта гиперплоскость называется* ***ортогональной*** *данной прямой.* ••

Теорема 6.5.2. *В пространстве E^n гиперплоскость, ортогональная данной прямой линии, неизотропна и не содержит данной прямой (даже если у них есть общая точка) в том и только в том случае, если данная прямая сама неизотропна.*

Доказательство. Пусть данная прямая неизотропная. Тогда она не лежит в ортогональной к ней гиперплоскости, так как иначе направляющий вектор прямой

$$\vec{a} = \sum_{i=1}^{n} a^i \vec{e}_i \neq \vec{0}$$

должен быть ортогонален самому себе, а это невозможно в силу условия неизотропности прямой $\left(\vec{a}, \vec{a}\right) \equiv \vec{a}^{\,2} \neq 0$. Но тогда система $\left\{\vec{a}, \vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{n-1}\right\}$, образованная направляющим вектором прямой и направляющими векторами гиперплоскости, линейно независима и все эти векторы в совокупности можно принять за базисные векторы репера $\left\{\vec{a}, \vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{n-1}\right\}$ пространства E^n. Гиперплоскость тоже неизотропная, так как в противном случае можно найти вектор $\vec{x} \in H^{n-1}$, $\vec{x} \neq \vec{0}$, ортогональный ко всем векторам гиперплоскости, в частности, к её направляющим векторам $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{n-1}$. Кроме этого, вектор $\vec{x}$ будет ортогонален и к вектору $\vec{a} = \sum_{i=1}^{n} a^i \vec{e}_i$, так как гиперплоскость ортогональна прямой. Следовательно, вектор $\vec{x}$ будет ортогонален ко всем векторам репера и тем самым ко всем векторам пространства E^n, что противоречит условию невырожденности.

Пусть прямая изотропна. Тогда её направляющий вектор в силу условия $\left(\vec{a}, \vec{a}\right) = 0$ принадлежит множеству векторов $\vec{x}$, ортогональных к нему, и поэтому лежит в ортогональной гиперплоскости. Поэтому если изотропная прямая и ортогональная к ней гиперплоскость проходят через общую точку O, то вместе с направляющим вектором и вся прямая лежит в гиперплоскости. А так как гиперплоскость содержит вектор $\vec{a}$, ортогональный ко всем её векторам, то она также является изотропной. ••

Результат теоремы 6.5.2 противоречит нашим повседневным представлениям о геометрии. Таким образом, обычные представления в многомерной геометрии часто непригодны. Эти различия относятся, однако, лишь к метрическим свойствам многомерного евклидова пространства. Аффинные свойства остаются неизменными, так как все типы евклидовых пространств конструируются на базе одного и того же вещественного (или комплексного) аффинного пространства.

Рассмотрим аналогичные построения для плоскостей.

Определение 6.5.2. *Две плоскости в пространстве E^n называются* ***ортогональными****, если каждый вектор одной из них ортогонален каждому вектору другой.* •

Теорема 6.5.3. *В пространстве E^n через начало репера $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ проходит одна и только одна плоскость $n-m$ измерений, ортогональная к данной m-мерной плоскости.*

Д о к а з а т е л ь с т в о. Пусть дана плоскость $H^m \subset E^n$ с направляющими векторами $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$. Рассмотрим совокупность векторов $\vec{x}$, ортогональных данной плоскости. Тогда векторы $\vec{x}$ должны быть ортогональны к направляющим векторам $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$, то есть для любого $k = 1, 2, \ldots, m$ выполняется условие

$$(\vec{a}_k, \vec{x}) = 0. \tag{6.5.3}$$

Перепишем соотношения (6.5.3) аналогично соотношению (6.5.1). Преобразуем сначала скалярные произведения в левой части (6.5.3):

$$(\vec{a}_k, \vec{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_k^i x^j \left(\vec{e}_i, \vec{e}_j\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} a_k^i x^j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} g_{ij} a_k^i\right) x^j = \sum_{j=1}^{n} a_{kj} x^j. \tag{6.5.4}$$

Здесь введено обозначение

$$a_{kj} = \sum_{i=1}^{n} g_{ij} a_k^i.$$

Теперь вместо (6.5.3) имеем СЛАУ вида

$$\sum_{j=1}^{n} a_{kj} x^{j} = 0. \tag{6.5.5}$$

Учитывая, что $k = 1, 2, \ldots, m$, перепишем суммы (6.5.5) в развёрнутом виде

$$\begin{cases} a_{11}x^1 + a_{12}x^2 + \ldots + a_{1n}x^n = 0, \\ a_{21}x^1 + a_{22}x^2 + \ldots + a_{2n}x^n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x^1 + a_{m2}x^2 + \ldots + a_{mn}x^n = 0. \end{cases} \tag{6.5.6}$$

Уравнения СЛАУ (6.5.6) в совокупности линейно независимы, так как в противном случае линейно зависимой была бы система направляющих векторов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$, что невозможно. Отложив векторы $\vec{x}$ от начала O, видим, что концы векторов $\vec{x}$ образуют $n-m$-мерную плоскость, ортогональную исходной плоскости. Действительно, уравнения (6.5.6) можно записать в виде, разрешённом относительно m неизвестных, а такая система уравнений определяет плоскость $n-m$ измерений. ••

Теорема 6.5.4. *Пусть $H^m \subset E^n$ — некоторая m-мерная плоскость в n-мерном евклидовом пространстве. Тогда плоскость H^{n-m}, ортогональная плоскости H^m, неизотропна в том и только в том случае, если данная плоскость неизотропна. Причём если обе плоскости H^m и H^{n-m} неизотропны, то они не пересекаются и их направляющие векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ и $\vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n$ можно в совокупности принять за векторы репера пространства E^n.*

Д о к а з а т е л ь с т в о. Пусть плоскости H^m и H^{n-m} не пересекаются. Тогда система

$$\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m, \vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n\right\}$$

направляющих векторов линейно независима. Действительно, если система линейно зависима, то можно подобрать такие числа $\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n$, не равные нулю одновременно, что

$$\alpha_1 \vec{a} + \alpha_2 \vec{a}_2 + \ldots + \alpha_m \vec{a}_m + \alpha_{m+1} \vec{a}_{m+1} + \alpha_{m+2} \vec{a}_{m+2} + \ldots + \alpha_n \vec{a}_n = \vec{0},$$

откуда следует существование вектора

$$\alpha_1 \vec{a} + \alpha_2 \vec{a}_2 + \ldots + \alpha_m \vec{a}_m =$$

$$= -\alpha_{m+1} \vec{a}_{m+1} - \alpha_{m+2} \vec{a}_{m+2} - \ldots - \alpha_n \vec{a}_n,$$

общего для обеих плоскостей. Последнее, в свою очередь, влечёт существование общей прямой, проходящей через общую точку O и имеющей этот общий вектор направляющим вектором. А это противоречит предположению о непересечении плоскостей. Поэтому система векторов $\left\{\vec{a}_1, \ldots, \vec{a}_m, \vec{a}_{m+1}, \ldots, \vec{a}_n\right\}$ линейно независима и из её векторов можно составить базис репера

$$\left\{O, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m, \vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n\right\} \subset E^n.$$

Из аксиомы невырожденности следует, что определитель матрицы метрических коэффициентов

$$\det\left(g_{ij}\right) \equiv \det\left(\vec{e}_i, \vec{e}_j\right) \neq 0. \tag{6.5.7}$$

В рассматриваемом случае имеем:

$$\vec{e}_1 = \vec{a}_1; \ldots; \vec{e}_m = \vec{a}_m; \vec{e}_{m+1} = \vec{a}_{m+1}; \ldots; \vec{e}_n = \vec{a}_n. \tag{6.5.8}$$

В равенствах (6.5.8) векторы из наборов $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ и $\left\{\vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n\right\}$ попарно ортогональны друг другу, так как принадлежат ортогональным плоскостям. Следовательно, матрица метрических коэффициентов (6.5.7) имеет вид

$$G = \begin{pmatrix} \left(\vec{a}_p, \vec{a}_q\right) & O \\ O & \left(\vec{a}_r, \vec{a}_s\right) \end{pmatrix}, \tag{6.5.9}$$

где $p, q = 1, 2, \ldots, m;\ r, s = m+1, m+2, \ldots, n$. Из вида матрицы (6.5.9) следует, что

$$\det G = \det\left(g_{ij}\right) = \det\left(\vec{a}_p, \vec{a}_q\right)\det\left(\vec{a}_r, \vec{a}_s\right).$$

Теперь условие невырожденности принимает вид:

$$\det\left(\vec{a}_p, \vec{a}_q\right)\det\left(\vec{a}_r, \vec{a}_s\right) \neq 0. \tag{6.5.10}$$

Поэтому в произведении (6.5.10) отличен от нуля каждый сомножитель, что означает невырожденность метрики на каждой из плоскостей, а следовательно, их неизотропность.

Пусть плоскости H^m и H^{n-m} пересекаются. Тогда у них существует, по крайней мере, одна общая прямая и, следовательно, её направляющий вектор $\vec{c}$ также является общим для обеих плоскостей. Разложим его по векторам $\left\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right\}$ и по векторам $\left\{\vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n\right\}$, имеем:

$$\vec{c} = c^1\vec{a}_1 + c^2\vec{a}_2 + \ldots + c^m\vec{a}_m = c^{m+1}\vec{a}_{m+1} + c^{m+2}\vec{a}_{m+2} + \ldots + c^n\vec{a}_n.$$

Отсюда следует, что векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ и $\vec{a}_{m+1}, \vec{a}_{m+2}, \ldots, \vec{a}_n$ в совокупности образуют линейно зависимую систему и вектор $\vec{c}$ принадлежит как плоскости H^m, так и плоскости H^{n-m}. Так как $\vec{c} \in H^m$, то $\vec{c} \perp H^{n-m}$, а так как $\vec{c} \in H^{n-m}$, то $\vec{c} \perp H^m$. Получили, что вектор $\vec{c}$ принадлежит и одновременно ортогонален каждой из двух плоскостей H^m и H^{n-m}. Поэтому каждая из плоскостей изотропна. ••

Ортонормированный репер в собственно евклидовом пространстве; ортогонализация Шмидта. По теореме 6.5.1 в собственно евклидовом пространстве не существует изотропных плоскостей, а следовательно, и изотропных векторов. Поэтому собственно евклидовы пространства в определённом смысле слова устроены наиболее просто. В главе 2 мы, по существу, уже исследовали их. Основываясь на введённых в главе 2 понятиях, рассмотрим обоснование ***процедуры ортогонализации Шмидта*** для случая собственно евклидова пространства.

Найдём в пространстве E^n некоторый вектор $\vec{x}_1 \neq \vec{0}$ и нормируем его, положив

$$\vec{e}_1 = \frac{1}{\left\|\vec{x}_1\right\|}\vec{x}_1 = \frac{1}{\sqrt{\left(\vec{x}_1, \vec{x}_1\right)}}\vec{x}_1. \tag{6.5.11}$$

Построим гиперплоскость $H^{n-1} \subset E^n$, ортогональную вектору $\vec{e}_1$ и проходящую через начало репера $\left\{O, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\}$ (подпространство, ортогональное вектору $\vec{e}_1$). Эта гиперплоскость несёт на себе евклидову геометрию $n-1$ измерений.

Выберем в гиперплоскости вектор $\vec{x}_2 \in H^{n-1}$, $\vec{x}_2 \neq \vec{0}$, и нормируем его, положив

$$\vec{e}_2 = \frac{1}{\left\|\vec{x}_2\right\|}\vec{x}_2 = \frac{1}{\sqrt{\left(\vec{x}_2, \vec{x}_2\right)}}\vec{x}_2. \tag{6.5.12}$$

Построим гиперплоскость $H^{n-2} \subset H^{n-1}$, ортогональную вектору $\vec{e}_2$ и проходящую через начало репера $\left\{O, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\}$.

Выберем вектор $\vec{x}_3 \in H^{n-2}$, $\vec{x}_3 \neq \vec{0}$, и нормируем его. И так далее ...

Повторяя процесс n раз, получим ***флаг гиперплоскостей*** (подпространств)

$$H^1 \subset H^2 \subset \ldots \subset H^{n-1} \subset E^n,$$

проходящих через начало репера $\left\{O, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\}$, и множество единичных векторов — ортов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$. Из проведённого построения следует, что

$$(\forall k = 1, 2, \ldots, n)\ \vec{e}_k \in H^{n-k+1} \left(H^n \equiv E^n\right) \text{ и } \vec{e}_k \perp H^{n-k}.$$

Получили ортонормированный базис:

$$g_{ij} \equiv \left(\vec{e}_i, \vec{e}_j\right) = \delta_{ij} = \begin{cases} 1, i = j; \\ 0, i \neq j. \end{cases} \tag{6.5.13}$$

Напомним, что в ортонормированном репере

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i y^j = x^1 y^1 + x^2 y^2 + \ldots + x^n y^n.$$

Кроме этого, в ортонормированном репере контравариантные координаты совпадают с ковариантными координатами, то есть имеем

$$x^k = \sum_{j=1}^{n} g_{kj} x^j = x_k.$$

Поэтому в ортонормированном базисе для скалярного произведения имеем:

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x_i y_j = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n,$$

$$\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x_i x_j = x_1^2 + x_2^2 + \ldots + x_n^2.$$

Расстояние между точками $A\left(x_1; x_2; \ldots; x_n\right)$ и $B\left(y_1; y_2; \ldots; y_n\right)$ выражается формулой

$$\rho\left(A, B\right) = \sqrt{\left(y_1 - x_1\right)^2 + \left(y_2 - x_2\right)^2 + \ldots + \left(y_n - x_n\right)^2},$$

причём расстояние всегда вещественно и строго положительно, за исключением случая, когда точки A и B совпадают, тогда $\rho\left(A, B\right) = 0$.

Напомним ***конструктивную реализацию*** процесса ортогонализации Шмидта.

Пусть $\left\{O, \vec{g}_1, \vec{g}_2, \ldots, \vec{g}_n\right\} \subset E^n$ — некоторый репер. Для построения ортогонального базиса положим $\vec{a}_1 = \vec{g}_1$, $\vec{a}_2 = \vec{g}_2 + \alpha_1 \vec{a}_1$. Число α_1 найдём из условия ортогональности векторов $\vec{a}_1$ и $\vec{a}_2$:

$$\left(\vec{a}_1, \vec{a}_2\right) = 0 \Rightarrow \left(\vec{a}_1, \vec{g}_2 + \alpha_1 \vec{a}_1\right) = 0 \Rightarrow \left(\vec{a}_1, \vec{g}_2\right) + \alpha_1 \left(\vec{a}_1, \vec{a}_1\right) = 0.$$

Отсюда $\alpha_1 = -\dfrac{\left(\vec{a}_1, \vec{g}_2\right)}{\left(\vec{a}_1, \vec{a}_1\right)} = -\dfrac{\left(\vec{a}_1, \vec{g}_2\right)}{\left\|\vec{a}_1\right\|^2}$.

Итак, имеем

$$\vec{a}_2 = \vec{g}_2 + \alpha_1 \vec{a}_1 = \vec{g}_2 - \frac{\left(\vec{a}_1, \vec{g}_2\right)}{\left(\vec{a}_1, \vec{a}_1\right)} \vec{a}_1 = \vec{g}_2 - \frac{\left(\vec{a}_1, \vec{g}_2\right)}{\left\|\vec{a}_1\right\|^2} \vec{a}_1. \tag{6.5.14}$$

Положим теперь $\vec{a}_3 = \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2$ и потребуем выполнения условий

$$\left(\vec{a}_1, \vec{a}_3\right) = 0, \left(\vec{a}_2, \vec{a}_3\right) = 0.$$

Из этих условий имеем:

$$\left(\vec{a}_1, \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2\right) = 0 \Rightarrow \left(\vec{a}_1, \vec{g}_3\right) + \alpha_1 \left(\vec{a}_1, \vec{a}_1\right) = 0,$$

так как $\left(\vec{a}_1, \vec{a}_2\right) = 0$. Откуда получаем

$$\alpha_1 = -\frac{\left(\vec{a}_1, \vec{g}_3\right)}{\left(\vec{a}_1, \vec{a}_1\right)} = -\frac{\left(\vec{a}_1, \vec{g}_3\right)}{\left\|\vec{a}_1\right\|^2}.$$

Аналогично для α_2 имеем:

$$\left(\vec{a}_2, \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2\right) = 0 \Rightarrow \left(\vec{a}_2, \vec{g}_3\right) + \alpha_2 \left(\vec{a}_2, \vec{a}_2\right) = 0,$$

так как $\left(\vec{a}_1, \vec{a}_2\right) = 0$. Откуда

$$\alpha_2 = -\frac{\left(\vec{a}_2, \vec{g}_3\right)}{\left(\vec{a}_2, \vec{a}_2\right)} = -\frac{\left(\vec{a}_2, \vec{g}_3\right)}{\left\|\vec{a}_2\right\|^2}.$$

Итак, аналогично формуле (6.5.14), получаем

$$\vec{a}_3 = \vec{g}_3 + \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 = \vec{g}_3 - \frac{\left(\vec{a}_1, \vec{g}_3\right)}{\left\|\vec{a}_1\right\|^2}\vec{a}_1 - \frac{\left(\vec{a}_2, \vec{g}_3\right)}{\left\|\vec{a}_2\right\|^2}\vec{a}_2. \tag{6.5.15}$$

Продолжая этот процесс, для некоторого номера $k < n$ получим:

$$\vec{a}_k = \vec{g}_k - \frac{\left(\vec{a}_1, \vec{g}_k\right)}{\left\|\vec{a}_1\right\|^2}\vec{a}_1 - \frac{\left(\vec{a}_2, \vec{g}_k\right)}{\left\|\vec{a}_2\right\|^2}\vec{a}_2 - \ldots - \frac{\left(\vec{a}_{k-1}, \vec{g}_k\right)}{\left\|\vec{a}_{k-1}\right\|^2}\vec{a}_{k-1}. \tag{6.5.16}$$

В разложении (6.5.16) вектор $\vec{a}_k \neq \vec{0}$. Действительно, $\vec{a}_k$ является линейной комбинацией $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-1}, \vec{g}_k$. Вектор $\vec{a}_{k-1}$ является линейной комбинацией $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k-2}, \vec{g}_{k-1}$, и так далее. Окончательно будем иметь

$$\vec{a}_k = {}^{2}{}_{1}\vec{g}_1 + {}^{2}{}_{2}\vec{g}_2 + \ldots + {}^{2}{}_{k-1}\vec{g}_{k-1} + \vec{g}_k,$$

откуда и следует, что $\vec{a}_k \neq \vec{0}$.

Продолжая процесс до номера $k = n$, получим ортогональный репер $\left\{O, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right\} \subset E^n$, нормируя который обычным образом, получим ортонормированный репер $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$.

Ортонормированный репер в комплексном евклидовом пространстве. Рассмотрим процесс ортогонализации в комплексном евклидовом пространстве. Предварительно докажем простую лемму.

Лемма 6.5.1. *В евклидовом пространстве не могут быть изотропными все векторы.*

Д о к а з а т е л ь с т в о. Пусть все векторы в евклидовом пространстве изотропны, то есть

$$\left(\forall \vec{x} \in E^n\right)\left(\vec{x}, \vec{x}\right) = 0.$$

Тогда, в частности, имеем:

$$\left(\forall \vec{x}, \vec{y} \in E^n\right)\left(\frac{\vec{x}+\vec{y}}{2}, \frac{\vec{x}+\vec{y}}{2}\right) = 0 \text{ и } \left(\frac{\vec{x}-\vec{y}}{2}, \frac{\vec{x}-\vec{y}}{2}\right) = 0.$$

Вычитая эти два равенства почленно, получаем

$$\left(\forall \vec{x}, \vec{y} \in E^n\right)\left(\vec{x}, \vec{y}\right) = 0.$$

Таким образом, любой вектор ортогонален ко всем векторам пространства, что противоречит аксиоме невырожденности. ●●

Итак, в комплексном евклидовом пространстве U^n, согласно лемме 6.5.1, всегда можно найти неизотропный вектор $\vec{x}_1 \in U^n$, то есть такой вектор, для которого выполняется равенство $\left(\vec{x}_1, \vec{x}_1\right) \neq 0$. Нормируем этот вектор, полагая орт

$$\vec{e}_1 = \frac{1}{\left\|\vec{x}_1\right\|}\vec{x}_1, \tag{6.5.17}$$

причём число $\frac{1}{\left\|\vec{x}_1\right\|}$ теперь, вообще говоря, является комплексным. Очевидно, что

$$\left(\vec{e}_1, \vec{e}_1\right) = \frac{1}{\left\|\vec{x}_1\right\|^2}\left(\vec{x}_1, \vec{x}_1\right) = 1.$$

Построим гиперплоскость $H_+^{n-1} \subset U^n$, ортогональную орту $\vec{e}_1$ и проходящую через начало некоторого репера $O \in U^n$. По теореме 6.5.2 эта гиперплоскость, будучи ортогональной неизотропному вектору, сама является неизотропной и по этой причине несёт на себе комплексную евклидову геометрию $n-1$ измерений, то есть является подпространством $U^{n-1} \subset U^n$. Выбирая на гиперплоскости H_+^{n-1} некоторый вектор $\vec{x}_2$ и нормируя его, получим орт

$$\vec{e}_2 = \frac{1}{\left\|\vec{x}_2\right\|}\vec{x}_2.$$

Обозначим $H_+^{n-2} \subset H_+^{n-1}$ гиперплоскость, ортогональную неизотропному орту $\vec{e}_2$ и проходящую через начало $O \in U^n$. Гиперплоскость H_+^{n-2}, как ортогональная неизотропному вектору $\vec{e}_2$, сама неизотропная и несёт на себе комплексную евклидову геометрию размерности $n-2$, то есть является подпространством $U^{n-2} \subset U^{n-1} \subset U^n$. Выбирая на ней некоторый вектор $\vec{x}_3$ и нормируя его, получим орт $\vec{e}_3 = \frac{1}{\left\|\vec{x}_3\right\|}\vec{x}_3$. Процесс продолжаем до получения одномерной гиперплоскости $H_+^1 \subset H_+^2$, на которой строим орт

$$\vec{e}_n = \frac{1}{\left\|\vec{x}_n\right\|}\vec{x}_n.$$

Получаем флаг неизотропных гиперплоскостей

$$H_+^1 \subset H_+^2 \subset \ldots \subset H_+^{n-2} \subset H_+^{n-1} \subset U^n$$

и систему ортов $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$, для которых аналогично случаю собственно евклидова пространства $\left(\vec{e}_i, \vec{e}_j\right) = \delta_{ij}$. Очевидно, что по построению система $\left\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$ является линейно независимой и является базисом ортонормированного репера $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset U^n$.

Так же, как и в собственно евклидовом пространстве, в комплексном евклидовом пространстве ковариантные координаты вектора в построенном репере совпадают, то есть $\left(\forall i = 1, 2, \ldots, n\right)\ x_i = x^i$. Справедливы также формулы для скалярного произведения и метрики, аналогичные формулам для собственно евклидова пространства:

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x_i y_j = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n;$$

$$\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}x_i x_j = x_1^2 + x_2^2 + \ldots + x_n^2;$$

$$\rho\left(A, B\right) = \sqrt{\left(y_1 - x_1\right)^2 + \left(y_2 - x_2\right)^2 + \ldots + \left(y_n - x_n\right)^2}.$$

Существенное отличие состоит в том, что координаты всех векторов и метрические коэффициенты теперь, вообще говоря, будут комплексными и, следовательно, комплексными будут скалярное произведение и расстояние.

Ортонормированный репер в вещественном евклидовом пространстве [21]. В этом случае также начинаем с выбора неизотропного вектора $\vec{x}_1 \in E^n$, то есть такого вектора, для которого выполняется условие $\left(\vec{x}_1, \vec{x}_1\right) \neq 0$. Однако при нормировке этого вектора могут реализоваться два случая.

В первом случае $\left(\vec{x}_1, \vec{x}_1\right) > 0$ и имеем:

$$\vec{e}_1 = \frac{1}{\left\|\vec{x}_1\right\|}\vec{x}_1, \tag{6.5.18}$$

$$\left(\vec{e}_1, \vec{e}_1\right) = \frac{1}{\left\|\vec{x}_1\right\|^2}\left(\vec{x}_1, \vec{x}_1\right) = 1. \tag{6.5.19}$$

Во втором случае $\left(\vec{x}_1, \vec{x}_1\right) < 0$ и знаменатель $\left\|\vec{x}_1\right\|$ в формулах (6.5.18), (6.5.19) является ***чисто мнимым***. Так как пространство E^n вещественно, формулы (6.5.18), (6.5.19) не имеют смысла. Чтобы обойти эту трудность, проводим нормировку вектора $\vec{x}_1 \in E^n$ по формуле

$$\vec{e}_1 = \frac{1}{\sqrt{-\left(\vec{x}_1, \vec{x}_1\right)}}\vec{x}_1 = \frac{1}{\left\|\vec{x}_1\right\|}\vec{x}_1, \tag{6.5.20}$$

где $\left\|\vec{x}_1\right\|$ является уже вещественным числом, и формула (6.5.20) законна. Из формулы (6.5.20) имеем:

$$\left(\vec{e}_1, \vec{e}_1\right) = \frac{1}{-\left(\vec{x}_1, \vec{x}_1\right)}\left(\vec{x}_1, \vec{x}_1\right) = \frac{1}{\left\|\vec{x}_1\right\|^2}\left(\vec{x}_1, \vec{x}_1\right) = -1. \tag{6.5.21}$$

Определение 6.5.3. *Вектор $\vec{e}$, такой, что*

$$\left(\vec{e}, \vec{e}\right) = -1,$$

*называется **мнимоединичным**.* •

Отметим особо, что *мнимоединичные векторы не являются «мнимыми» — это обычные вещественные векторы вещественного евклидова пространства, имеющие, однако, мнимую норму*

$$\left\|\vec{e}\right\| = \sqrt{-1} = i.$$

Построив единичный или мнимоединичный орт $\vec{e}_1$, аналогично случаю комплексного евклидова пространства проводим через фиксированную точку O ортогональную к нему гиперплоскость H^{n-1}. Будучи ортогональной к неизотропному вектору $\vec{e}_1$ эта гиперплоскость сама неизотропная и несёт на себе евклидову геометрию $n-1$ измерений. На гиперплоскости H^{n-1} снова выбираем единичный или мнимоединичный вектор $\vec{e}_2$, и так далее ...В результате получаем ортонормированный репер $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$, орты которого удовлетворяют условию:

$$\left(\vec{e}_i, \vec{e}_j\right) = {}'_{ij} = \begin{cases} \pm 1, i = j; \\ 0, i \neq j. \end{cases} \tag{6.5.22}$$

Таким образом, в вещественном евклидовом пространстве часть ортов являются ***единичными***, а часть — ***мнимоединичными***.

Пусть сначала следуют мнимоединичные орты

$$\left(\vec{e}_1, \vec{e}_1\right) = \left(\vec{e}_2, \vec{e}_2\right) = \ldots = \left(\vec{e}_k, \vec{e}_k\right) = -1, \tag{6.5.23}$$

а затем единичные орты

$$\left(\vec{e}_{k+1}, \vec{e}_{k+1}\right) = \left(\vec{e}_{k+2}, \vec{e}_{k+2}\right) = \ldots = \left(\vec{e}_n, \vec{e}_n\right) = 1. \tag{6.5.24}$$

В равенствах (6.5.23) для числа мнимоединичных ортов возможны значения $k = 0, 1, 2, \ldots, n$. Соответственно этому для метрических коэффициентов имеем:

$$g_{ij} = \begin{cases} 0, i \neq j; \\ -1, i = j = 1, 2, \ldots, k; \\ 1, i = j = k+1, k+2, \ldots, n. \end{cases} \tag{6.5.25}$$

Число $0 < k \leq n$ *мнимоединичных ортов, входящих в репер* $\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\} \subset E^n$ *вещественного евклидова пространства, называется* ***индексом пространства***, а само пространство E^n — ***псевдоевклидовым пространством индекса*** k. Связь ковариантных и контравариантных координат вектора в таком репере принимает вид:

$$x_i = -x^i \left(\forall i = 1, 2, \ldots, k\right);\ x_i = x^i \left(\forall i = k+1, k+2, \ldots, n\right). \qquad (6.5.26)$$

Таким образом, разница между двумя типами координат не исчезает полностью, хотя и становится малосущественной.

Для скалярного произведения и расстояния получаем:

$$\left(\vec{x}, \vec{y}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i y^j = -\sum_{i=1}^{k} x_i y_i + \sum_{i=k+1}^{n} x_i y_i;$$

$$\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i x^j = -\sum_{i=1}^{k} \left(x_i\right)^2 + \sum_{i=k+1}^{n} \left(x_i\right)^2;$$

$$\rho\left(A, B\right) = \sqrt{-\sum_{i=1}^{k}\left(y_i - x_i\right)^2 + \sum_{i=k+1}^{n}\left(y_i - x_i\right)^2}.$$

Определение 6.5.4. *Форма второй степени*

$$G\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i x^j$$

*называется **метрической квадратичной формой** (или **метрикой**).* •

Из приведённого рассмотрения вытекает теорема.

Теорема 6.5.5. *Метрическая квадратичная форма*

$$G\left(\vec{x}, \vec{x}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} x^i x^j$$

*приводится к **каноническому виду** суммы-разности только в ортонормированных системах координат.*

Справедлива также следующая важнейшая теорема.

Теорема 6.5.6. *Независимо от выбора ортонормированного репера*

$$\left\{O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\right\}$$

в вещественном евклидовом пространстве E^n число k мнимоединичных ортов (индекс пространства), а следовательно, и число $n-k$ единичных ортов остаётся постоянным.

Доказательство. Пусть

$$\left\{O, \vec{e}_1, \ldots, \vec{e}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n\right\} \text{ и } \left\{O, \vec{a}_1, \ldots, \vec{a}_m, \vec{a}_{m+1}, \ldots, \vec{a}_n\right\}$$

— два ортонормированных репера, где k — число мнимоединичных ортов первого, а m — число мнимоединичных ортов второго реперов.

Предположим, например, что $m > k$. Покажем, что это предположение приводит к противоречию. Для этого рассмотрим систему мнимоединичных ортов второго репера и единичных ортов первого репера:

$$\left\{ \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m, \vec{e}_{k+1}, \ldots, \vec{e}_n \right\}.$$

Их число $m + n - k > n$ (так как $m > k$ и, следовательно, $m - k > 0$). Поэтому данная система должна быть линейно зависимой. Откуда имеем, например, такое равенство:

$$\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_m \vec{a}_m = \beta_{k+1} \vec{e}_{k+1} + \beta_{k+2} \vec{e}_{k+2} + \ldots + \beta_n \vec{e}_n.$$

Приравнивая скалярные квадраты обеих частей последнего равенства (если векторы равны, то равны и их скалярные квадраты) с учётом равенств (6.5.26), получаем

$$-\alpha_1^2 - \alpha_2^2 - \ldots - \alpha_m^2 = \beta_{k+1}^2 + \beta_{k+2}^2 + \ldots + \beta_n^2.$$

Последнее равенство справедливо лишь в случае, когда все коэффициенты одновременно обращаются в нуль. Но тогда система ортов должна быть линейно независима, что невозможно. ••

Из доказанной теоремы следует, *что при любом способе приведения вещественной (метрической) квадратичной формы к каноническому виду число положительных и отрицательных квадратов не меняется*. Это — ***закон инерции квадратичных форм***, с которым мы уже встречались выше в случае собственно евклидова пространства.

В заключение отметим, что введённые в этом параграфе понятия индекса пространства и псевдоевклидова пространства широко применяются в современной физике, а именно в специальной теории относительности [38].

Глава 7.
Физические приложения теории конечномерных векторных пространств, линейных операторов и матриц

В этой главе в весьма сжатом виде приведены примеры физических приложений теории конечномерных векторных пространств, систем линейных алгебраических уравнений и теории линейных операторов. Таких приложений очень много и даже для их краткого обзора потребовалось бы написание отдельной книги. Мы, однако, остановимся только на нескольких простейших приложениях из теоретической механики, электротехники и квантовой механики. Для понимания приведённых в данном параграфе сведений требуется наличие первоначальных познаний из классической механики, теории постоянного и переменного электрического тока и простейших понятий квантовой механики. Большинство из этих понятий даётся ещё в курсе физики средней школы и конечно в вузовском курсе общей физики.

7.1. Инерциальные системы координат в классической механике [37]

В классической механике Ньютона большое значение имеют так называемые инерциальные системы координат. Их существование постулируется в первом законе динамики, который в современной формулировке может быть представлен в виде аксиомы. Формулировка первой аксиомы классической динамики базируется на первоначальных понятиях о пространстве и времени, фундаментом которых как раз и является при современном уровне подходов теория векторных пространств и отображений в этих пространствах. Отвлекаясь от экспериментальных основ, которые достаточно полно излагаются в курсах физики, сформулируем базовые понятия классической механики.

1) Моделью реального ***физического пространства***, заполненного материей, в классической механике является четырёхмерное аффинное пространство $A^4 = E^3 \otimes R^1$, являющееся прямым произведением трёхмерного собственно евклидова пространства E^3 (определение 3.5) и действительной числовой оси R^1. Это пространство A^4 называется ***Мир*** или ***Вселенная***. Предполагается, что в пространстве E^3 зафиксирован ортонормированный репер.

2) Под ***временем*** понимается отображение $\hat{T} : A^4 \to R^1$ Вселенной на ***вещественную ось времени***. Время считается абсолютным и однородным, характеризует длительность процессов и не зависит от каких-либо физических характеристик пространства E^3.

3) Под ***событием*** в классической механике понимают фиксацию некоторых наблюдаемых величин, характеризующих положение материального тела и изменение его положения в пространстве E^3 с течением времени.

Обозначим событие как четырёхмерный вектор в пространстве A^4, например, вектор $\vec{X}$. Если имеем два события $\vec{X}_1$ и $\vec{X}_2$, то ***промежутком времени от события*** $\vec{X}_1 \in A^4$ ***до события*** $\vec{X}_2 \in A^4$ называется значение действительной функции — число

$$t = t\left(\vec{X}_2 - \vec{X}_1\right).$$

Если

$$\left(\forall \vec{X}_1, \vec{X}_2 \in R^4\right) t\left(\vec{X}_2 - \vec{X}_1\right) = 0,$$

то события $\vec{X}_1$ и $\vec{X}_2$ называются ***одновременными***. Подпространство E^3 (определения 3.4.1 и 3.4.2) Вселенной A^4 является, очевидно, множеством событий, происходящих в один и тот же момент времени, и называется ***пространством одновременных событий***.

Определение 7.1.1. *Прямое произведение $A^4 = E^3 \otimes R^1$ пространства одновременных событий E^3 и оси времени R^1 в классической механике называется* ***пространством событий***. •

Таким образом, ***Вселенная в классической механике*** является как бы ***сценой***, на которой происходят различные события.

Расстояние между одновременными событиями определяется через скалярное произведение известной формулой (геометрические векторы в пространстве одновременных событий обозначаем малыми буквами латинского алфавита со стрелкой)

$$\rho\left(\vec{x}_1, \vec{x}_2\right) = \left\|\vec{x}_2 - \vec{x}_1\right\| = \sqrt{\left(\vec{x}_2 - \vec{x}_1, \vec{x}_2 - \vec{x}_1\right)},$$

которая следует из определения собственно евклидова пространства. Пространство одновременных событий является в классической механике ***однородным и изотропным***. Однородность и изотропность пространства означают равноправность любых выбираемых реперов, независимо от положения их начала и от их ориентации. Напомним, что термин «собственно евклидово пространство» означает выполнение в пространстве E^3 аксиомы невырожденности скалярного произведения.

Если в пространстве одновременных событий зафиксирована декартова система координат, то есть в качестве пространства одновременных событий рассматривается пространство R^3, то пространство событий называется ещё ***пространством Галилея***, или галилеевым пространством.

4) Материальной точкой в классической механике называется геометрическая точка $\vec{x} \in E^3$ пространства одновременных событий, которой приписано некоторое неотрицательное число m, называемое ***массой***.

Совокупность материальных точек, объединённых некоторыми условиями, образует ***систему материальных точек***.

Множество точек в галилеевом пространстве $R^1 \otimes E^3$, определённое для каждой материальной точки как множество вида

$$\left\{\vec{X} = \left(\vec{x}, t\right) \equiv \left(x^1, x^2, x^3, t\right) \in R^1 \otimes E^3 : \vec{x} \in E^3 \wedge t \in R^1\right\},$$

называется ***мировой линией*** материальной точки. Изобразить мировую линию материальной точки можно лишь для случая трёхмерного галилеева пространства $R^1 \otimes R^2$. На рис. 7.1 изображены мировые линии системы четырёх различных материальных точек.

Теперь можно сформулировать первую аксиому динамики.

Аксиома 1. *Во Вселенной существует система координат S, относительно которой изолированная материальная точка сохраняет состояние покоя или равномерного прямолинейного движения.*

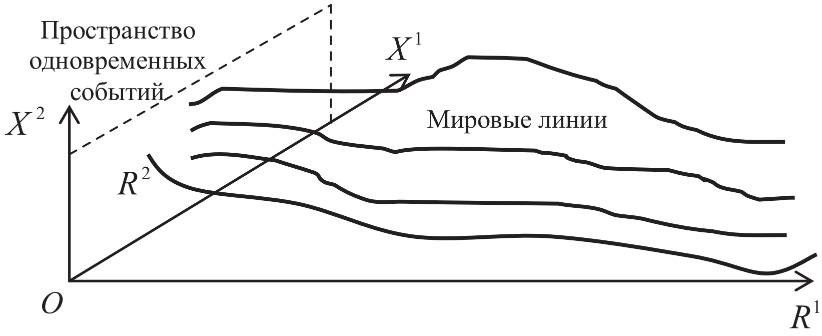


Рис. 7.1

Определение 7.1.2. *Система координат S, фигурирующая в первой аксиоме динамики, называется* ***инерциальной системой координат***. •

Определение 7.1.3. ***Системой отсчёта*** *в классической механике называется совокупность системы координат в пространстве одновременных событий* E^3 *и системы координат на числовой оси времени* R^1. •

Обычно систему координат связывают с некоторым ***отсчётным телом***, а в качестве системы координат на оси времени выбирают некоторые ***часы*** — циклическое природное явление или прибор, реализующий циклический процесс. Совокупность отсчётного тела, системы координат и часов называется ***системой отсчёта***. Далее для простоты будем называть систему отсчёта просто системой координат.

Теорема 7.1.1. *Система координат* S'*, движущаяся относительно инерциальной системы координат S поступательно с постоянной скоростью, также является инерциальной.*

Доказательство. Пусть закон движения материальной точки в инерциальной системе координат S имеет вид:

$$\vec{R} = \vec{R}_0 + \vec{V}_0 t, \tag{7.1.1}$$

где

$$t \in R^1,\ \vec{R} = \sum_{i=1}^{3} x^i \vec{e}_i.$$

Здесь векторы

$$\vec{R}_0 = \sum_{i=1}^{3} x_0^i \vec{e}_i,\ \vec{V}_0 = \sum_{i=1}^{3} V_0^i \vec{e}_i$$

— это постоянные векторы. Подстановка в закон движения (7.1.1) с последующим разложением по базису декартова репера приводит к координатным равенствам

$$x^i = x_0^i + V_0^i t,\ \left(i = 1, 2, 3\right).$$

Пусть начало A системы координат S' движется относительно начала системы координат S по закону

$$\vec{R}_A = \vec{a} + \vec{b}\,t, \tag{7.1.2}$$

где

$$\vec{R}_A = \sum_{i=1}^{3} x_A^i \vec{e}_i,$$

а векторы

$$\vec{a} = \sum_{i=1}^{3} a^i \vec{e}_i,\ \vec{b} = \sum_{i=1}^{3} b^i \vec{e}_i$$

— постоянные векторы. Подстановка этих разложений в сумму (7.1.2) даёт

$$x_A^i = a^i + b^i t.$$

Очевидно, что в соответствии с правилом треугольника имеет место векторное равенство

$$\vec{R} = \vec{R}_A + \vec{r}\left(t\right) = \vec{R}_A + \sum_{j=1}^{3} y^j \vec{a}_j. \tag{7.1.3}$$

Здесь орты системы S' выражаются через орты инерциальной системы координат по формулам:

$$\vec{a}_j = \sum_{i=1}^{3} A_j^i \vec{e}_i,$$

где $\left(A_j^i\right)$ — матрица перехода (формулы (3.3.2) и (3.3.4)) от старого базиса к новому базису. Эта матрица ортогональная, так как переводит ортонормированный базис в ортонормированный базис (теорема 4.2.10 — критерий ортогональности оператора).

Подставляя в правило треугольника (7.1.3) выражения (7.1.1) и (7.1.2), получаем (рис. 7.2)

$$\vec{R}_0 + \vec{V}_0 t = \vec{a} + \vec{b}\,t + \sum_{j=1}^{3} y^j \vec{a}_j.$$

Переходя в этом выражении к разложениям по базисным векторам, получим:

$$\sum_{i=1}^{3}\left(x_0^i - a^i\right)\vec{e}_i + \left[\sum_{i=1}^{3}\left(V_0^i - b^i\right)\vec{e}_i\right]t = \sum_{j=1}^{3} y^j \sum_{i=1}^{3} A_j^i \vec{e}_i.$$

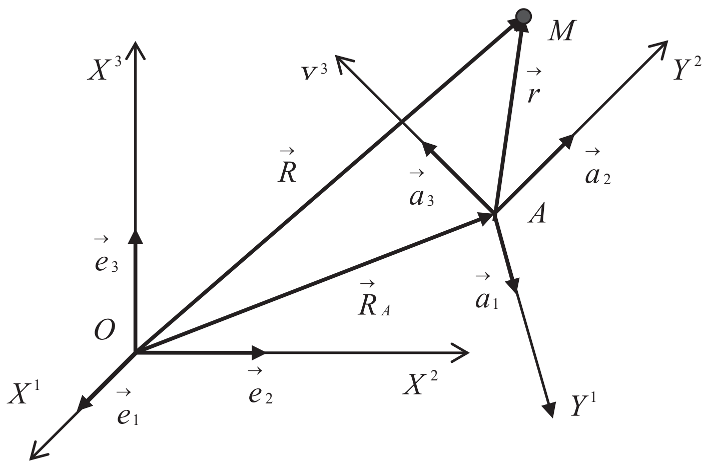


Рис. 7.2

Меняя порядок суммирования в правой части предыдущего равенства, имеем

$$\sum_{i=1}^{3}\left[\left(x_0^i - a^i\right)+\left(V_0^i - b^i\right)t\right]\vec{e}_i = \sum_{i=1}^{3}\left(\sum_{j=1}^{3} A_j^i y^j\right)\vec{e}_i,$$

откуда следует координатное равенство

$$\left(x_0^i - a^i\right)+\left(V_0^i - b^i\right)t = \sum_{j=1}^{3} A_j^i y^j. \tag{7.1.4}$$

Так как матрица перехода ортогональная, то $A^{-1} = A^T$. Умножая обе части равенства (7.1.4) на обратную матрицу $A^{-1} = A^T$, имеем

$$y^j = \sum_{i=1}^{3} A_i^j\left[\left(x_0^i - a^i\right)+\left(V_0^i - b^i\right)t\right].$$

Вводя обозначения

$$\vec{r}_0 = \sum_{j=1}^{3}\left[\sum_{i=1}^{3} A_i^j\left(x_0^i - a^i\right)\right]\vec{a}_j,\ \vec{v}_0 = \sum_{j=1}^{3}\left[\sum_{i=1}^{3} A_i^j\left(V_0^i - b^i\right)\right]\vec{a}_j,$$

получаем равенство

$$\vec{r} = \vec{r}_0 + \vec{v}t, \tag{7.1.5}$$

по существу являющееся векторным параметрическим уравнением прямой линии следующего вида:

$$\vec{x} = \vec{x} + t\vec{a}.$$

Из уравнения (7.1.5) следует, что изолированная точка движется относительно системы координат S' равномерно и прямолинейно, то есть система координат S' является инерциальной. ••

Следствие 1 из теоремы 7.1.1. *Для любой изолированной материальной точки найдётся такая система координат, в которой материальная точка покоится и совпадает с началом этой системы координат.*

Д о к а з а т е л ь с т в о. Полагая в уравнении (7.1.5) $\vec{a} = \vec{R}_0$ и $\vec{b} = \vec{V}_0$, получаем $\vec{r}(t) = \vec{0}$. ••

Для понимания следующего утверждения необходимо владение простейшими понятиями математического анализа.

Следствие 2 из теоремы 7.1.1. *Ускорение материальной точки инвариантно относительно выбора инерциальной системы координат.*

Д о к а з а т е л ь с т в о. Закон движения материальной точки имеет вид:

$$\vec{R}(t) = \vec{a} + \vec{b}\,t + \sum_{i=1}^{3}\left(\sum_{j=1}^{3} A_j^i y^j\right)\vec{e}_i. \tag{7.1.6}$$

Дифференцируя обе части векторного уравнения (7.1.6) по времени дважды, получаем

$$\frac{d^2\vec{R}(t)}{dt^2} = \sum_{i=1}^{3}\left(\sum_{j=1}^{3} A_j^i \frac{d^2 y^j}{dt^2}\right)\vec{e}_i. \text{ ••}$$

В классической механике все инерциальные системы координат считаются неразличимыми, а именно, постулируется справедливость следующего принципа.

Принцип относительности Галилея: *все инерциальные системы координат во Вселенной эквивалентны в механическом смысле, то есть все законы классической механики формулируются одинаково в любых инерциальных системах координат.*

Экспериментально установлено, что неподвижных систем координат во Вселенной не существует. По этой причине под неподвижной системой координат в механике понимается инерциальная система координат, конкретный выбор которой зависит от решаемой задачи. Так, например, при решении большинства задач небесной механики в качестве инерциальной (неподвижной) системы координат можно выбрать ***систему Коперника*** — начало системы координат помещается в центр масс Солнечной системы, а оси направляются на «неподвижные» звёзды.

7.2. Структура кинетической энергии системы материальных точек в обобщённых координатах [37]

В теоретической механике (и в других разделах теоретической физики) для описания движения системы (материальных) точек широко используется понятие ***обобщённых координат***. Далее обозначаем систему точек M_k, как это принято в теории множеств, то есть

$$\left\{M_k \middle| k = \overline{1, N}\right\}.$$

Обобщённые координаты вводятся как параметры, характеризующие возможность точкам системы совершать перемещения в тот или иной фиксированный момент времени, то есть в пространстве одновременных состояний, которое очевидно является гиперплоскостью (параграф 2.8) в галилеевом пространстве $R^1 \otimes E^3$. Число таких ***виртуальных*** перемещений, при условии их независимости, как раз и называется **числом степеней свободы** системы $\left\{M_k \middle| k = \overline{1, N}\right\}$.

Обобщённые координаты выбирают, исходя из физического смысла решаемой задачи. Обобщённые координаты могут быть расстояниями, углами, площадями и тому подобное. Очень часто обобщённые координаты вообще не имеют геометрического толкования.

Чтобы некоторые величины можно было принять за обобщённые координаты, они должны удовлетворять следующим требованиям:

1) ***обобщённые координаты должны быть линейно независимыми в совокупности и их число должно быть равно числу степеней свободы системы***;

2) ***декартовы координаты точек системы должны единственным образом выражаться через обобщённые координаты***.

Дальше будем обозначать обобщённые координаты в совокупности $\left\{q^1, q^2, \ldots, q^m\right\}$ и писать

$$\begin{cases} \vec{r}_1 = \vec{r}_1\left(q^1, q^2, \ldots, q^m, t\right), \\ \vec{r}_2 = \vec{r}_2\left(q^1, q^2, \ldots, q^m, t\right) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \vec{r}_N = \vec{r}_N\left(q^1, q^2, \ldots, q^m, t\right). \end{cases} \qquad (7.2.1)$$

Отметим, что обобщённые координаты произвольны, в частности, в качестве обобщённых можно выбрать криволинейные координаты, которые будут изучаться в соответствующем месте.

Здесь $\vec{r}_k$ — радиус-вектор точки системы с номером k.

Обобщённые координаты можно рассматривать как координаты

$$\left\{q^1, q^2, \ldots, q^m\right\} \subset X^m$$

в абстрактном векторном пространстве X^m, которое называется ***конфигурационным пространством*** (***пространством конфигураций***) системы $\left\{M_k \middle| k = \overline{1, N}\right\}$. Действительно, из системы функциональных зависимостей (7.2.1) следует, что в каждый момент времени набору $\left\{q^1, q^2, \ldots, q^m\right\}$ обобщённых координат соответствует единственное возможное положение системы и наоборот (более строго и подробно данное соответствие будет рассмотрено в разделе, посвящённом теории дифференцируемых преобразований). Таким образом, каждому возможному положению точек системы $\left\{q^1, q^2, \ldots, q^m\right\}$ в конфигурационном пространстве соответствует ***изображающая точка***. Движению системы в реальном пространстве R^3 соответствует движение изображающей точки в конфигурационном пространстве X^m.

Первая и вторая производные обобщённых координат по времени в механике обозначаются точками

$$\dot{q}^j \equiv \frac{dq^j}{dt}, \ddot{q}^j \equiv \frac{d^2q^j}{dt^2} \left(j = 1, 2, \ldots, m\right)$$

и называются, соответственно, ***обобщёнными скоростями*** и ***обобщёнными ускорениями*** точек системы.

Продифференцируем вектор-функции (7.2.1) по времени дважды:

$$\vec{v}_k = \frac{d\vec{r}_k}{dt} = \sum_{j=1}^{m} \frac{\partial \vec{r}_k}{\partial q^j} \frac{dq^j}{dt} + \frac{\partial \vec{r}_k}{\partial t}, \tag{7.2.2}$$

$$\vec{w}_k = \frac{d^2 \vec{r}_k}{dt^2} = \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 \vec{r}_k}{\partial q^i \partial q^j} \frac{\partial q^i}{\partial t} \frac{\partial q^j}{\partial t} + \sum_{j=1}^{m} \frac{\partial \vec{r}_k}{\partial q^j} \frac{d^2 q^j}{dt^2} +$$

$$+ \sum_{j=1}^{m} \frac{\partial^2 \vec{r}_k}{\partial q^j \partial t} \frac{dq^j}{dt} + \frac{\partial^2 \vec{r}_k}{\partial t^2}, \tag{7.2.3}$$

где $k = 1, 2, \ldots, N$. Получили выражения скоростей (7.2.2) и ускорений (7.2.3) точек системы через обобщённые скорости и обобщённые ускорения.

Дальше нам потребуется только выражение для скорости точек системы. Для так называемой ***склерономной системы материальных точек*** выражение для скорости (7.2.2) упрощается и принимает вид

$$\vec{v}_k = \frac{d\vec{r}_k}{dt} = \sum_{j=1}^{m} \frac{\partial \vec{r}_k}{\partial q^j} \frac{dq^j}{dt}. \tag{7.2.4}$$

Для точки системы с номером $k = 1, 2, \ldots, N$ функция

$$T_k \stackrel{def}{=} \frac{1}{2} m_k \left(\vec{v}_k, \vec{v}_k \right) = \frac{1}{2} m_k \left\| \vec{v}_k \right\|^2 \equiv \frac{1}{2} m_k v_k^2 \tag{7.2.5}$$

называется ***кинетической энергией*** материальной точки с номером k. Кинетическая энергия системы материальных точек является величиной аддитивной и в декартовой системе координат выражается формулой

$$T = \frac{1}{2} \sum_{k=1}^{N} m_k \left(\frac{d\vec{r}_k}{dt}, \frac{d\vec{r}_k}{dt} \right). \tag{7.2.6}$$

Рассмотрим структуру выражения (7.2.6) для кинетической системы из N материальных точек подробнее:

$$T = \frac{1}{2} \sum_{k=1}^{N} m_k \left(\frac{d\vec{r}_k}{dt}, \frac{d\vec{r}_k}{dt} \right) =$$

$$= \frac{1}{2} \sum_{k=1}^{N} m_k \left(\sum_{j=1}^{m} \frac{\partial \vec{r}_k}{\partial q^j} \frac{dq^j}{dt} + \frac{\partial \vec{r}_k}{\partial t}, \sum_{i=1}^{m} \frac{\partial \vec{r}_k}{\partial q^i} \frac{dq^i}{dt} + \frac{\partial \vec{r}_k}{\partial t} \right) =$$

$$= \frac{1}{2} \sum_{k=1}^{N} m_k \left[\sum_{j=1}^{m} \sum_{i=1}^{m} \left(\frac{\partial \vec{r}_k}{\partial q^j}, \frac{\partial \vec{r}_k}{\partial q^i} \right) \frac{dq^j}{dt} \frac{dq^i}{dt} + \right.$$

$$\left. + \sum_{j=1}^{m} \left(\frac{\partial \vec{r}_k}{\partial q^j}, \frac{\partial \vec{r}_k}{\partial t} \right) \frac{dq^j}{dt} + \sum_{i=1}^{m} \left(\frac{\partial \vec{r}_k}{\partial q^i}, \frac{\partial \vec{r}_k}{\partial t} \right) \frac{dq^i}{dt} + \left(\frac{\partial \vec{r}_k}{\partial t}, \frac{\partial \vec{r}_k}{\partial t} \right) \right] =$$

$$
=\sum_{j=1}^{m}\sum_{i=1}^{m}\left[\frac{1}{2}\sum_{k=1}^{N}m_k\left(\frac{\partial \vec{r}_k}{\partial q^j},\frac{\partial \vec{r}_k}{\partial q^i}\right)\right]\frac{dq^j}{dt}\frac{dq^i}{dt}+
$$

$$
+\sum_{j=1}^{m}\left[\sum_{k=1}^{N}m_k\left(\frac{\partial \vec{r}_k}{\partial q^j},\frac{\partial \vec{r}_k}{\partial t}\right)\right]\frac{dq^j}{dt}+\frac{1}{2}\sum_{k=1}^{N}m_k\left(\frac{\partial \vec{r}_k}{\partial t},\frac{\partial \vec{r}_k}{\partial t}\right).
$$

Введём обозначения:

$$
a_{ji}\equiv\frac{1}{2}\sum_{k=1}^{N}m_k\left(\frac{\partial \vec{r}_k}{\partial q^j},\frac{\partial \vec{r}_k}{\partial q^i}\right),
$$

$$
a_j\equiv\sum_{k=1}^{N}m_k\left(\frac{\partial \vec{r}_k}{\partial q^j},\frac{\partial \vec{r}_k}{\partial t}\right), \tag{7.2.7}
$$

$$
a_0\equiv\frac{1}{2}\sum_{k=1}^{N}m_k\left(\frac{\partial \vec{r}_k}{\partial t},\frac{\partial \vec{r}_k}{\partial t}\right).
$$

С учётом формул (7.2.7) выражение для кинетической энергии принимает следующий вид:

$$
T=\sum_{j=1}^{m}\sum_{i=1}^{m}a_{ji}\frac{dq^j}{dt}\frac{dq^i}{dt}+\sum_{j=1}^{m}a_j\frac{dq^j}{dt}+a_0. \tag{7.2.8}
$$

В выражении (7.2.8) коэффициенты a_{ji}, a_j и a_0 являются, очевидно, функциями от обобщённых координат $\{q^1,q^2,\ldots,q^m\}$ и времени t.

Итак, кинетическая энергия системы материальных точек является ***многочленом второй степени относительно обобщённых скоростей*** и может быть записана в виде

$$
T=T_2+T_1+T_0, \tag{7.2.9}
$$

где введены обозначения

$$
T_2\equiv\sum_{j=1}^{m}\sum_{i=1}^{m}a_{ji}\frac{dq^j}{dt}\frac{dq^i}{dt},\ T_1\equiv\sum_{j=1}^{m}a_j\frac{dq^j}{dt},\ T_0\equiv a_0.
$$

Для склерономной системы

$$\frac{\partial \vec{r}_k}{\partial t} = \vec{0}\left(k = 1, 2, \ldots, N\right)$$

и из формул (7.2.7) следует, что $a_j = 0, a_0 = 0$. Таким образом, для склерономной системы

$$T \equiv T_2 = \sum_{j=1}^{m}\sum_{i=1}^{m} a_{ji}\frac{dq^j}{dt}\frac{dq^i}{dt} \equiv \sum_{j=1}^{m}\sum_{i=1}^{m} a_{ji}\dot{q}^j\dot{q}^i. \qquad (7.2.10)$$

Получили следующий результат.

Лемма 7.2.1. *Кинетическая энергия склерономной системы является квадратичной формой относительно обобщённых скоростей с коэффициентами, не зависящими от времени явно.*

Напомним, что квадратичная форма называется невырожденной, если её ранг равен размерности пространства (в рассматриваемом случае это размерность конфигурационного пространства m). По определению ранга квадратичной формы или, что то же самое, ранга её матрицы для невырожденной формы выполняется условие

$$\det\left(a_{ij}\right) \neq 0,$$

где $j, i = 1, 2, \ldots, m$.

Далее в курсах теоретической механики доказывается следующая лемма.

Лемма 7.2.2. *Квадратичная форма кинетической энергии* $T \equiv T_2$ *склерономной системы невырождена.*

7.3. Движение по орбитам. Конические сечения [22, 37]

В небесной механике большое значение имеет так называемая ***«задача двух тел»*** — задача о совместном движении двух материальных точек, взаимодействующих по закону всемирного тяготения. Эта задача является частным случаем ***«основной задачи небесной механики»*** [22].

Основная физическая модель небесной механики представляет собой эволюционирующую (во времени) в абсолютно пустом пространстве изолированную систему материальных точек постоянной массы, взаимодействующих между собой по закону Всемирного тяготения Ньютона [37]. В соответствии с терминологией теории систем так называемым ***носителем*** основной математической модели небесной механики является система упорядоченных пар $\left\{m_k, \vec{r}_k\right\}$, где m_k — масса материальной точки с номером k, а $\vec{r}_k$ — её радиус-вектор относительно фиксированной системы координат. ***Сигнатурой*** модели является множество внутренних сил, действующих между точками системы в соответствии с законом Всемирного тяготения И. Ньютона. Напомним известную ещё из элементарного курса физики формулировку закона ***Всемирного тяготения*** И. Ньютона.

Каждая материальная частица во Вселенной притягивает каждую другую материальную частицу с силой, прямо пропорциональной произведению масс этих частиц, обратно пропорциональной квадрату расстояния между ними и направленной по прямой линии, соединяющей взаимодействующие частицы.

Следовательно, для двух материальных частиц, расположенных на расстоянии r, имеет место соотношение

$$F_{12} = f\frac{m_1 m_2}{r^2}, \tag{7.3.1}$$

где m_1 и m_2 — массы материальных частиц, F_{12} — сила взаимодействия частиц, r — расстояние между ними, $f = 6{,}67\cdot 10^{-11}$ м$^3\cdot$кг$^{-1}\cdot$с$^{-2}$ — универсальная гравитационная постоянная. Формула (7.3.1) может быть записана в векторном виде

$$\vec{F}_{12} = f\frac{m_1 m_2}{r_{12}^3}\vec{r}_{12}, \tag{7.3.2}$$

где $\vec{r}_{12}$ — вектор, направленный от точки 1 к точке 2 (см. рис. 7.3). Система уравнений второго закона И. Ньютона для рассматриваемого случая имеет вид

$$m_k\frac{d^2\vec{r}_k}{dt^2} = \vec{F}_k. \tag{7.3.3}$$

Используя выражение (7.3.1), систему уравнений второго закона Ньютона (7.3.3) для двух точек, взаимодействующих по закону Всемирного тяготения, запишем в виде

$$m_1 \frac{d^2 \vec{R}_1}{dt^2} = f \frac{m_1 m_2}{r_{12}^3} \vec{r}_{12}, \qquad (7.3.4)$$

$$m_2 \frac{d^2 \vec{R}_2}{dt^2} = f \frac{m_1 m_2}{r_{12}^3} \vec{r}_{21}. \qquad (7.3.5)$$

Сила

$$\vec{F}_{12} = f \frac{m_1 m_2}{r_{12}^3} \vec{r}_{12}$$

действует на точку 1 с массой m_1 и направлена по вектору $\vec{r}_{12}$ от точки 1 к точке 2 (рис. 7.3).

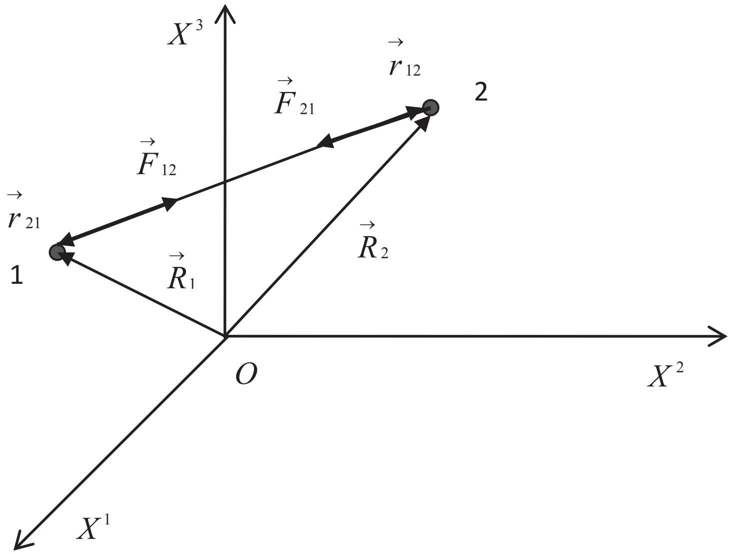


Рис. 7.3

Сила

$$\vec{F}_{21} = f \frac{m_1 m_2}{r_{12}^3} \vec{r}_{21}$$

действует на точку 2 с массой m_2 и направлена по вектору $\vec{r}_{21}$ от точки 2 к точке 1 (рис. 7.3). По третьей аксиоме динамики эти силы равны по норме и противоположны по направлению.

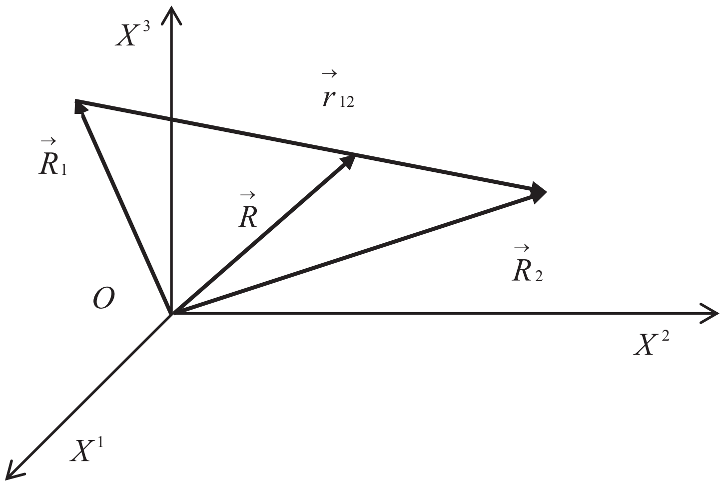


Рис. 7.4

Сложим уравнения (7.3.4) и (7.3.5) и, учитывая, что $\left\|\vec{r}_{12}\right\| = \left\|\vec{r}_{21}\right\|$ и $\vec{r}_{12} = -\vec{r}_{21}$, получаем (рис. 7.4)

$$m_1 \frac{d^2 \vec{R}_1}{dt^2} + m_2 \frac{d^2 \vec{R}_2}{dt^2} = \vec{0},$$

откуда имеем два интеграла

$$m_1 \frac{d \vec{R}_1}{dt} + m_2 \frac{d \vec{R}_2}{dt} = \vec{a}, \qquad (7.3.6)$$

$$m_1 \vec{R}_1 + m_2 \vec{R}_2 = \vec{a}\, t + \vec{b}, \qquad (7.3.7)$$

где $\vec{a}$ и $\vec{b}$ — постоянные векторы.

Определим радиус-вектор центра масс частиц m_1 и m_2 формулой

$$\vec{R} = \frac{m_1 \vec{R}_1 + m_2 \vec{R}_2}{m_1 + m_2},$$

тогда, вводя обозначение

$$M = m_1 + m_2,$$

получаем, что

$$M \vec{R} = m_1 \vec{R}_1 + m_2 \vec{R}_2,$$

и уравнения (7.3.6) и (7.3.7) перепишутся в виде

$$M\frac{d\vec{R}}{dt}=\vec{a}, \tag{7.3.8}$$

$$M\vec{R}=\vec{a}t+\vec{b}. \tag{7.3.9}$$

Из полученных соотношений следует, что центр масс движется с постоянной скоростью.

Уравнения (7.3.4) и (7.3.5) перепишем в виде

$$\frac{d^2\vec{R}_1}{dt^2}=fm_2\frac{\vec{r}}{r^3}, \tag{7.3.10}$$

$$\frac{d^2\vec{R}_2}{dt^2}=-fm_1\frac{\vec{r}}{r^3}, \tag{7.3.11}$$

где учтено, что

$$\vec{r}_{12}=-\vec{r}_{21}$$

и обозначено $\vec{r}_{12}\equiv\vec{r}$. Вычитая из уравнения (7.3.10) уравнение (7.3.11) почленно, получим

$$\frac{d^2}{dt^2}\left(\vec{R}_1-\vec{R}_2\right)=f\left(m_1+m_2\right)\frac{\vec{r}}{r^3}.$$

Так как $\vec{R}_1-\vec{R}_2=-\vec{r}$, имеем

$$\frac{d^2\vec{r}}{dt^2}+f\left(m_1+m_2\right)\frac{\vec{r}}{r^3}=\vec{0},$$

или обозначая

$$f\left(m_1+m_2\right)\overset{def}{=}\mu,$$

получим

$$\frac{d^2\vec{r}}{dt^2}+\mu\frac{\vec{r}}{r^3}=\vec{0}. \tag{7.3.12}$$

Находя векторное произведение обеих частей уравнения (7.3.12) и вектора $\vec{r}$, получаем

$$\left[\vec{r}, \frac{d^2 \vec{r}}{dt^2}\right] = \vec{0}.$$

Преобразуя левую часть полученного равенства

$$\left[\vec{r}, \frac{d^2 \vec{r}}{dt^2}\right] = \frac{d}{dt}\left[\vec{r}, \frac{d\vec{r}}{dt}\right] - \left[\frac{d\vec{r}}{dt}, \frac{d\vec{r}}{dt}\right] = \frac{d}{dt}\left[\vec{r}, \frac{d\vec{r}}{dt}\right],$$

имеем

$$\frac{d}{dt}\left[\vec{r}, \frac{d\vec{r}}{dt}\right] = \vec{0}.$$

Интегрируя последнее равенство, приходим к результату:

$$\left[\vec{r}, \frac{d\vec{r}}{dt}\right] = \vec{h}. \qquad (7.3.13)$$

Соотношение (7.3.13) является ***интегралом кинетического момента*** (***момента импульса***). Так как вектор $\vec{h}$ постоянен во времени, то движение частиц m_1 и m_2 происходит в плоскости, для которой этот вектор является нормальным вектором. Эта плоскость называется ***плоскостью эклиптики***.

Помещая начало системы координат в точку m_1 и вводя ***полярную систему координат*** (рис. 7.5), можно продолжить решение задачи.

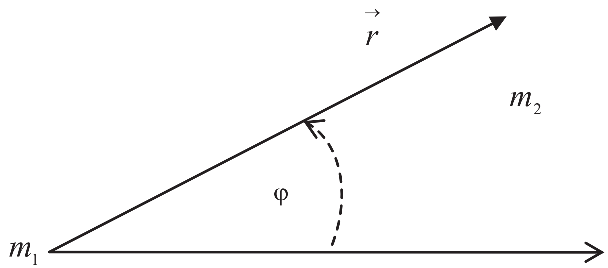


Рис. 7.5

Как известно из дифференциальной геометрии [35], проекции скорости на полярные оси равны $\dot{r}$ и $r\dot{\varphi}$. Тогда в полярной системе координат

$$\frac{d\vec{r}}{dt} = \dot{r}\vec{g}_{\langle r \rangle} + r\dot{\varphi}\vec{g}_{\langle \phi \rangle}, \tag{7.3.14}$$

где $\vec{g}_{\langle r \rangle}$ и $\vec{g}_{\langle \varphi \rangle}$ — орты ***радиальной*** и ***трансверсальной*** осей полярной системы координат соответственно. Проводя преобразования уравнения (7.3.12) в полярной системе координат, можно привести уравнение к виду [22]

$$\frac{d^2u}{d\varphi^2} + u = \frac{\mu}{h^2}, \tag{7.3.15}$$

где $\mu = f \cdot (m_1 + m_2)$, f — гравитационная постоянная, а $u = \dfrac{1}{r}$.

Решение уравнения (7.3.15) имеет вид [22]

$$u = \frac{\mu}{h^2} + A\cos(\varphi - \omega),$$

где A и ω — постоянные интегрирования. Переходя к переменной $r = \dfrac{1}{u}$, получим

$$r = \frac{h^2/\mu}{1 + \dfrac{Ah^2}{\mu}\cos(\varphi - \omega)}. \tag{7.3.16}$$

Если ввести обозначения $p = \dfrac{h^2}{\mu}$ и $\varepsilon = \dfrac{Ah^2}{\mu}$, то получаем уравнение ***конического сечения*** в полярной системе координат

$$r = \frac{p}{1 + \varepsilon\cos(\varphi - \omega)}. \tag{7.3.17}$$

Поясним приведённый результат с точки зрения геометрии.

Коническим сечением называется кривая линия, по которой пересекает круговой конус произвольная плоскость, не проходящая через его вершину (см. рис. 7.6).

Каждое коническое сечение, кроме окружности, представляет собой геометрическое место точек плоскости, отношение расстояний которых от некоторой точки F и некоторой прямой L постоянно. Доказательство этого свойства можно найти в книге [32]. Точка F называется ***фокусом*** конического сечения, а прямая L — его ***директрисой***.

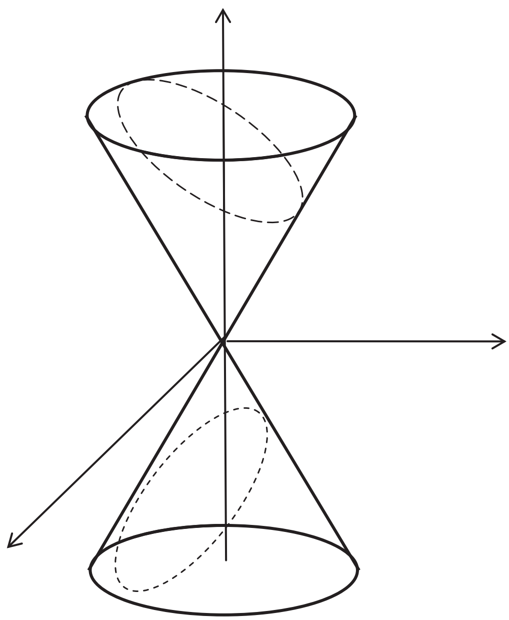


Рис. 7.6

В зависимости от величины отношения ε расстояния произвольной точки конического сечения от фокуса до директрисы, кривая линия называется ***эллипсом*** при $\varepsilon<1$, ***параболой*** при $\varepsilon=1$ и ***гиперболой*** при $\mu>1$. Само число ε называется ***эксцентриситетом*** конического сечения.

Составим уравнение конического сечения в полярной системе координат. Для этого обозначим буквой *p* расстояние между фокусом и директрисой. Расстояние произвольной точки *A* конического сечения от фокуса обозначим буквой *r*, а расстояние точки *A* от директрисы $p-r\cos\varphi$, или $r\cos\varphi-p$ в зависимости от того, как расположены точки *A* и F — по одну сторону директрисы или по разные стороны. Тогда уравнение конического сечения принимает вид

$$\frac{r}{p-r\cos\varphi}=\varepsilon \qquad (7.3.18)$$

в случае эллипса и параболы, и

$$\frac{r}{p-r\cos\varphi}=\pm\varepsilon \qquad (7.3.19)$$

в случае гиперболы. В последнем случае разные знаки при эксцентриситете соответствуют разным ветвям гиперболы.

Если из уравнений (7.3.18) и (7.3.19) выразить расстояние r, то получим

$$r = \frac{\varepsilon p}{1 + \varepsilon \cos\varphi} \tag{7.3.20}$$

— уравнения эллипса и параболы и

$$r = \frac{\pm\varepsilon p}{1 \pm \varepsilon \cos\varphi} \tag{7.3.21}$$

— уравнения гиперболы.

Очевидно, что уравнения (7.3.20) и (7.3.21) являются иной записью уравнения (7.3.17). Таким образом, можно сделать вывод о том, что при решении задачи двух тел орбиты материальных точек, взаимодействующих по закону всемирного тяготения И. Ньютона, могут быть ***эллиптическими***, ***параболическими*** или ***гиперболическими*** в зависимости от условий решаемой задачи. Анализ условий реализации вида орбиты проводится в курсах теоретической механики в рамках так называемой ***ограниченной задачи двух тел*** [22].

7.4. Законы Кирхгофа для электрических цепей [38]

Все электрические цепи подчиняются первому и второму ***законам Кирхгофа***, для каждого из которых возможны эквивалентные формулировки.

Первый закон Кирхгофа:

1) *алгебраическая сумма токов, притекающих к любому узлу, равна нулю;*

2) *сумма притекающих к любому узлу токов равна сумме утекающих от этого узла токов*.

Математическая запись первого закона Кирхгофа имеет следующий вид:

$$\sum_{k=1}^{n} I_k = 0. \tag{7.4.1}$$

Первый закон Кирхгофа для электрических цепей (см. рис. 7.7) является прямым следствием закона сохранения заряда

$$div\,\vec{j} = -\frac{\partial\rho}{\partial t},$$

который формулируется в курсе электродинамики.

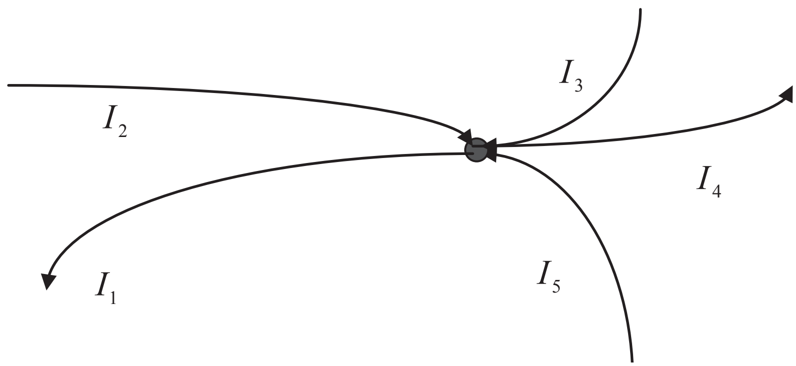


Рис. 7.7

Второй закон Кирхгофа:

1) *алгебраическая сумма падений напряжений в любом замкнутом контуре равна алгебраической сумме ЭДС вдоль того же контура, то есть*

$$\sum_{k=1}^{n} I_{k,k+1} R_{k,k+1} = \sum_{k=1}^{n} E_{k,k+1}^{стор}. \tag{7.4.2}$$

В формуле (7.4.2) нужно учитывать, что контур замкнут и, следовательно, точка под номером $k = n+1$ совпадает с точкой под номером $k = 1$. Для выяснения смысла этого замечания рассмотрим пример цепи квазилинейных токов с произвольным числом разветвлений (см. рис. 7.8).

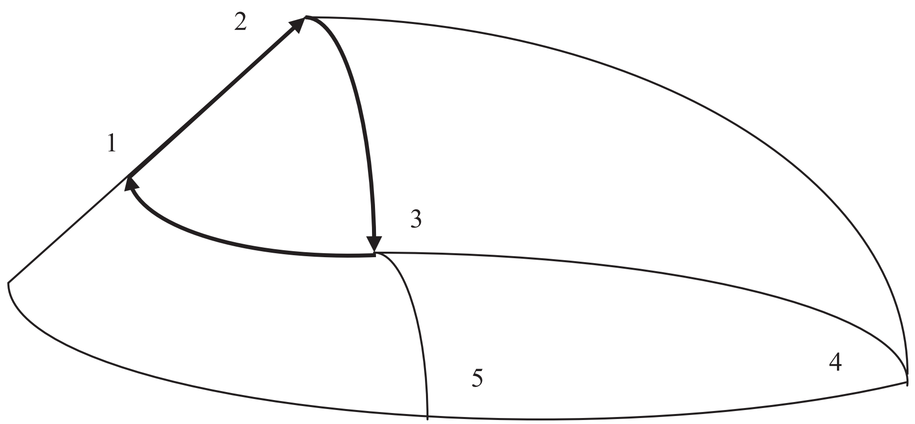


Рис. 7.8

Пусть контур L есть контур 1231. Обозначим $I_{k,k+1}$ силу тока на участке цепи с номерами k, $k+1$, например, на участке $1, 2$ сила тока обозначается $I_{1,2}$. Составляя для каждого участка цепи уравнение обобщённого закона Ома

$$I_{k,k+1}R_{k,k+1} = \varphi_k - \varphi_{k+1} + E_{k,k+1}^{стор}$$

и суммируя по всем участкам цепи, получаем уравнение

$$\sum_{k=1}^{n} I_{k,k+1}R_{k,k+1} = \sum_{k=1}^{n}\left(\varphi_k - \varphi_{k+1}\right) + \sum_{k=1}^{n} E_{k,k+1}^{стор}.$$

Учитывая, что

$$\sum_{k=1}^{n}\left(\varphi_k - \varphi_{k+1}\right) =$$
$$= \left(\varphi_1 - \varphi_2\right) + \left(\varphi_2 - \varphi_3\right) + \ldots + \left(\varphi_{n-1} - \varphi_n\right) + \left(\varphi_n - \varphi_1\right) = 0,$$

где учтено, что $n+1 \equiv 1$, получаем уравнение второго закона Кирхгофа (7.4.2).

Второй закон Кирхгофа можно сформулировать также следующим образом:

2) *алгебраическая сумма напряжений вдоль любого замкнутого контура равна нулю, то есть*

$$\sum_{ij} U_{ij} = 0. \qquad (7.4.3)$$

Приведём некоторые правила составления уравнений для расчёта цепей на основе законов Кирхгофа.

Обозначим число всех ветвей в схеме символом x; число ветвей, содержащих источники тока, символом x_i; число узлов символом y. Так как токи в ветвях с источниками тока известны, то число неизвестных токов равно $x - x_i$. Для составления уравнений предварительно нужно выбрать:

1) положительные направления токов в ветвях и обозначить их на схеме;

2) положительные направления обхода контуров для составления уравнений по второму закону Кирхгофа.

Для единообразия предполагается, что направления обхода всех контуров ориентированы одинаково.

Для получения линейно независимой системы уравнений по первому закону Кирхгофа составляются уравнения, число которых равно $y-1$. Уравнение для последнего узла не составляется, так как оно является линейной комбинацией предыдущих уравнений системы.

По второму закону Кирхгофа составляются уравнения, число которых равно $x - x_i$, минус число уравнений, составленных по первому закону Кирхгофа, то есть

$$(x - x_i) - (y - 1) = x - x_i - y + 1.$$

При составлении уравнений по второму закону Кирхгофа нужно учесть все ветви схемы, за исключением ветвей с источниками тока.

В качестве примера применения теории СЛАУ для расчёта электрических цепей постоянного тока рассмотрим математическую модель электрической цепи постоянного тока, составленную на основе законов Кирхгофа.

Пример 7.4.1. Для элементов электрической цепи, схема которой приведена на рис. 7.9, приняты следующие значения:

$E_1 = 80$ В, $E_2 = 64$ В, $R_1 = 6$ Ом, $R_2 = 4$ Ом, $R_3 = 3$ Ом, $R_4 = 1$ Ом.

Требуется найти токи в ветвях схемы.

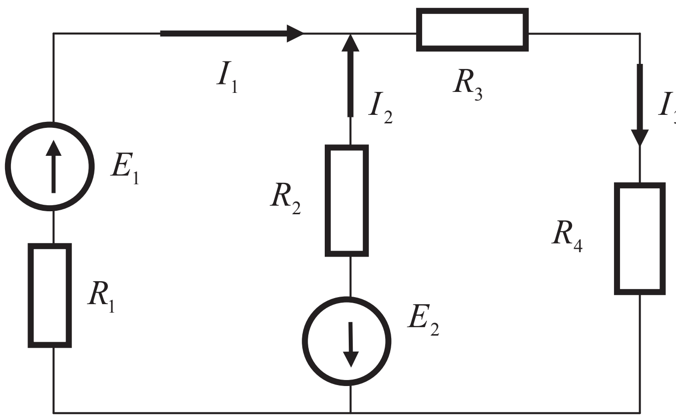


Рис. 7.9

Р е ш е н и е. В схеме выбираем ветви. В рассматриваемом случае

$$x = 3,\ x_i = 0,\ y = 2.$$

В ветвях схемы произвольно выбираем положительные направления токов.

По первому закону Кирхгофа можно составить одно уравнение

$$I_1 + I_2 = I_3. \qquad (1)$$

По второму закону Кирхгофа составляем уравнения в количестве

$$\left(x-x_i\right)-\left(y-1\right)=3-0-1=2.$$

Положительные направления обхода контуров выбираем по часовой стрелке.

Для контура $R_1E_1R_2E_2$ уравнение имеет вид:

$$I_1R_1+I_2R_2=E_1-E_2. \qquad (2)$$

Для контура $E_2R_2R_3R_4$ уравнение имеет вид:

$$I_2R_2-I_3R_3\left(R_3+R_4\right)=-E_2. \qquad (3)$$

Объединяя уравнения (1), (2) и (3) в систему, получаем:

$$\begin{cases} I_1+I_2-I_3=0, \\ R_1I_1+R_2I_2=E_1-E_2 \\ R_2I_2-R_3\left(R_3+R_4\right)I_3=-E_2. \end{cases} \qquad (4)$$

В матричной форме система уравнений имеет вид

$$\begin{pmatrix} 1 & 1 & -1 \\ R_1 & R_2 & 0 \\ 0 & R_2 & -R_3\left(R_3+R_4\right) \end{pmatrix}\begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix}=\begin{pmatrix} 0 \\ E_1-E_2 \\ -E_2 \end{pmatrix}.$$

Нетрудно видеть, что для принятых значений параметров электрической цепи

$$\det\begin{pmatrix} 1 & 1 & -1 \\ R_1 & R_2 & 0 \\ 0 & R_2 & -R_3\left(R_3+R_4\right) \end{pmatrix}=\det\begin{pmatrix} 1 & 1 & -1 \\ 6 & 4 & 0 \\ 0 & 4 & -12 \end{pmatrix}=-20\neq 0.$$

Таким образом, система совместна и определённа.

Решение системы уравнений (4), например по формулам Крамера, даёт следующие значения токов:

$$I_1=14\text{ A},\ I_2=-15\text{ A},\ I_3=-1\text{ A}.$$

Получившиеся отрицательные значения токов I_2 и I_3 означают, что принятые произвольно положительные направления токов не соответствуют истинным направлениям токов I_2 и I_3. Последние протекают в противоположном направлении. ⊗

7.5. Представление чистых состояний в квантовой механике векторами в унитарном пространстве [8, 9, 11, 15, 23]

В качестве примера применения теории векторных пространств в квантовой механике рассмотрим ***микроскопическую материальную систему***, полагая, что некоторая её физическая характеристика может принимать конечное число дискретных значений, которые называются ***квантовыми числами***.

Такая физическая модель системы является идеализированной по причине того, что значения многих физических величин в квантовой механике могут заполнять, например, некоторый промежуток числовой оси, то есть физические величины могут принимать «непрерывный ряд» значений. В этом случае для описания эволюции системы требуется применение более сложной теории линейных операторов, которая изучается в функциональном анализе.

Сначала рассмотрим опыт (эксперимент) Штерна — Герлаха, который лёг в основу интерпретации квантовой механики и послужил, по существу, мотивацией для применения аппарата теории векторных пространств и операторов в квантовой механике. Этот опыт (эксперимент) был осуществлён в 1922 году немецкими физиками О. Штерном и В. Герлахом.

Эксперимент состоял в следующем: пучок атомов серебра пропускали через сильно неоднородное магнитное поле, создаваемое мощным постоянным магнитом. При прохождении атомов через магнитное поле, в силу наличия у них магнитных моментов, на них действовала зависящая от проекции магнитного момента на направление магнитного поля сила, отклонявшая атомы от их первоначального направления движения. Причём, если предположить, что магнитные моменты атомов ориентированы хаотично (непрерывно), то тогда на расположенном далее по направлению движения атомов экране (фотопластинке) должна была проявиться размытая полоса. Однако вместо этого на пластинке образовались две достаточно чёткие узкие полосы, что свидетельствовало в пользу того, что магнитные моменты атомов вдоль выделенного направления принимали лишь два определённых значения. Таким образом, опыт подтвердил квантование проекции вектора магнитного момента атомов, а также стал одним из главных аргументов в пользу существования у электронов собственного магнитного момента и связанного с ним момента импульса, который получил название ***спин***.

Схема опыта Штерна — Герлаха приведена на рис. 7.10. В колбе с вакуумом 10^{-5} мм рт. ст. нагревался серебряный шарик S до температуры испарения. Атомы серебра, пролетая с тепловой скоростью (около 100 м/с) через щелевые диафрагмы и проходя резко неоднородное магнитное поле определённой ориентации, то есть через поле очень мощного постоянного магнита $N - S$, попадали на фотопластинку — экран.

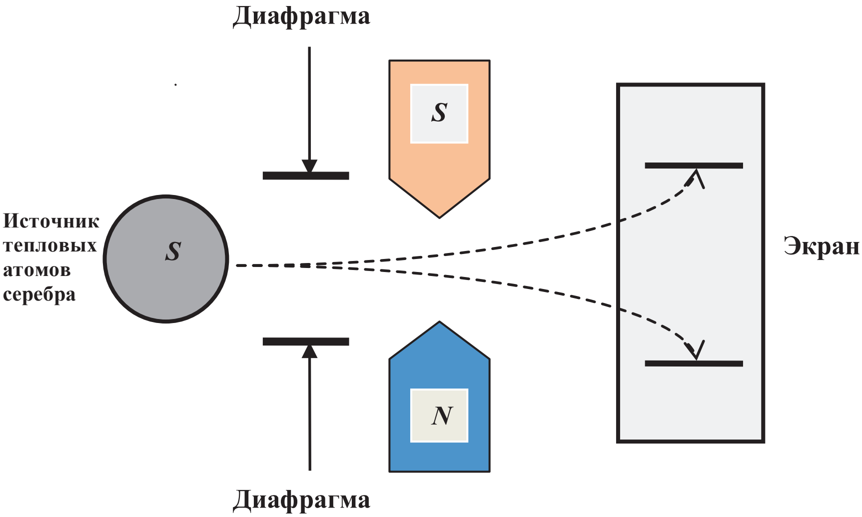


Рис. 7.10

Если бы момент импульса атома (и его магнитный момент) мог принимать в магнитном поле произвольное направление, то можно было бы ожидать непрерывного распределения попаданий атомов серебра на экран с большой плотностью попаданий в середине. Но на опыте был получен совершенно другой результат: на фотопластинке получились две резкие полосы — все атомы отклонялись в магнитном поле двояким образом, соответствующим лишь двум возможным ориентациям магнитного момента.

Таким образом, опыт Штерна — Герлаха доказал квантовый характер магнитных моментов электронов. В дальнейшем количественный анализ показал, что проекция магнитного момента электрона на направление магнитного поля равна ***магнетону Бора***:

$$\mu_B = \frac{e\hbar}{2m_e} = 9,27 \cdot 10^{-24} \text{ Дж Тл}^{-1}.$$

Таким образом, для атомов серебра из опыта Штерна-Герлаха следует, что проекция магнитного момента атома (электрона) на направление магнитного поля численно равна магнетону Бора.

Опыты Штерна и Герлаха не только подтвердили пространственное квантование моментов импульсов в магнитном поле, но и дали экспериментальное подтверждение тому, что магнитные моменты атомов тоже состоят из некоторого числа «элементарных моментов», то есть имеют дискретную природу. Единицей измерения магнитных моментов электронов и атомов является магнетон Бора ($\hbar$ — единица измерения механического момента импульса).

Кроме этого, в этих опытах было обнаружено новое явление. Валентный электрон в основном состоянии атома серебра имеет орбитальное квантовое число $l = 0$ (s-состояние). Но при $l = 0$ проекция момента импульса на направление внешнего поля также равна нулю. Возник вопрос, пространственное квантование какого момента импульса обнаружилось в этих опытах и проекция какого магнитного момента равна магнетону Бора.

В 1925 г. Гаудсмит и Уленбек предположили существование собственного механического момента импульса у электрона — ***спина*** и, соответственно, собственного магнитного момента электрона P_{ms}.

Введение понятия спина сразу объяснило ряд затруднений, имевшихся к тому времени в квантовой механике. И в первую очередь результаты опытов Штерна и Герлаха.

Гаудсмит и Уленбек отождествили спин электрона с механическим моментом вращающегося волчка, которым по их представлениям и был электрон. Но тогда «поверхность» волчка (электрона) должна вращаться с линейной скоростью, равной 300 c, где c — скорость света. От такого толкования спина пришлось отказаться.

П. Дирак впоследствии показал, что существование спина вытекает из решения релятивистского волнового уравнения Шредингера и является, таким образом, свойством электрона, таким же, как заряд и масса. В настоящее время известно, что спином обладают все элементарные частицы.

Перейдём к математическому оформлению приведённых экспериментальных данных.

Математическую интерпретацию проведём на примере простейшей физической системы с дискретным набором квантовых чисел, которая называется ***дихотомической*** системой. В такой системе некоторая физическая величина имеет дихотомический характер, то есть может принимать только два значения. Примерами дихотомической физической величины являются проекция спина электрона на направление внешнего магнитного поля, заряд нуклона, число фермионов в данном квантовом состоянии и так далее. Опыт Штерна — Герлаха демонстрирует как раз поведение дихотомической системы.

Пусть пучок атомов со спином $s = \frac{1}{2}$, ориентированный вдоль направления поля, задаваемого углами θ и φ. В квантовой механике состояние частиц в такой ситуации описывается нормированным ***вектором состояния*** в унитарном пространстве U^2:

$$|a_1\rangle = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}. \tag{7.5.1}$$

Здесь $\alpha_1 = \alpha_1(\theta, \varphi)$, $\beta_1 = \beta_1(\theta, \varphi)$.

Норма вектора $|a_1\rangle$ есть квантовое число, соответствующее проекции спина «на направление», задаваемое углами θ и φ. Из требования нормированности вектора состояния (7.5.1) следует, что скалярный квадрат

$$\langle a_1 | a_1 \rangle = \begin{pmatrix} \alpha_1^* & \beta_1^* \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} = |\alpha_1|^2 + |\beta_1|^2 = 1. \tag{7.5.2}$$

Условие (7.5.2) позволяет дать вероятностную трактовку величин $|\alpha_1|^2$ и $|\beta_1|^2$ в опыте Штерна — Герлаха. Действительно, если пучок атомов со спином $s = \frac{1}{2}$, ориентированный вдоль направления, задаваемого углами θ и φ, и описываемый вектором состояния $|a_1\rangle$, повторно пропустить через установку Штерна — Герлаха, направив магнитное поле по оси OX_3, то в результате мы получим, что спины одних частиц направлены вдоль оси OX_3, а спины других частиц направлены в противоположном направлении. Соотношение числа частиц с разнонаправленными спинами можно предсказать только статистически. Таким образом, норму вектора состояния можно интерпретировать так: квадрат модуля первой компоненты $|\alpha_1|^2$ равен вероятности обнаружения спина данной частицы, направленной в положительном направлении оси OX_3, а квадрат модуля второй компоненты $|\beta_1|^2$ равен вероятности обнаружения спина данной частицы, направленной в отрицательном направлении оси OX_3.

При такой интерпретации векторы

$$|a_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |a_-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

описывают состояния, когда проекция спина на положительное направление оси OX_3 с достоверностью принимает значение $+\frac{1}{2}$ и $-\frac{1}{2}$ соответственно. Эти векторы составляют базис пространства U^2. Следовательно, если вектор состояния

$$|a_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

описывает состояние, в котором спин достоверно направлен в положительном направлении оси OX_3, а вектор состояния

$$|a_1\rangle = \begin{pmatrix} \pm_1 \\ {}_2 \quad {}_1 \end{pmatrix}$$

описывает состояние, в котором спин достоверно направлен в направлении, описываемом углами θ и φ, в силу того, что

$$\langle a_+ | a_+ \rangle^{1/2} = 1,$$

скалярное произведение

$$\langle a_+ | \alpha_1 \rangle = 1 \cdot \alpha_1 + 0 \cdot \beta_1 = \alpha_1.$$

И, следовательно, $|\alpha_1|^2$ является вероятностью обнаружения спина $s = +\frac{1}{2}$ у объекта, находящегося в состоянии $|a_1\rangle$.

Таким образом, в случае квантовой дихотомической системы для описания экспериментов типа опыта Штерна — Герлаха, естественно использование геометрии двумерного унитарного пространства с ортонормированным базисом, а именно пространства U^2 (комплексное евклидово пространство).

Подобная вероятностная интерпретация может быть обобщена и на более общий случай. При таком обобщении проекция вектора состояния

$$|a\rangle = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}^T$$

на состояние

$$|b\rangle = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix}^T$$

даёт амплитуду вероятности обнаружения в результате измерения квантового числа b, если объект первоначально находился в состоянии $|a\rangle$. В силу нормированности векторов состояния проекция равна

$$\langle b|a\rangle = b_1^* a_1 + b_2^* a_2 + \ldots + b_n^* a_n = \langle a|b\rangle^*. \qquad (7.5.3)$$

Постулируется, что по аналогии с результатами опыта Штерна — Герлаха величина $|\langle a|b\rangle|^2$ интерпретируется как вероятность обнаружения квантового числа b, если объект находится в квантовом состоянии с квантовым числом a.

Таким образом, в случае квантовой физической системы с большим числом размерностей, нежели 2, естественно использовать геометрию конечномерного унитарного пространства с ортонормированным базисом, а именно пространства U^n.

7.6. Наблюдаемые величины в квантовой механике [8, 9, 11, 15, 23]

Как и выше, рассматриваем только случай системы с конечномерным пространством состояний.

Наблюдаемой в квантовой механике называют любую физическую величину, значения которой могут быть найдены в эксперименте, причём измеренные в эксперименте значения наблюдаемой величины обязательно являются вещественными, то есть принадлежат полю действительных чисел. Сам процесс измерения влияет на значения наблюдаемых величин, что должно учитываться при построении математического аппарата, позволяющего рассчитывать значения наблюдаемых величин.

В качестве математического аппарата, позволяющего учесть влияние процесса измерения на состояние системы и вычислить среднее значение наблюдаемой величины A в любом данном состоянии $|b\rangle$, в квантовой механике используется теория линейных эрмитовых (самосопряжённых) операторов. В рамках этой теории каждой наблюдаемой величине A ставится в соответствие линейный эрмитов оператор $\hat{A}$, собственные векторы которого являются векторами состояния системы, в которых наблюдаемая величина A принимает определённые значения a_k. Причём указанное соответствие осуществляется таким образом, чтобы значения, принимаемые наблюдаемой величиной, были собственными значениями оператора $\hat{A}$:

$$\hat{A}|a_k\rangle = a_k|a_k\rangle. \tag{7.6.1}$$

Как известно, все корни характеристического многочлена эрмитова (самосопряжённого) оператора различные, а соответствующие собственные векторы ортогональны. Собственные векторы можно нормировать и получить ортонормированный базис собственных векторов оператора. Удобство представления наблюдаемой величины A линейным эрмитовым оператором $\hat{A}$ из уравнения (7.6.1) подтверждается следующей теоремой.

Теорема 7.6.1. *Пусть $\hat{A}$ — эрмитов оператор и a_k — его собственные значения из уравнения (7.6.1). Тогда проекция вектора состояния*

$$|c\rangle = \hat{A}|b\rangle$$

на вектор состояния $|b\rangle$ есть среднее значение наблюдаемой величины A в состоянии $|b\rangle$, то есть

$$\overline{A} = \langle b|c\rangle = \langle b|\hat{A}|b\rangle. \tag{7.6.2}$$

Д о к а з а т е л ь с т в о. Разложим вектор $|b\rangle$ по векторам базиса

$$\{|a_1\rangle, |a_2\rangle, \ldots, |a_n\rangle\},$$

воспользовавшись ортогональным разложением (2.5.12):

$$|b\rangle = \langle a_1|b\rangle|a_1\rangle + \langle a_2|b\rangle|a_2\rangle + \ldots + \langle a_n|b\rangle|a_n\rangle. \tag{7.6.3}$$

В разложении (7.6.3) величины $\langle a_k|b\rangle$ — это амплитуды вероятности обнаружения значения a_k наблюдаемой A в состоянии $|b\rangle$. Действуя оператором $\hat{A}$ на обе части разложения (7.6.3) и учитывая соотношения (7.6.1), получаем

$$\hat{A}|b\rangle = \langle a_1|b\rangle \hat{A}|a_1\rangle + \langle a_2|b\rangle \hat{A}|a_2\rangle + \ldots + \langle a_n|b\rangle \hat{A}|a_n\rangle =$$
$$= a_1|a_1\rangle\langle a_1|b\rangle + a_2|a_2\rangle\langle a_2|b\rangle + \ldots + a_n|a_n\rangle\langle a_n|b\rangle.$$

Находя значение скалярного произведения образа вектора $\hat{A}|b\rangle$ и самого вектора $|b\rangle$, получаем

$$\langle b|\hat{A}|b\rangle = a_1\langle b|a_1\rangle\langle a_1|b\rangle + a_2\langle b|a_2\rangle\langle a_2|b\rangle + \ldots + a_n\langle b|a_n\rangle\langle a_n|b\rangle =$$
$$= a_1|\langle a_1|b\rangle|^2 + a_2|\langle a_2|b\rangle|^2 + \ldots + a_m|\langle a_m|b\rangle|^2.$$

Последнее выражение является суммой произведений значений a_k случайной величины A на соответствующие им вероятности, и по определению есть математическое ожидание величины A, то есть её среднее значение $\overline{A}$. ••

Напомним, что физическое требование вещественности среднего значения $\overline{A}$ наблюдаемой величины A для эрмитовых операторов следует из того, что все собственные значения эрмитова оператора вещественны, а модуль $|\langle a_k|b\rangle|$ есть по определению величина вещественная. Можно этот простой факт доказать и при помощи следующей выкладки:

$$\overline{A} = \langle b|\hat{A}|b\rangle = \langle A^*b|b\rangle = \langle b|A^*b\rangle = \langle b|A^*|b\rangle^* = \langle b|A|b\rangle^* = \overline{A}^*.$$

Напомним также, в соответствии с теоремой 3.1.5 о конструкции линейного оператора образ $|c\rangle = \hat{A}|b\rangle$ может быть записан в векторно-матричной форме, а именно:

$$\begin{pmatrix} c^1 \\ c^2 \\ \ldots \\ c^n \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_m^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ a_1^n & a_2^n & \ldots & a_m^n \end{pmatrix} \begin{pmatrix} b^1 \\ b^2 \\ \ldots \\ b^m \end{pmatrix}.$$

На этом мы закончим краткий обзор применений теории конечномерных векторных пространств и действующих в них линейных операторов. С дальнейшими применениями этой теории читатель сможет ознакомиться в специальных дисциплинах.

Библиографический список

1. Александров П. С. Введение в теорию множеств и общую топологию / П. С. Александров. — Москва : Наука, 1977. — 368 с.
2. Архангельский А. В. Конечномерные векторные пространства / А. В. Архангельский. — Москва : МГУ, 1982. — 248 с.
3. Винберг Э. Б. Курс алгебры / Э. Б. Винберг. — Москва : ЛитРес МЦНМО, 2015. — 592 с. — ISBN 5–88688–050-X.
4. Воеводин В. В. Линейная алгебра / В. В. Воеводин. — Москва : Лань, 2009. — 400 с. — ИБ 11654.
5. Гантмахер Ф. Р. Теория матриц / Ф. Р. Гантмахер. — Москва : Наука, 1988. — 552 с. — ISBN 5-02-013722-7.
6. Головина Л. И. Линейная алгебра и некоторые её приложения / Л. И. Головина. — Москва : Наука, 1985. — 392 с. — ИБ 12629.
7. Грауэрт Г. Дифференциальное и интегральное исчисление / Г. Грауэрт, И. Либ, В. Фишер. — Москва : Мир, 1971. — 680 с.
8. Дирак П. А. М. Принципы квантовой механики / П. А. М. Дирак. — Москва : Наука, 1979. — 480 с. — ISBN 978-5-45833-292-7.
9. Зайцев Г. А. Алгебраические проблемы математической и теоретической физики / Г. А. Зайцев. — Москва : Наука, 1974. — 191 с.
10. Зельдович Я. Б. Высшая математика для начинающих и ее приложения к физике / Я. Б. Зельдович. — Москва : Наука, 2010. — 520 с. — ISBN 978-5-9221-1667-1.
11. Кемпфер Ф. Основные положения квантовой механики / Ф. Кемпфер. — Москва : Мир, 1967. — 391 с.
12. Кострикин А. И. Введение в алгебру / А. И. Кострикин. — Москва : Наука, 1977. — 496 с.
13. Кострикин А. И. Линейная алгебра и геометрия / А. И. Кострикин, Ю. И. Манин. — Москва : Наука, 1986. — 304 с. — ИБ 12692.
14. Кострикин А. И. Сборник задач по алгебре // А. И. Кострикин ; под ред. А. И. Кострикина. — Москва : Факториал, 1995. — 454 с. — ISBN 5-88688-001-1 (Факториал) ; ISBN 985-410-024-3 (Просперус).
15. Мессиа А. Квантовая механика. Т. I / А. Мессиа. — Москва : Наука, 1978. — 480 с. — ИБ 2264
16. Никольский С. М. Курс математического анализа. Т. I / С. М. Никольский. — Москва : Наука, 1990. — 528 с. — ISBN 5-02-014424-X.

17. Никольский С. М. Курс математического анализа. Т. II / С. М. Никольский. — Москва : Наука, 1991. — 544 с. — ISBN 5-02-014425-8.
18. Пизо Ш. Курс математики. Алгебра и анализ / Ш. Пизо, М. Заманский. — Москва : Наука, 1971. — 656 с.
19. Погорелов А. В. Геометрия / А. В. Погорелов. — Москва : Наука, 1983. — 288 с. — ИБ 12350.
20. Размыслович Г. П. Сборник задач по геометрии и алгебре / Г. П. Размыслович, М. М. Феденя, В. М. Ширяев. — Минск : Універсітэцкае, 1999. — 383 с. — ISBN 985–09–0288–4.
21. Рашевский П. К. Риманова геометрия и тензорный анализ / П. К. Рашевский. — Москва : Наука, 1967. — 664 с.
22. Рой А. Движение по орбитам / А. Рой. — Москва : Мир, 1981. — 544 с.
23. Садбери А. Квантовая механика и физика элементарных частиц / А. Садбери. — Москва : Мир, 1989. — 488 с. — ISBN 5-03-000985-X.
24. Смирнов В. И. Курс высшей математики. Т. I / В. И. Смирнов. — Москва : Наука, 1974. — 480 с.
25. Смирнов В. И. Курс высшей математики. Т. II / В. И. Смирнов. — Москва : Наука, 1974. — 656 с.
26. Смирнов В. И. Курс высшей математики. Т. III. Ч. 1/ В. И. Смирнов. — Москва : Наука, 1974. — 324 с.
27. Смирнов В. И. Курс высшей математики. Т. III. Ч. 2 / В. И. Смирнов. — Москва : Наука, 1974. — 672 с.
28. Смирнов В. И. Курс высшей математики. Т. IV. Ч. 1 / В. И. Смирнов. — Москва : Наука, 1974. — 336 с.
29. Смирнов В. И. Курс высшей математики. Т. IV. Ч. 2 / В. И. Смирнов. — Москва : Наука, 1981. — 550 с.
30. Стренг Г. Линейная алгебра и её применения / Г. Стренг. — Москва : Мир, 1980. — 454 с.
31. Сурнев В. Б. Основы высшей математики. Ч. 1 и Ч. 2 / В. Б. Сурнев. — Екатеринбург : Изд-во УГГГА, 2000. — 400 с.
32. Сурнев В. Б. Алгебра и аналитическая геометрия / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 2003. — 656 с.
33. Сурнев В. Б. Основы высшей математики. Ч. 1. Алгебра и аналитическая геометрия / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 2006. — 191 с. ISBN 5-8019-0078-0.
34. Сурнев В. Б. Основы высшей математики. Ч. 2. Анализ функций одного действительного переменного / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 2006. — 133 с. — ISBN 5-8019-0080-2.

35. Сурнев В. Б. Дифференциальная геометрия / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 2007. — 186 с. — ISBN 5-8019-0142-6.
36. Сурнев В. Б. Основы высшей математики. Ч. 3. Анализ функций нескольких действительных переменных / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 2010. — 297 с. — ISBN 978-5-8019-0232-6.
37. Сурнев В. Б. Принципы классической механики (краткий курс для математиков-прикладников) / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 20011. — 153 с. — ISBN 978-5-8019-0278-4.
38. Сурнев В. Б. Математическое моделирование. Непрерывные детерминированные модели / В. Б. Сурнев. — Екатеринбург : Изд-во УГГУ, 2013. — 690 с. — ISBN 978-5-8019-0310-1
39. Фаддеев Д. К. Лекции по алгебре / Д. К. Фаддеев. — Москва : Наука, 1984. — 416 с. — ИБ 12076.
40. Фаддеев Д. К. Задачи по высшей алгебре / Д. К. Фаддеев, И. С. Соминский. — Санкт-Петербург : Лань, 1998. — 288 с. — ISBN 5-8114-0113-2.
41. Шварц Л. Анализ. Т. 1 / Л. Шварц. — Москва : Мир, 1972. — 824 с.
42. Шварц Л. Анализ. Т. 2 / Л. Шварц. — Москва : Мир, 1972. — 528 с.
43. Шилов Г. Е. Математический анализ. Конечномерные линейные пространства / Г. Е. Шилов. — Москва : Наука, 1969. — 432 с.
44. Шнеперман Л. Б. Сборник задач по алгебре и теории чисел / Л. Б. Шнеперман. — Минск : Дизайн ПРО, 2000. — 240 с. — ISBN 985-452-018-8.

СУРНЕВ ВИКТОР БОРИСОВИЧ

доктор физико-математических наук, заведующий кафедрой математики Уральского государственного горного университета (УГГУ), профессор кафедры физики высокоэнергетических процессов ФТИ УрФУ. Автор более 60 научных работ и 10 учебных пособий.

Специалист в области теории непрерывных динамических систем. Научные интересы: теория рассеяния акустических, упругих и электромагнитных волн; дифракционная томография; прямая и обратная задачи динамики многих тел в небесной механике; динамика непрерывных моделей экономических систем.