ГОСУДАРСТВЕННЫЙ КОМИТЕТ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПО ВЫСШЕМУ ОБРАЗОВАНИЮ

САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П.КОРОЛЕВА

А.Ф.ТАРАСКИН

СТАТИСТИЧЕСКИЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ АВТОРЕГРЕССИИ И СКОЛЬЗЯЩЕГО СРЕДНЕГО

УЧЕБНОЕ ПОСОБИЕ

САМАРА 1998
УДК 519.2 (075)

Статистический анализ временных рядов авторегрессии и скользящего среднего: Учебное пособие / А.Ф.Тараскин; Самар. гос. аэрокосм. ун-т. Самара, 1998. 64 с.
ISBN 5-230-16 956-7

Кратко излагаются основные факты теории случайных временных рядов авторегрессии и скользящего среднего. Рассматривается статистические задачи для процессов при условии их стационарности.

Предназначено для студентов специальности «Прикладная математика» при изучении курса «Случайные процессы» и при выполнению курсовой работы по этому курсу. Подготовлено на кафедре «Техническая кибернетика».

Ил.2 Библиогр.: 7 назв. Табл 1.

Печатается по решению редакционно-издательского совета
Самарского государственного аэрокосмического университета
имени академика С.П.Королева

Рецензенты: А.И.Жданов, В.М.Климкин

ISBN 5-230-16 956-7

ОГЛАВЛЕНИЕ

1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

1.1. Основные понятия и терминология

При исследовании реальных устройств, функционирующих в условиях случайных возмущений, экспериментатор может наблюдать и фиксировать реализации случайных процессов, связанных с работой устройства. При этом статистические закономерности процессов и параметры исследуемого устройства частично или полностью оказываются априори неизвестными. Поскольку получение точных значений интересующих характеристик и параметров, как правило, бывает невозможным, приходится оценивать из на основе обработки экспериментальных данных с учетом априорной информации, указывающей, например, класс к которому принадлежит исследуемый процесс.

Для широкого класса устройств модель функционирования может быть представлена как реакция на входные возмущения и начальное состояние. Модель, описывающую работу устройств как преобразование входных возмущений и начального состояния в выходную реакцию, называют *системой*.

Для математического описания системы удобно использовать принятую терминологию: так, входные возмущения и начальное состояние называют *входным сигналом*, реакцию системы - *выходным сигналом*. Входные и выходные сигналы в общем случае являются элементами произвольных пространств. Например, для механических устройств входными сигналами могут быть силы и моменты, а выходными - перемещения, скорости и ускорения. Для радиотехнических и электронных систем входными сигналами являются электромагнитные поля, токи и напряжения, а выходными - сигналы той же природы или звуковые сигналы, а возможно и телевизионные изображения. Для организационных систем в качестве входных сигналов можно рассматривать проблемы, а качестве выходных - решения проблем.

Обозначая входной сигнал через X, а выходной - через Y, можно схематически изобразить систему (рис.1).

Мы будем рассматривать стационарные (установившиеся) режимы функционирования систем, а это означает, что входной и

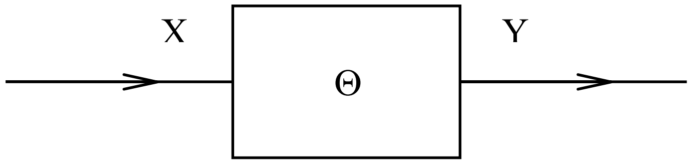


Рис. 1.

выходной процессы являются стационарными в широком смысле. Кроме этого, предположим, что X и Y, являются процессами с целочисленным временем: $X_t, Y_t, t = 0, \pm 1, \pm 2, \ldots$ Такие процессы чаше называются *случайными последовательностями* (с.п.) или временными рядами.

1.2. Элементы теории стационарных случайных процессов

С.п. $X_t, t = 0, \pm 1, \pm 2, \ldots,$ принимающая, вообще говоря, комплексные значения, называется стационарной в широком смысле, если для любого целого t $MX_t = m = \mathrm{const}$ и корреляционная функция

$$R(t-s) = M(X_t - m)\overline{(X_s - m)} = R_1(t-s)$$

зависит только от разности моментов времени t и s. Таким образом, корреляционная функция стационарной с.п. является комплекснозначной функцией целочисленного аргумента:

$$R(k), k = 0, \pm 1, \pm 2, \ldots$$

Она обладает следующими свойствами:

а.) $R(0) > 0$; если же X_t принимает только вещественные значения, то $R(0) = DX_t = \sigma_X^2$;

б.) $R(-k) = \overline{R(k)}$; если X_t вещественнозначная с.п., то $R(k) = R(-k)$;

в.) $R(k)$ неотрицательно определена, т.е. для любого целого $n > 0$, любых целых $k_1, k_2, \ldots, k_n$ и любого набора комплексных чисел $z_1, z_2, \ldots, z_n$ выполнятся неравенство

$$\sum_{r=1}^{n} \sum_{j=1}^{n} R(k_r - k_j) z_k \overline{z_j} \geq 0$$

Согласно теореме А.Я.Хинчина для корреляционной функции стационарной с.п. имеет представление

$$R(k)=\int_{-\pi}^{\pi} e^{ik\omega}dF(\omega) \qquad (1)$$

в котором $F(\omega)$- неубывающая неотрицательная ограниченная функция на $[-\pi,\pi]$, называемая, как и в случае процессов с непрерывным временем, спектральной функцией. Если

$$\sum_{k=-\infty}^{\infty}|R(k)|<\infty \qquad (2)$$

то функция $F(\omega)$ будет дифференцируемой, и она может быть представлена в виде

$$F(\omega)=\int_{-\pi}^{\omega} f(u)du \qquad (3)$$

где $f(\omega)=F'(\omega)\geq 0$. При этом (1) можно заменить формулой

$$R(k)=\int_{-\pi}^{\pi} e^{ik\omega}f(\omega)d\omega \qquad (4)$$

Функция $f(\omega)$, где $\omega\in[-\pi б\pi]$, называется *спектральной* плотностью случайной последовательности. Из (4) видно, что величины $\frac{1}{2\pi}R(-k), k=0,\pm1,\pm2,...$, являются коэффициентами Фурье функции $f(\omega)$, так что разложение этой функции в ряд Фурье будет иметь вид

$$f(\omega)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{-ik\omega}R(k) \qquad (5)$$

Эту формулу можно рассматривать как дискретный аналог известной формулы обращения для спектральной плотности непрерывного в среднем квадратическом (с.к.) стационарного в широком смысле процесса.

Используя теорему Карунена для стационарной последовательности $X_t, t=0,\pm1,\pm2,...$, с $MX_t=0$ и с корреляционной функцией (1), получаем интегральное представление (теорема Хинчина)

$$X_t=\int_{-\pi}^{\pi} e^{it\omega}dV_\omega, \qquad (6)$$

где в правой части имеем стохастический интеграл по процессу $V_\omega, \omega\in[-\pi,\pi]$с ортогональными приращениями и со спектральной функцией $F(\omega), \omega\in[-\pi,\pi]$, совпадающей со спектральной функцией в представлении (1) корреляционной функции последовательности X_t.

2. ПРОЦЕССЫ АВТОРЕГРЕССИИ И СКОЛЬЗЯЩЕГО СРЕДНЕГО

2.1. Значение процессов авторегрессии и скользящего среднего

Разнообразные данные в физике, космических исследованиях, экономике, медицине и других областях поступают исследователю в виде случайных последовательностей (временных рядов). Совокупность существующих методов изучения таких рядов зависимых наблюдений называется *анализом временных рядов*.

В начале 70-х годов многие зарубежные исследователи стали аппроксимировать изучаемый временной ряд наиболее близкой (например, с точки зрения асимптотической среднеквадратичной теории) последовательностью авторегрессии, скользящего среднего или смешанной последовательностью авторегрессии - скользящего среднего. Это позволило характеризовать целый ряд наблюдений всего несколькими параметрами. Простота структуры последовательностей авторегрессии и скользящего среднего и в то же время возможность использования их для аппроксимации широкого класса с.п. определяют как практический, так и теоретический интерес к ним. Моделирование этих последовательностей позволяет решать самые разнообразные прикладные задачи, связанные с изучением реальных процессов в науке и технике.

2.2 Случайные последовательности авторегрессии

Так называемая *модель авторегрессии* описывается стохастическим разностным уравнением вида

$$Y_t = \sum_{j=1}^{p} \beta_j Y_{i-j} + X_t, \left(\beta_j \neq 0, j = \overline{1,p}\right), \qquad (7)$$

где t пробегает целые значения, а X_t - последовательность вещественнозначных некоррелированных и одинаково распределенных $\left(MX_t = 0, DX_t = \sigma_X^2\right)$ случайных величин (с.в.). Последовательность Y_t называется с.п. *авторегрессии порядка p*, сокращенно АР(p). Большое практическое значение имеют последовательности авторегрессии первого (p=1) и (p=2) порядков

$$Y_t = \beta_1 Y_{t-1} + X_t, \qquad (8)$$

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + X_t. \qquad (9)$$

Характеристическим уравнением для стохастического разностного уравнения (7) называют [2] алгебраическое уравнение

$$x^p = \sum_{j=1}^{p} \beta_j x^{p-j}. \qquad (10)$$

Оно имеет p корней, которые мы обозначим $z_1,...,z_p$. Корни характеристического уравнения (10) определяют важные свойства процесса Y_t, удовлетворяющего уравнению (7). Эти свойства можно сформулировать следующими теоремами.

Теорема 1. Если все корни характеристического уравнения (10) по абсолютной величине меньше 1, то Y_t представим бесконечной линейной комбинацией с.в. $X_t, X_{t-1},...,$ и, следовательно, Y_t не коррелирует со с.в. $X_{t+1}, X_{t+2},...$. При этом последовательность Y_t будет стационарной (в широком смысле).

Замечание. Для стационарности с.п. Y_t, определяемой уравнением (7), от последовательности с.в. X_t в теореме 1 достаточно требовать лишь стационарности в широком смысле.

Теорема 2. Если стационарная с.п. Y_t удовлетворяет уравнению (7), характеристическое уравнение (10) которого имеет хотя бы один корень, равный единице, то с вероятностью 1 все значения этой последовательности совпадают.

Доказательства этих теорем можно найти в работах [2,3].

В дальнейшем мы будем ограничиваться рассмотрением случая, когда все корни характеристического уравнения (10) по абсолютной величине меньше единицы. Получим важное рекуррентное соотношение, связывающее параметры $\beta_j, j = \overline{1,p}$ авторегрессии со значениями корреляционной функции выходной с.п.АР(p). Для этого умножим (7) на Y_{i-k} и перейдем к математическому ожиданию. Учитывая при этом, что $MX_{t-k}Y_t = 0$ при k>0, получаем

$$R_Y(k) = \beta_1 R_Y(k-1) + \beta_2 R_Y(k-2) + \\ +...+\beta_p R_Y(k-p), k > 0 \qquad (11)$$

Уравнение (11) - разностное уравнение относительно R(k). Общее решение этого уравнения задается в виде

$$R_Y(k) = A_1 z_1^{|k|} + A_2 z_2^{|k|} +...+A_p z_p^{|k|}, \qquad (12)$$

если корни $z_1,...,z_p$ характеристического уравнения (10) различны. Граничными условиями будут p-1 соотношения

$$R_Y(-k) = R_Y(k), k = 1,2,...,p-1 \qquad (13)$$

и соотношение

$$\sigma_Y^2 = R_Y(0) = \beta_1 R_Y(-1) + \beta_2 R_Y(-2) + \\ +...+\beta_p R_Y(-p) + \sigma_X^2 \qquad (14)$$

которое получается умножением (7) на Y_t и переходом к математическому ожиданию в обеих частях полученного неравенства. Для определения коэффициентов $A_1,...,A_p$ сначала можно решить относительно $R_Y(0), R_Y(1),..., R_Y(p)$ систему из p уравнений вида (11) при k=1,2,...,p и уравнения (14), а затем относительно $A_1,...,A_p$ решить систему уравнений вида (12) при k=0,1,...,p-1, подставляя вместо $R_Y(r), r = \overline{0, p-1}$ предварительно найденные значения.

Вводя нормированную корреляционную функцию $r_Y(k) = R_Y(k) / \sigma_Y^2$ из (14), учитывая равенства (13), получаем выражение для дисперсии с.п.АР(p)

$$\sigma_Y^2 = \frac{\sigma_X^2}{1 - \beta_1 r_Y(1) - ...\beta_p r_Y(p)} \qquad (15)$$

Поделим все члены уравнения (11) на σ_Y^2 и подставив в него значения k=1,2,...,p, получим систему линейных уравнений для $\beta_1,...\beta_2$ со свободными членами $r_Y(1),...,r_Y(p)$

$$\left.\begin{aligned} r_Y(1) &= \beta_1 + \beta_2 r_Y(1) + \ldots + \beta_p r_Y(p-1) \\ r_Y(2) &= \beta_1 r_Y(1) + \beta_2 \ldots + \beta_p r_Y(p-2) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ r_Y(p) &= \beta_1 r_Y(p-1) + \beta_2 r_Y(p-2) + \ldots + \beta_p \end{aligned}\right\}$$

Эти уравнения обычно называются уравнениями Юла-Уокера и могут использоваться для получения оценок параметров $\beta_k, k = \overline{1,p}$.

Перейдем к рассмотрению спектра с.п.АР(p). Заметим, что входная с.п. $X_t, t = 0, \pm 1, \pm 2, \ldots$ имеет корреляционную функцию $R_X(k) = 0$ при $k \neq 0$ и $R_X(0) = \sigma_X^2$. Этой корреляционной функции соответствует согласно (5) постоянная на $[-\pi, \pi]$ спектральная плотность

$$f(\omega) = \frac{\sigma_X^2}{2\pi}. \tag{16}$$

В силу теоремы Хинчина для последовательности X_t будет справедливо представление (6) с некоторым процессом с ортогональными приращениями $W_\omega, \omega \in [-\pi, \pi]$. Выходная с.п.АР(p) $Y_t, t = 0, \pm 1, \pm 2, \ldots$ будет иметь корреляционную функцию $R_Y(k), k = 0, \pm 1, \pm 2, \ldots$ удовлетворяющая рекуррентному соотношению (11), и сам процесс будет иметь также в силу теоремы Хинчина интегральное представление

$$Y_t = \int_{-\pi}^{\pi} e^{it\omega} dZ_\omega \tag{17}$$

с процессом, с ортогональными приращениями $Z_\omega, \omega \in [-\pi, \pi]$, со структурной функцией $F_Y(\omega)$, являющейся одновременно спектральной функцией последовательности Y_t. Подставляя в (7) вместо X_t его выражение из (6), а вместо Y_t - выражение из (17), получаем

$$\int_{-\pi}^{\pi} e^{it\omega} \left[1 - \sum_{j=1}^{p} \beta_j e^{-ij\omega}\right] dZ_\omega = \int_{-\pi}^{\pi} e^{ij\omega} dV_\omega \tag{18}$$

Обе части этого неравенства - стохастические интегралы при любом целом t. Умножая (18) на такое же равенство при t-k(k-целое) и беря затем математическое ожидание от обеих частей полученного равенства, находим (с учетом свойств стохастических интегралов)

$$\int_{-\pi}^{\pi} e^{ik\omega}\left|1-\sum_{j=1}^{p}\beta_j e^{-ij\omega}\right|^2 dF_Y(\omega)=\int_{-\pi}^{\pi} e^{ik\omega} f_X(\omega)d\omega. \qquad (19)$$

Из (12) видно, что для стационарной с.п.АР(р) Y_t условие (2) выполняется и, следовательно, существует спектральная плотность $f_Y(\omega)=F^{'}(\omega)$. Тогда из равенства (19) получаем

$$f_Y(\omega)=\frac{f_X(\omega)}{\left|1-\sum_{j=1}^{p}\beta_j e^{-ij\omega}\right|^2}=\frac{\sigma_X^2}{2\pi\left|1-\sum_{j=1}^{p}\beta_j e^{-ij\omega}\right|^2}. \qquad (20)$$

Приведем основные характеристики с.п.АР(р) первого и второго порядков. При р=1 мы имеем дело со с.п.АР(1), определяемой уравнением (8). Ему соответствует характеристическое уравнение $x=\beta_1$ с очевидным корнем $z_1=\beta_1$. Это означает, что для стационарной с.п.АР(1) должно выполнятся условие $|\beta_1|<1$. При этом из формулы (12) следует, что $R(k)=A_1\beta_1^k$ для $k\geq 0$, где $A_1=R_Y(0)=\sigma_Y^2$. Согласно (15) для дисперсии Y_t получаем формулу

$$\sigma_Y^2=\frac{\sigma_X^2}{1-\beta_1^2}, \qquad (21)$$

а для корреляционной функции

$$R_Y(k)=\frac{\sigma_X^2}{1-\beta_1^2}\beta_1^k, k\geq 0. \qquad (22)$$

Спектральная плотность с.п.АР(1) согласно (20) имеет вид

$$f(\omega)=\frac{\sigma_X^2}{2\pi\left|1-\beta_1 e^{-i\omega}\right|^2}=\frac{\sigma_X^2}{2\pi\left(1-2\beta_1\cos(\omega)+\beta_1^2\right)}. \qquad (23)$$

Теперь рассмотрим с.п.АР(2), определяемую уравнением (3). Характеристическое уравнение будет иметь вид

$$z^2-\beta_1 z-\beta_2=0 \qquad (24)$$

и с.п.АР(2) будет стационарной, если корни уравнения (24) лежат внутри единичного круга $|z|<1$, т.е.

$$\left|z_{1,2}\right|=\left|\frac{\beta_1\pm\sqrt{\beta_1^2+4\beta_2}}{2}\right|<1. \qquad (25)$$

В случае действительных различных корней $\left(\beta_1^2+\beta_2>0\right)$ имеем неравенства $-2<\beta_1+\sqrt{\beta_1^2+4\beta_2}<2$ и $-2<\beta_1-\sqrt{\beta_1^2+4\beta_2}<2$. Это означает, что

$$2+\beta_1>\sqrt{\beta_1^2+4\beta_2} \text{ и } \sqrt{\beta_1^2+4\beta_2}<2-\beta_1, \qquad (26)$$

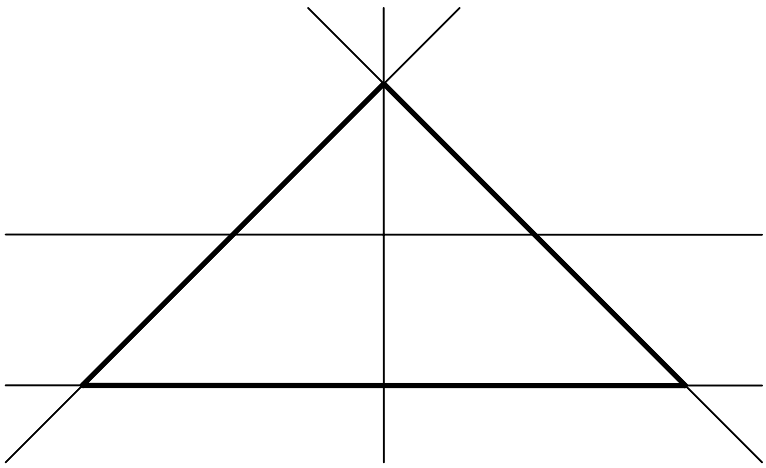
откуда получаем

$$\beta_1+\beta_2<1, \beta_2-\beta_1<1. \qquad (27)$$

Итак, область параметров β_1,β_2, обеспечивающих различные действительные корни $\left|z_{1,2}\right|<2$ уравнения (24), задается неравенством (27) и неравенством $\beta_1^2+4\beta_2>0$. Найдем область параметров β_1,β_2, дающих комплексные (сопряженные) корни $\left|z_{1,2}\right|<1$. При этом дискриминант $\beta_1^2+4\beta_2<0$ и, следовательно, $\beta_2<0$. Условие (25) тогда можно представить в виде

$$\left|z_{1,2}\right|^2=-\beta_2<1.$$

Следовательно, интересующая нас область параметров β_1,β_2, определяется неравенствами $\beta_1^2+4\beta_2<0$ и $-1<\beta_2<0$. Результатам проведенного анализа дадим геометрическую иллюстрацию (рис.2) на плоскости (β_1,β_2).



Корреляционная функция с.п.АР(2) в соответствии с формулой (12) будет иметь вид

$$R_Y(k)=A_1z_1^{|k|}+A_2z_2^{|k|}$$

Записывая уравнения (14) и (11) для p=2 при k=1,2, получаем систему относительно

$$R_Y(0), R_Y(1), R_Y(2)$$

$$\left.\begin{aligned}R_Y(0)&=\beta_1R_Y(1)+\beta_2R_Y(2)+\sigma_X^2\\R_Y(1)&=\beta_1R_Y(0)+\beta_2R_Y(1)\\R_Y(2)&=\beta_1R_Y(1)+\beta_2R_Y(0)\end{aligned}\right\}$$

Решая ее, находим

$$\sigma_Y^2=R_Y(0)=\frac{(1-\beta_2)\sigma_X^2}{1-2\beta_1-\beta_1^2(1+\beta_2)+\beta_2^2} \qquad (28)$$

$$R_Y(1)=\frac{\beta_1}{1-\beta_2}R_Y(0),R_Y(2)=\frac{1+\beta_1^2-\beta_2^2}{1-\beta_2}R_Y(0). \quad (29)$$

Полагая, k=0 и k=1, получим систему

$$\left.\begin{aligned}R_Y(0)&=A_1+A_2,\\R_Y(1)&=A_1z_1+A_2z_2.\end{aligned}\right\}$$

из которой находим

$$A_1=\frac{z_1(1-z_2^2)}{(1+z_1z_2)(z_1-z_2)}R_Y(0),$$

$$A_1=-\frac{z_2(1-z_1^2)}{(1+z_1z_2)(z_1-z_2)}R_Y(0). \qquad (30)$$

Подставляем в эти формулы значения корней z_1 и z_2 квадратного уравнения (24) и $R_Y(0)$ из (28), получим по формуле (12) значения корреляционной функции $R_Y(k),k>2$, выраженные через параметры β_1 и β_2. Для $|k|\le 2$ значения корреляционной функции определяются формулами (28) и (29).

В соответствии с формулой (20) получаем спектральную плотность с.п.АР(2)

$$f(\omega)=\frac{\sigma_X^2}{2\pi\left[1+\beta_1^2+\beta_2^2-2\beta_1(1-\beta_2)\cos\omega-2\beta_2\cos2\omega\right]},$$

$$\omega\in[-\pi,\pi]. \qquad (31)$$

2.3. Случайные последовательности скользящего среднего

Последовательность с.в. $Y_t,t=0,\pm1,\pm2,...$ называют случайной последовательностью скользящего среднего порядка $q\ge 1$ (с.п.СС(q)), если она задается равенством

$$Y_t = \sum_{j=0}^{q} \alpha_j X_{i-j} \qquad (32)$$

в котором X_t - последовательность вещественных некоррелированных и одинаково распределенных с $MX_t = 0$ и $DX_t = \sigma_X^2$ с.в., а $\alpha_j, j = \overline{0,q}$ $(\alpha_0 = 1, \alpha_q \neq 0)$- вещественные параметры. Из определения с.п.СС(q) следует, что $MY_t = 0$. Вычислим ее корреляционную функцию

$$MY_t Y_s = M\left(\sum_{j=0}^{q}\sum_{r=0}^{q} \alpha_j \alpha_r X_{t-j} X_{s-r}\right) = \sum_{j,r=0}^{q} \alpha_j \alpha_r MX_{t-j} X_{s-r} .$$

Так как $MX_t X_s = 0$ при $s \neq t$ и $MX_t^2 = \sigma_X^2$, то в сумме в правой части последнего неравенства ненулевыми будут только слагаемые, индексы которых удовлетворяют равенству $t - j = s - r$ или $t - s = j - r$. Если $|t - s| > q$, то все слагаемые в сумме нулевые. При $|t - s| \le q$

$$MY_t Y_s = \sigma_X^2 \sum_{r=0}^{q-|t-s|} \alpha_r \alpha_{r+|t-s|} .$$

Итак, с.п.СС(q) оказывается стационарной в широком смысле без всяких ограничений на параметры $\alpha_r, r = \overline{1,q}$. Обозначая ее корреляционную функцию через $R_Y(k)$, получаем формулу

$$R_Y(k) = \begin{cases} \sigma_X^2 \sum\limits_{j=0}^{q-|k|} \alpha_j \alpha_{j+|k|} & ,\text{если} |k| \le q, \\ 0 & ,\text{если} |k| > q. \end{cases} \qquad (33)$$

Для с.п.СС(q) отсюда имеем формулу

$$\sigma_Y^2 = R_Y(0) = \sigma_X^2 \sum_{j=0}^{q} \alpha_j^2 . \qquad (34)$$

Условие (2) для корреляционной функции (33), очевидно, выполняется и, следовательно, существует спектральная плотность с.п.СС(q), которую обозначим $f_Y(\omega)$. Спектральная плотность «входной» последовательности X_t согласно (16) постоянна на $[-\pi, \pi]$ и $f_X(\omega) = \sigma_X^2 / 2\pi$. С.в. X_t имеет интегральное представление (6). В соответствии с теоремой Хинчина «выходная» последовательность Y_t будет иметь представление вида (17). Учитывая эти представления, равенство (32) можно записать в виде

$$\int_{-\pi}^{\pi} e^{it\omega} dZ_{\omega} = \int_{-\pi}^{\pi} e^{it\omega} \left[\sum_{j=0}^{q} \alpha_j e^{-ij\omega} \right] dV_{\omega} .$$

Умножим это равенство на такое же равенство, в котором вместо t положено $t-k$, и в обеих частях произведен переход к комплексно-сопряженным величинам. Беря затем математические ожидания от обеих частей полученного равенства, находим

$$\int_{-\pi}^{\pi} e^{ik\omega} f_Y(\omega) d\omega = \int_{-\pi}^{\pi} e^{ik\omega} \left| \sum_{j=0}^{q} \alpha_j e^{-ij\omega} \right|^2 \frac{\sigma_X^2}{2\pi} d\omega .$$

Из единственности интегрального представления корреляционной функции отсюда имеем

$$f_Y(\omega) d\omega = \frac{\sigma_X^2}{2\pi} \left| \sum_{j=0}^{q} \alpha_j e^{-ij\omega} \right|^2, \omega \in [-\pi, \pi]. \qquad (35)$$

Оказывается, что спектральная плотность с.п.СС(q), вообще говоря, неоднозначно определятся параметрами $\sigma_X^2, \alpha_1, ..., \alpha_k$. Во избежание этого достаточно условится, чтобы корни характеристического уравнения

$$z^q + \alpha_1 z^{q-1} + ... + \alpha_q = 0 \qquad (36)$$

лежали, например, внутри единичного круга. Это условие однозначности иногда в литературе называется *условием обратимости* с.п.СС(q).

Большое практическое применение имеют последовательности скользящего среднего первого и второго порядков. с.п.СС(1) определяются уравнением

$$Y_t = X_1 + \alpha_1 X_{t-1}, \qquad (37)$$

и ее корреляционная функция имеет вид

$$R_Y(k) = \begin{cases} \sigma_X^2 \left(1 + \alpha_1^2\right) = \sigma_Y^2 & , k = 0, \\ \sigma_X^2 \alpha_1 & , k = 1, \\ 0 & , k \geq 2. \end{cases} \qquad (38)$$

Спектральная плотность с.п.СС(1) имеет вид

$$f_Y(\omega) = \frac{\sigma_X^2}{2\pi} \left| 1 + \alpha_1 e^{-i\omega} \right|^2 = \left(1 + 2\alpha_1 \cos\omega + \alpha_1^2\right) \frac{\sigma_X^2}{2\pi} . \qquad (39)$$

С.п.СС(2) определяется уравнением

$$Y_t = X_1 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2}, \qquad (40)$$

а ее корреляционная имеет вид

$$R_Y(k)=\begin{cases}\sigma_X^2\left(1+\alpha_1^2+\alpha_2^2\right)=\sigma_Y^2 & ,k=0,\\ \sigma_X^2\alpha_1\left(1+\alpha_2\right) & ,k=1,\\ \sigma_X^2\alpha_2 & ,k=2,\\ 0 & ,k\ge 3.\end{cases} \qquad (41)$$

Спектральная плотность с.п.СС(2) имеет вид

$$f_Y(\omega)=\frac{\sigma_X^2}{2\pi}\left|1+\alpha_1 e^{-i\omega}+\alpha_2 e^{-2i\omega}\right|^2=$$

$$=\frac{\sigma_X^2}{2\pi}\left[1+\alpha_1^2+\alpha_2^2+2\alpha_1\left(1+\alpha_2\right)\cos\omega+2\alpha_2\cos 2\omega\right].(42)$$

2.4. Смешанная модель авторегрессии - скользящего среднего

В теореме 1 утверждалось, что в условиях стационарности с.п.АР(p) Y_t может быть представлена бесконечной линейной комбинацией с.в. $X_t, X_{t-1}, \ldots$, т.е. может рассматриваться как с.п.СС(∞) с последовательностью параметров $\alpha_1, \alpha_2, \ldots$. Известно также, что и с.п.СС(q) может быть (при условии обратимости) представлена в виде с.п.АР(∞) с последовательностью параметров $\beta_1, \beta_2, \ldots$. Это ставит вопрос об экономичности (в смысле числа используемых параметров) представления данной с.п. На практике для получения экономичной параметризации иногда бывает необходимо включать в модель как члены, описывающие авторегрессию, так и члены, моделирующие скользящее среднее. Такая с.п. может определена уравнением

$$Y_t=\sum_{j=0}^{p}\beta_j Y_{t-j}+\sum_{r=0}^{q}\alpha_r X_{t-r}\,, \qquad (43)$$

где $\beta_j, j=\overline{1,p}, \alpha_0=1, \alpha_r, r=\overline{1,q}$ - вещественные параметры, а $X_t, t=0,\pm1,\pm2,\ldots$ - последовательность некоррелированных одинаково распределенных с.в. с $MX_t=0, DX_t=\sigma_X^2$ и называется смешанной с.п. авторегрессии-скользящего среднего порядка (p,q). В дальнейшем такую последовательность сокращенно будем обозначать АРСС(p,q).

В соответствии с замечанием к теореме 1 члены со скользящим средним в правой части (43) не повлияют на условия стационарности последовательности Y_t. Поэтому с.п. АРСС(p,q) будет стационарной в широком смысле при условии, что все корни характеристического уравнения

$$z^p = \sum_{j=1}^{p} \beta_j z^{p-j} \qquad (44)$$

лежат внутри единичного круга $|z| < 1$. Аналогично для обратимости АРСС(p,q) корни характеристического уравнения

$$z^q + \alpha_1 z^{q-1} + \ldots + \alpha_q = 0 \qquad (45)$$

должны лежать внутри единичного круга $|z| < 1$.

Предполагая с.п. АРСС(p,q) стационарной, найдем, как и для с.п. АР(p), рекуррентные соотношения, связывающие параметры $\beta_j, j = \overline{1,p}$ и $\alpha_r, r = \overline{1,q}$ со значениями корреляционной функции. Для этого все члены в (43) умножим на Y_{t-k} и, перейдя к математическим ожиданиям получаем

$$R_Y(k) = \sum_{j=1}^{p} \beta_j R_Y(k-j) + \sum_{r=1}^{q} \alpha_j R_{YX}(k-r), \qquad (46)$$

где $R_{YX}(k) = MX_t Y_{t-k}$ - взаимная корреляционная функция последовательностей X и Y. Так как Y_{t-k} зависит только от членов входной последовательности X до момента $t-k$, то, очевидно, что $R_{YX}(k) = 0$ при $k > 0$ и $R_{YX}(k) \neq 0$ для $k \leq 0$. Из (46) следует, что

$$R_Y(k) = \sum_{j=1}^{p} \beta_j R_Y(k-j), k \geq q+1 \qquad (47)$$

и для нормированной корреляционной функции

$$r_Y(k) = \sum_{j=1}^{p} \beta_j r_Y(k-j), k \geq q+1 \qquad (47')$$

Это означает, что для с.п. АРСС(p,q) существует q значений корреляционной функции $R_Y(q), R_Y(q-1), \ldots, R_Y(1)$, которые связаны зависимостью (46) с q параметрами скользящего среднего α и p параметрами авторегрессии β. Для решения разностных уравнений (47) и (47’) (для больших k) в качестве начальных необходимы p значений, например, $R_Y(q), R_Y(q-1), \ldots, R_Y(q-p+1)$.

Дисперсию с.п. АРСС(p,q) $\sigma_X^2 = R_Y(0)$ вместе с $R_Y(1), \ldots, R_Y(p)$ получим, решая систему уравнений, получающаяся из (46) при k=0,1,2,...,p.

Спектральную плотность можно получить аналогично случаям «чистых» последовательностей АР(p) и СС(q).

$$f_Y(\omega)=\frac{\sigma_X^2}{2\pi}\frac{\left|\sum_{r=0}^{q}\alpha_r e^{-ir\omega}\right|^2}{\left|1-\sum_{j=1}^{p}\beta_j e^{-ij\omega}\right|^2},\omega\in[-\pi,\pi]. \qquad (48)$$

Рассмотрим подробнее случай АРСС(2,1):

$$Y_t=\beta_1 Y_{t-1}+\beta_2 Y_{t-2}+X_t+\alpha_1 X_{t-1}. \qquad (49)$$

Из (46) имеем

$$R_Y(0)=\beta_1 R_Y(1)+\beta_2 R_Y(2)+R_{YX}(0)+\alpha_1 R_{YX}(-1) \qquad (50)$$

$$R_Y(1)=\beta_1 R_Y(0)+\beta_2 R_Y(1)+\alpha_1 R_{YX}(0). \qquad (51)$$

Чтобы найти $R_{YX}(0)$ и $R_{YX}(-1)$, умножим поочередно (49) на X_t и X_{t-1} и перейдем к математическим ожиданиям. В результате получим

$$R_{YX}(0)=\sigma_X^2 \text{ и } R_{YX}(-1)=(\alpha_1+\beta_1)\sigma_X^2.$$

Тогда уравнениям (50) и (51) приобретают вид

$$\left.\begin{array}{l}R_Y(0)=\beta_1 R_Y(1)+\beta_2 R_Y(2)+\sigma_X^2\left(1+\alpha_1^2+\alpha_1\beta_2\right),\\(1-\beta_2)R_Y(1)=\beta_1 R_Y(0)+\alpha_1\sigma_X^2.\end{array}\right\} \qquad (52)$$

При $k\ge 2$ $(k\ge q+1)$ уравнения (47) в рассматриваемом случае имеют вид

$$R_Y(k)=\beta_1 R_Y(k-1)+\beta_2 R_Y(k-2) \qquad (53)$$

и вместе с уравнениями (52) позволяют определить последовательность $R_Y(k), k=0,1,\ldots$. В частности, из системы уравнений (52) и уравнения (53) при $k=2$ получаем формулу для дисперсии

$$\sigma_Y^2=R_Y(0)=\sigma_X^2\frac{2\alpha_1\beta_1+\left(1+\alpha_1^2\right)(1-\beta_2)}{(1-\beta_2)\left[(1-\beta_2)^2-\beta_1^2\right]}. \qquad (54)$$

Спектральная плотность с.п. АРСС(2,1) согласно (48) имеет вид

$$f_Y(\omega)=\frac{\sigma_X^2}{2\pi}\frac{\left|1+\alpha_1 e^{-i\omega}\right|^2}{\left|1-\beta_j e^{-i\omega}-\beta_j e^{-2i\omega}\right|^2}=$$

$$=\frac{\sigma_X^2}{2\pi}\frac{1+2\alpha_1\cos\omega+\alpha_1^2}{1+\beta_1^2+\beta_2^2-2\beta_1(1-\beta_2)\cos\omega-2\beta_2\cos 2\omega}. \qquad (55)$$

Наконец, рассмотрим часто употребляемую в различных прикладных науках с.п. АРСС(1,1). Она определяются разностным уравнением

$$Y_t = \beta_1 Y_{t-1} + X_t + \alpha_1 X_{t-1}. \qquad (56)$$

В этом случае входящая в с.п. авторегрессия имеет порядок p=1, и корень ее характеристического уравнения равен β_1. Последовательность будет стационарной, если $-1 < \beta_1 < 1$. Уравнения для корреляционной функции получаются из формул (46) и (47)

$$\left.\begin{aligned} &R_Y(0) = \beta_1 R_Y(1) + R_{YX}(0) + \alpha_1 R_{YX}(-1) \\ &R_Y(1) = \beta_1 R_Y(0) + \alpha_1 R_{YX}(0) \\ &R_Y(k) = \beta_1 R_Y(k-1)\_\text{при}\_k \geq 2 \end{aligned}\right\} \qquad (57)$$

Выражения для $R_{YX}(0)$ и $R_{YX}(-1)$ получаются аналогично предыдущему умножением (56) на X_t и X_{t-1} и переходом к математическим ожиданиям.

$$R_{YX}(0) = \sigma_X^2,$$

$$R_{YX}(-1) = (\alpha_1 + \beta_1)\sigma_X^2. \qquad (58)$$

Из (57) и (58) получаем выражения корреляционной функции с.п. АРСС(1,1)

$$R_Y(0) = \sigma_X^2 \frac{1 + \alpha_1^2 + 2\alpha_1\beta_1}{1 - \beta_1^2},$$

$$R_Y(1) = \sigma_X^2 \frac{(1 + \alpha_1\beta_1)(\alpha_1 + \beta_1)}{1 - \beta_1^2}, \qquad (59)$$

$$R_Y(k) = \beta_1 R_Y(k-1), \quad k \geq 2.$$

Из (59) следует, что при $\alpha_1 = -\beta_1$ имеем $R_Y(1) = 0$. Поэтому $R_Y(k) \equiv 0$ при всех значениях $k \neq 0$, т.е. последовательность $Y_t, t = 0, \pm 1, \pm 2, \ldots$ является некоррелированной.

Спектральная плотность с.п. АРСС(1,1) будет иметь вид

$$f_Y(\omega) = \frac{\sigma_X^2}{2\pi} \frac{1 + 2\alpha_1 \cos\omega + \alpha_1^2}{1 - 2\beta_1 \cos\omega + \beta_1^2}. \qquad (60)$$

3.СТАТИСТИЧЕСКИЕ ВЫВОДЫ ПО НАБЛЮДЕНИЯМ СЛУЧАЙНЫХ ВЕРОЯТНОСТЕЙ

3.1 Общая характеристика задач статистических случайных процессов

При изучении систем, осуществляющих преобразования входных случайных сигналов, приходится решать ряд статистических задач. Прежде всего это задачи определения вероятностных характеристик входных и выходных сигналов, а также задачи проверки гипотез об этих сигналах. Кроме того, часто неизвестной или частично известной является сама преобразующая система. В простых случаях она бывает известной с точностью до конечной совокупности определяющих ее параметров. Задачи статистических выводов в этом случае связаны с этими параметрами и включаются в так называемые методы идентификации, призванные по экспериментальным данным определить тип преобразующей системы. Все упомянутые задачи решаются на основе обработки экспериментального материала и относятся к сравнительно новой ветви математической статистики, называемой *статистикой случайных процессов*. Применяемые в статистике случайных процессов (и, в частности, случайных последовательностей) методы принципиально не отличаются от обычных методов обработки, используемых в математической статистике (МС), однако имеются некоторые особенности, осложняющие решение задач. Как известно, все оценки и критерии в МС строятся на основе выборки

$$x_1,...,x_n \qquad (61)$$

где n - объем выборки, а ее элементы x_k являются значениями исследуемой с.в. X при независимых измерениях. Поэтому выборку (16) в МС рассматривают как реализацию случайной последовательности «длины» n независимых с.в. $X_1, X_2,...$ с общим распределением, совпадающим с распределением исследуемой с.в.X. В статистике случайных процессов по ряду причин основой не могут служить независимые реализации процесса (или последовательности) Y, и поэтому теряет смысл понятие объема выборки n, входящего явно в известные в МС формулы для критериев и оценок. Аналогичную объему выборки роль в статистике случайных процессов играет «длина» времени, в течении которого наблюдается реализация процесса. Если Y- случайная последовательность: $Y=\{Y_t, t=0,\pm1,\pm2,...\}$, то реализация, наблюдаемая в моменты t=1,...T, будет представлять собой совокупность значений

$$y_1, y_2, ..., y_T, \qquad (62)$$

которая и служит основой для статистических выводов.

Мы здесь имеем возможность рассмотреть только некоторые задачи статистической обработки реализаций с.п. в предположении их стационарности. Предположим известными основные факты МС независимых наблюдений.

Некоторые дополнительные сведения содержаться в прил.1 и 2 настоящего пособия. При исследовании входной случайной последовательности $X=(X_1, X_2, ...)$ одной из важнейших является задача проверки гипотезы о том, что случайные величины $X_k, k=1,2,...$ независимы и одинаково распределены. Один из критериев проверки такой гипотезы изложен в прил.1. Прил.2. содержит описание одного из возможных критериев проверки гипотезы стационарности.

Следующий раздел данного учебного пособия содержит традиционный материал по оцениванию среднего значения и корреляционной функции стационарных последовательностей и не связан прямо с последовательностями авторегрессии и скользящего среднего. Остальной материал раздела посвящен решению статистических задач АР и СС. Следует отметить, что основная статистическая задача оценивания параметров АР и СС может иметь различные интерпретации. Ее можно, в частности, трактовать как оценивание спектра, а также и как задачу идентификации системы, структура которой определена с точностью до неизвестных параметров.

3.2. Оценка среднего значения и корреляционной функции

Пусть $y_1,...y_T$- T последовательных наблюдений с.п. $Y=\{Y_t, t=0,\pm1,\pm2,...\}$, стационарной в широком смысле со средним значением $m^Y = MY_t, t = 0,\pm1,...,$ и корреляционной функцией $R_Y(k) = M(Y_t - m^Y)(Y_{t+k} - m^Y)$, $t = 0,\pm1,\pm2,...,$ $k = 0,\pm1,\pm2,...,$ $R_Y(k) = R_Y(-k)$.

Рассмотрим сначала оценивание среднего значения m^Y. Будем искать оценку величины m^Y в классе линейных оценок, т.е. среди линейных комбинаций наблюдений

$$l_T = \sum_{t=1}^{T} a_t y_t \,. \tag{63}$$

Среднее значение и дисперсия произвольной линейной комбинации (63) соответственно равны

$$ML_T = \sum_{t=1}^{T} a_t MV_t = m^Y \sum_{t=1}^{T} a_t \tag{64}$$

и

$$DL_T = M\left[\sum_{t=1}^{T} a_t\left(Y_t - m^Y\right)\right]^2 = \sum_{t,s=1}^{T} a_t a_s R_t(t-s) =$$

$$= \sum_{r=-(T-1)}^{T-1} R_Y(r) \sum_{s\in S_t} a_{s+r} a_s = \int_{-\pi}^{\pi} \left|\sum_{t=1}^{T} a_t e^{it\omega}\right| f_Y(\omega) d\omega \,, \tag{65}$$

где $S_r = \{1,2,..,T-r\}$ при $r \ge 0$ и $S_r = \{1-r, 2-r,..,T\}$ при $r \le 0$. Для того, чтобы с.в. L_T была несмещенной оценкой среднего значения m^Y, должно выполнятся равенство

$$\sum_{t=1}^{T} a_t = 1 \tag{66}$$

По обычным правилам отыскания условного экстремума находим вектор $a = (a_1, a_2,.., a_T)$, обеспечивающий минимальную дисперсию несмещенной оценке

$$\widetilde{m}^Y = \sum_{t=1}^{T} a_t y_t \,,\ \sum_{t=1}^{T} a_t = 1. \tag{67}$$

Несложные выкладки показывают, что несмещенная линейная оценка с минимальной дисперсией задается формулой (67) при векторе коэффициентов

$$a^0 = \left[\left(R^{-1}e, e\right)\right]^{-1} R^{-1} e, \tag{68}$$

где R^{-1} - матрица, обратная к корреляционной матрице наблюдений $R = \left\|R_Y(t-s)\right\|_{t,s=1}^{T}; e = (1,1,....,1)$ - единичный вектор, а (.,.) - скалярное произведение n-мерном евклидовом пространстве. Таким образом, оптимальная линейная несмещенная оценка будет иметь вид

$$\hat{m} = (a, y) = \left(R^{-1}e, y\right) / \left(R^{-1}e, e\right). \qquad (69)$$

где $y = (y_1, .., y_T)$- вектор наблюдений, а дисперсия этой оценки определяется равенством

$$D\hat{m}^Y = \left(Ra^0, a^0\right) = \left(R^{-1}e, e\right)^{-1} = \left[\sum_{i,j=1}^{T} R_{ij}^{-1}\right]^{-1} \qquad (70)$$

где R_{ij}^{-1} - элементы матрицы R^{-1}.

На практике часто в качестве несмещенной оценки среднего значения стационарной с.п. используется среднее арифметическое наблюдений

$$\overline{y} = \frac{1}{T}\sum_{t=1}^{T} y_t \qquad (71)$$

Используя формулу (65) получим дисперсию этой оценки

$$D\overline{Y} = \frac{1}{T_t^2}\sum_{t,s=1}^{T} R_t(t-s) = \frac{1}{T}\sum_{r=-(T-1)}^{T-1}\left(1 - \frac{|r|}{T}\right) R_Y(r) =$$

$$= \int_{-\pi}^{\pi} \left(\frac{\sin\frac{\omega T}{2}}{T\sin\frac{\omega}{2}}\right)^2 f_Y(\omega) d\omega. \qquad (72)$$

Асимптотическое (при $T \to \infty$) поведение дисперсии среднеарифметической оценки величины m^Y дается следующей теоремой.

Т е о р е м а 3. Если выполняется условие

$$\sum_{r=-\infty}^{\infty} R_Y(r) < \infty, \qquad (73)$$

то оценка (71) состоятельна и

$$\lim_{T\to\infty} T D\overline{Y} = 2\pi f_Y(0). \qquad (74)$$

Доказательство состоятельности оценки $\overline{Y}$ следует из того, что при условии (73) $D\overline{Y} \to 0$ при $T \to \infty$. Основанием соотношения (74) является известный из анализа факт.

Л е м м а. Если ряд $\sum_{r=1}^{\infty} a_r$ сходится, то

$$\lim_{T\to\infty}\sum_{r=1}^{T-1}\left(1 - \frac{r}{T}\right) a_r = \sum_{r=1}^{\infty} a_r. \qquad (75)$$

Соотношение (75) следует из (74), если учесть разложение вида (5) спектральной плотности в ряд Фурье, полагая в нем $\omega = 0$.

Теперь рассмотрим *оценивание корреляционной функции*. Если m^Y известно, то обычно используется оценка

$$\widetilde{R}_Y(k) = \widetilde{R}_Y(-k) = \frac{1}{T-k}\sum_{t=1}^{T-k}\left(y_t - m^Y\right)\left(y_{t+k} - m^Y\right), \qquad (76)$$

где $k = 0, \pm1, \pm2, ..., T-1$.

Если m^Y неизвестно, то по аналогии можно построить следующую оценку:

$$\widetilde{R}_Y(k) = \widetilde{R}_Y(-k) = \frac{1}{T-k}\sum_{t=1}^{T-k}\left(y_t - \overline{y}\right)\left(y_{t+k} - \overline{y}\right), \qquad (77)$$

где $k = 0, \pm1, \pm2, ..., T-1$.

Возможны еще и другие оценки, в частности

$$R_Y^*(k) = R_Y^*(-k) = \frac{1}{T-k}\sum_{t=1}^{T-k}\left(y_t - y_k\right)\left(y_{t+k} - y_{k+}\right) =$$

$$= \frac{1}{T-k}\left[\sum_{t=1}^{T-k} y_t y_{t+k} - (T-k) y_k y_{k+}\right], k = 0, 1, ..., T-2,$$

где

$$\overline{y}_k = \frac{1}{T-k}\sum_{t=1}^{T-k} y_t, k = 0, \pm1, \pm2, ..., T-2,$$

$$\overline{y}_{k+} = \frac{1}{T-k}\sum_{t=1}^{T-k} y_{t+k}, k = 0, \pm1, \pm2, ..., T-2.$$

Рассмотрим моменты первого и второго порядков оценок величины $R_Y(k)$. Непосредственные вычисления дают следующие результаты. В случае известного среднего m^Y

$$M\widetilde{R}_Y(k) = \frac{1}{T-k}\sum_{t=1}^{T-k} M\left(Y_t - m^y\right)\left(Y_{t+k} - m^y\right) = MR_Y(k), \qquad (78)$$

т.е. оценка $\widetilde{R}_Y(k)$ является несмещенной. В случае неизвестного m^Y для математического ожидания оценки $\widetilde{R}_Y(k)$ после несложных, но утомительных вычислений получаем соотношения:

$$M\hat{R}_Y(0) = R_Y(0) - \frac{1}{T}\left\{R_Y(0) + 2\sum_{r=1}^{T-1}\left(1 - \frac{r}{T}\right)R_Y(r)\right\}; \qquad (79)$$

$$M\hat{R}_Y(k) = R_Y(k) - \frac{1}{T}\left\{R_Y(0) + 2\sum_{r=1}^{k}\left[1 - \frac{rk}{T(T-k)}\right]R_Y(r) + \right.$$

$$\left. + 2\sum_{r=k+1}^{T-k-1}\left[1 - \frac{rk}{T(T-k)} - \frac{r-k}{T-k}\right]R_Y(r) + 2\sum_{r=T-k}^{T-1}\frac{(T-r)k}{T(T-k)}R_Y(r)\right\} \qquad (80)$$

если $1 \le k < T-k-1$;

$$M\hat{R}_Y(k) = R_Y(k) - \frac{1}{T}\left\{R_Y(0) + 2\sum_{r=1}^{k}\left[1 - \frac{rk}{T(T-k)}\right]R_Y(r) + \right.$$

$$\left. + 2\sum_{r=k+1}^{T-1}\frac{(T-r)k}{T(T-k)}R_Y(r)\right\}, \qquad (81)$$

если $1 \le k = T-k-1$;

$$M\hat{R}_Y(k) = R_Y(k) - \frac{1}{T}\left\{R_Y(0) + 2\sum_{r=1}^{T-k-1}\left[1 - \frac{rk}{T(T-k)}\right]R_Y(r) + \right.$$

$$\left. + 2\sum_{r=T-k}^{k}\frac{r}{T}R_Y(r) + 2\sum_{r=k+1}^{T-1}\frac{(T-r)k}{T(T-k)}R_Y(r)\right\}, \qquad (82)$$

если $T-k-1 < k < T-1$;

$$M\hat{R}_Y(T-1) = R_Y(T-1) - \frac{1}{T}\left\{R_Y(0) + 2\sum_{r=1}^{T-1}\frac{r}{T}R_Y(r)\right\}. \qquad (83)$$

Формулы (79)-(83) показывают, что оценка $\hat{R}_Y(k)$ является смещенной, порядок смещения равен $1/T$.

Математическое ожидание оценки $\hat{R}_Y(k)$ можно также выразить с помощью спектральной плотности

$$M\hat{R}_Y(k) = \int_{-\pi}^{\pi}\left\{\cos\omega k - \frac{1}{T(T-k)}\sum_{t=1}^{T-k}\sum_{s=1}^{T}\left[e^{i\omega(t-s)} + e^{i\omega(t-s+k)}\right] + \right.$$

$$\left. + \frac{1}{T^2}\sum_{s,t=1}^{T}e^{i\omega(t-s)}\right\}f_Y(\omega)d\omega =$$

$$= \int_{-\pi}^{\pi}\left\{\cos\omega k - 2\frac{\sin\frac{1}{2}\omega T\sin\frac{1}{2}\omega(T-k)}{T\sin\frac{1}{2}\omega(T-k)\sin\frac{1}{2}\omega}\cos\frac{1}{2}\omega k + \right.$$

$$\left. + \left(\frac{\sin\frac{1}{2}\omega T}{T\sin\frac{1}{2}\omega}\right)^2\right\}f_Y(\omega)d\omega. \qquad (84)$$

Аналогичные выражения можно получить и для математического ожидания оценки $R_Y^*(k)$.

Ждя дисперсии несмещенной оценки $\hat{R}_Y(k)$ можно получить следующее выражение:

$$(T-k)D\widetilde{R}_Y(k)=\sum_{r=-(T-k-1)}^{T-k-1}\left(1-\frac{|r|}{T-k}\right)\left[R_Y^2(r)+\right.$$

$$\left.+R_Y(r+k)R_Y(r-k)+\chi(k,-r,k-r)\right]=$$

$$=\frac{1}{T-k}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left[\frac{\sin\frac{1}{2}(\omega-\nu)(T-k)}{\sin\frac{1}{2}(\omega-\nu)}\right]\left[1+e^{-i(\omega+\nu)k}\right]f_Y(\omega)f_Y(\nu)d\omega d\nu+$$

$$+\sum_{t,s=1}^{T-k}\chi(k,s-t,s-t+k)=$$

$$=2\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\frac{\sin^2\frac{1}{2}(\omega-\nu)(T-k)}{(T-k)\sin^2\frac{1}{2}(\omega-\nu)}\cos^2\frac{1}{2}(\omega+\nu)kf_Y(\omega)f_Y(\nu)d\omega d\nu+$$

$$+\sum_{t,s=1}^{T-k}\chi(k,s-t,s-t+k)$$

$$k=0,1,\ldots,T-1, \tag{85}$$

где $\chi(k,r,s)=M(Y_t-m^Y)(Y_{t+k}-m^Y)(Y_{t+r}-m^Y)(Y_{t+s}-m^Y)-$

$-\left[R_Y(k)R_Y(r-s)+R_Y(r)R_Y(k-s)+R_Y(s)R_Y(k-r)\right]$- семиинвариант четвертого порядка.

Если последовательность $Y=\{Y_t, t=0,\pm1,\ldots\}$ гауссовская, то семиинварианты четвертого порядка в выражениях (85) обращаются в нуль.

Более трудоемко вычисление выражений для дисперсий смещенных оценок $\hat{R}_Y(k)$ и $R_Y^*(k)$; сами выражения дисперсий более громоздки и мы их не приводим.

Отметим некоторые ассимптотические свойства оценок корреляционоой функции. При известном среднем m^Y оценка $\hat{R}_Y(k)$, как уже отмечалось выше, является несмещенной. При неизвестном среднем оценки $\hat{R}_Y(k)$ и $R_Y^*(k)$, как показывают, в частности, формулы (79)-(83), являются смещенными, причем смещение содержит множитель $1/T$. Более точно поведение оценок $\hat{R}_Y(k)$ и $R_Y^*(k)$ при $T\to\infty$ выражается следующей теоремой.

Т е о р е м а 4. Если $\sum_{r=-\infty}^{\infty} R_Y(r) < \infty$, то оценки $\hat{R}_Y(k)$ и $R_Y^*(k)$ являются ассимптотическими несмещенными и

$$\lim_{T\to\infty} T\left[M\hat{R}_Y(k) - R_Y(k)\right] = \lim_{T\to\infty} T\left[MR_Y^*(k) - R_Y(k)\right] = -\sum_{r=-\infty}^{\infty} R_Y(r).$$

Если $f_Y(\omega)$ непрерывна при $\omega = 0$, то

$$\lim_{T\to\infty} T\left[M\hat{R}_Y(k) - R_Y(k)\right] = \lim_{T\to\infty} T\left[MR_Y^*(k) - R_Y(k)\right] = -2\pi f_Y(0).$$

Вернемся к дисперсиям оценок корреляционной функции. Если для с.п. $Y = \{Y_t, t = 0, \pm 1, \ldots\}$, $\sum_{r=-\infty}^{\infty} R_Y^2(r) < \infty$ и $\left|\sum_{r=-\infty}^{\infty} \chi(k, -r, k-r)\right| < \infty$, то предельная дисперсия величины $\sqrt{T-k} \cdot \tilde{R}_Y(k)$ будет определятся соотношением

$$\lim_{T\to\infty} (T-k) D\hat{R}_Y(k) = \sum_{r=-\infty}^{\infty}\left[R_Y^2(r) + R_Y(r)R_Y(r+2k)\right] +$$

$$+ \sum_{r=-\infty}^{\infty} \chi(k, -r, k-r) =$$

$$= 4\pi\int_{-\pi}^{\pi} \cos^2 \omega k f_Y^2(\omega) d\omega + \sum_{r=-\infty}^{\infty} \chi(k, -r, k-r). \qquad (86)$$

Соотношение (86) вместе с асимптотической несмещенностью означают состоятельность оценки $\hat{R}_Y(k)$.

Если $\chi(r, s, t) = 0$ (что имеет место для гауссовских с.п.) и $\sum_{r=-\infty}^{\infty} \left|R_Y(r)\right| < \infty$, то предельные дисперсии величин $\sqrt{T-k} \cdot \tilde{R}_Y(k)$ стремятся к $\sum_{r=-\infty}^{\infty} R_Y^*(r)$, когда $k \to \infty$.

Можно установить, что если моменты с.п. $Y = \{Y_t, t = 0, \pm 1, \ldots\}$ до четвертого порядка включительно соответствуют стационарности, и

$$\sum_{r=-\infty}^{\infty} R_Y(r) < \infty \text{ и } \left|\sum_{r,s,t=-\infty}^{\infty} \chi(r, s, t)\right| < \infty,$$

то разность между $(T-k)D\tilde{R}_Y(k)$ и $(T-k)D\hat{R}_Y(k)$, а также между $(T-k)D\tilde{R}_Y(k)$ и $(T-k)DR_Y^*(k)$ имеет порядок $1/(T-k)$. Это означает, во-первых, что оценки $\hat{R}_Y(k)$ и $R_Y^*(k)$ состоятельны, а, вовторых, что для «больших выборок» величину $D\tilde{R}_Y(k)$ можно использовать как апроксимацию для $D\hat{R}_Y(k)$ и $DR_Y^*(k)$.

3.3. Оценивание параметров модели методом моментов

Одним из общих методов нахождения оценок параметров является метод моментов. Он заключается в приравнивании определенного количества выборочных моментов к соответствующим теоретическим параметрам, которые являются функциями от неизвестных параметров. Рассматривая количество моментов, равное числу оцениваемых параметров, получаем искомые оценки. На практике этод метод приводит к сравнительно простым вычислениям. Мы рассмотрим применение этого метода последовательно к оцениванию параметров последовательности АР(p), СС(q) и АРСС(p,q).

Для уравнений с.п.АР(p) метод моментов сводится к решению системы уравнений Юла-Уокера (11) для $k = 1,\dots,p$ относительно параметров $\beta_1,\dots,\beta_p$. Вместо теоретических значений корреляционной функции $R_Y(k)$ следует подставить их оценки. В качестве таких оценок можно использовать $\tilde{R}_Y(k)$, $\hat{R}_Y(k)$ и $R_Y^*(k)$ из предыдущего раздела. Поскольку эти оценки состоятельны, то при больших T с вероятностью, близкой к единице, выборочная корреляционная матрица будет невырожденной и уравнения Юла-Уокера будут иметь решение $\hat{\beta} = \left(\hat{\beta}_1,\dots,\hat{\beta}_p\right)$, которое будет также состоятельной оценкой вектора $\beta = \left(\beta_1,\dots,\beta_p\right)$. Однако оценки, найденные с помощью метода моментов, с точки зрения эффективности не являются наилучшими из возможных и даже при больших T они имеют наименьшую возможную дисперсию. В рассматриваемой модели с.п.АР(p) дисперсия «шума» σ_X^2 тоже может являтся неизвестным параметром и для ее оценки может быть использованно равенство (14), в котором значения $R_Y(k), k = \overline{0,p}$ следует заменить их оценками, как при оценивании β, а сами параметры $\beta_k, k=\overline{1,p}$ в (14) должны быть заменены уже найденными оценками $\hat{\beta}_k, k=\overline{1,p}$.

Приведем два примера оценок, полученных по методу моментов.

1. Последовательность АР(1), определяемая уравнением (8), содержит один неизвестный параметр β_1. Уравнение Юла-Уокера имеет вид $R_Y(1)=\beta_1 R_Y(0)$. Подставляя в него вместо $R_Y(0)$ и $R_Y(1)$ оценки $\hat{R}_Y(0)$ и $\hat{R}_Y(1)$ вида (77), получаем

$$\hat{\beta}_1 = \hat{R}_1(1)/\hat{R}_1(0) =$$

$$= \frac{T}{T-1}\left[\sum_{t=1}^{T-1}(y_t-\bar{y})(y_{t+1}-\bar{y})\right]\Big/\sum_{t=1}^{T}(y_t-\bar{y})^2 . \qquad (87)$$

В нашем примере равенство (14) имеет вид

$$R_Y(0)=\beta_1 R_Y(1)+\sigma_X^2 .$$

Отсюда получается оценка дисперсии «шума»

$$\hat{\sigma}_X^2 = \hat{R}_Y(0)-\hat{\beta}_1\hat{R}_Y(1)=\hat{R}_Y^2\ (0)\left[1-\left(\hat{R}_Y(1)/\hat{R}_Y(0)\right)^2\right].$$

2. Последовательность АР(2), определяемая уравнением (9), содержит неизвестные параметр β_1 и β_2. Система уравнений Юла-Уокера имеет вид

$$\left.\begin{array}{l} R_Y(1)=\beta_1 R_Y(0)+\beta_2 R_Y(1) \\ R_Y(2)=\beta_1 R_Y(1)+\beta_2 R_Y 0 \end{array}\right\}$$

и дает оценки

$$\hat{\beta}_1 = \frac{\hat{R}_1(1)\left[\hat{R}_1(0)-\hat{R}_1(2)\right]}{\hat{R}_Y^2(0)-\hat{R}_Y^2(1)}, \hat{\beta}_2 = \frac{\hat{R}_1(0)\hat{R}_1(2)-\hat{R}_Y^2(1)}{\hat{R}_Y^2(0)-\hat{R}_Y^2(1)} .$$

Оценка для дисперсии «шума» σ_X^2 будет иметь вид

$$\hat{\sigma}_X^2 = \frac{\hat{R}_Y(0)\left[\hat{R}_Y^2(0)-2\hat{R}_Y(1)+\hat{R}_Y^2(2)\right]}{\hat{R}_Y^2(0)-\hat{R}_Y^2(1)},$$

В случае последовательности СС(q) метод моментов приводит к системе нелинейных уравнений относительно параметров $\alpha_1,\ldots,\alpha_q$, если в (33) полагать $k=1,\ldots,q$. Два метода решения такой нелинейной системы изложены в [3, с.226-229]. Приведем здесь один из них. Это так называемый линейно сходящийся итеративный процесс.

Из выражений (34) и (33) для корреляционной функции процесса СС(q) можно найти оценки параметров σ_X^2, $\alpha_q, \ldots, \alpha_1$ точно в том порядке, как здесь указано, при помощи итераций

$$\hat{\sigma}_X^2 = \frac{\hat{R}_Y(0)}{1 + \alpha_1^2 + \ldots + \alpha_q^2},$$
$$\alpha_k = \frac{\hat{R}_Y(k)}{\hat{\sigma}_X^2} - \left(\alpha_1 \alpha_{k+1} + \ldots + \alpha_{q-k} \alpha_q\right) \tag{88}$$

с условием, что $\alpha_0 = 1$. Параметры $\alpha_1, \ldots, \alpha_q$ приравниваются к нулю в самом начале итеративной процедуры; значения α_k и σ_X^2, используемые в любом цикле вычисления - это последние из доступных оценок этих величин. Например, в случае $q = 2$ уравнения (88) имеют вид

$$\hat{\sigma}_X^2 = \frac{\hat{R}_Y(0)}{1 + \alpha_1^2 + \alpha_2^2},$$
$$\alpha_2 = \frac{\hat{R}_Y(2)}{\sigma_X^2},$$
$$\alpha_1 = \frac{\hat{R}_Y(1)}{\sigma_X^2} - \alpha_1 \alpha_2.$$

В случае $q = 1$ уравнения (88) приобретают вид

$$\hat{\sigma}_X^2 = \frac{\hat{R}_Y(0)}{1 + \alpha_1^2},$$
$$\alpha_1 = \frac{\hat{R}_Y(1)}{\sigma_X^2}.$$

После исключения из этих уравнений величины σ_X^2 приходим к квадратному уравнению относительно α_1 и находим оценку

$$\hat{\alpha}_1 = \frac{1 - \sqrt{1 - 4\hat{r}_Y^2(1)}}{2\hat{r}_Y(1)},$$
$$\left|\hat{r}_Y(1)\right| \le \frac{1}{2}, \tag{89}$$

в которой $\hat{r}_Y(1) = \hat{R}_Y(1) / \hat{R}_Y(0)$. Затем получаем оценку

$$\hat{\sigma}_X^2 = \frac{2\hat{r}_Y(1)\hat{R}_Y(1)}{1-\sqrt{1-4\hat{r}_Y^2(1)}}. \qquad (90)$$

Наконец, рассмотрим оценку параметров в смешанной модели АРСС(p,q). Для оценки параметров авторегрессии $\beta_1,\ldots,\beta_p$ можно использовать p-уравнения вида (47) для $k = q+1,\ldots,q+p$, в которых значения корреляционной функции $R_Y(r)$ заменяются выборочными значениями (оценками). Затем введем вспомогательный процесс W, полагая

$$W_t = Y_t - \sum_{j=1}^{p}\beta_j Y_{t-j}, \qquad t = 0,\pm1,\ldots \qquad (91)$$

Тогда уравнению (43) можно придать вид

$$W_t = \sum_{r=0}^{q}\alpha_j X_{t-r}$$

и, следовательно, с.п.W можно рассматривать как «чистую» последовательность СС(q). Исходя из равенства (91), можно выразить корреляционную функцию последовательности W через значения корреляционной функции последовательности Y. Можно показать, что

$$R_W(j) = \sum_{i=0}^{p}\beta_i^2 R_Y(j) + \sum_{i=1}^{p}\left(\beta_0\beta_i + \beta_1\beta_{i+1} + \ldots + \beta_{p-i}\beta_p\right)d_j, \qquad (92)$$

где $j = 0,1,\ldots,q$; $d_j = R_Y(j+i) + R_Y(j-i)$, $\beta_0 = -1$.

Далее, пользуясь уже найдеными оценками для $R_Y(r)$ и параметров $\beta_1,\ldots,\beta_p$, по формулам (92) будем иметь оценки для корреляционной функции $R_W(r), j = 0,1,\ldots,q$. На заключительном этапе используем описанный выше линейно сходящийся итеративный процесс для оценки неизвестных σ_X^2, $\alpha_1,\ldots,\alpha_q$ либо (в частном случае $q = 1$) воспользуемся готовыми формулами (89) и (90), в которых оценки $\hat{R}_Y(0)$ и $\hat{R}_Y(1)$ должны быть заменены оценками $\hat{R}_W(0)$ и $\hat{R}_W(1)$.

3.4. Оценивание параметров

методом максимального правдоподобия

Использованный в предыдущем разделе метод моментов приводит, при определенных условиях, к состоятельным оценкам. Однако оценки по методу моментов, вообще говоря, неэффективны. К тому же метод моментов неприменим, когда теоретические моменты нужного порядка не существуют. Одним из популярных общих методов получения оценок параметров, обладающих рядом преимуществ по сравнению с методом моментов, является метод максимального правдоподобия (м.м.п.).Напомним суть этого метода.

Пусть имеем вектор наблюдений $y^T = (y_1, \ldots y_T)$ случайной последовательности $Y = \{Y_t, t = 0, \pm 1, \ldots\}$ и $f(y^T; \theta) = f(y_1, \ldots y_T | \theta)$ - совместная плотность этих наблюдений. Векторный параметр $\theta = (\theta_1, \ldots \theta_k)$ принадлежит параметрическому множеству $\Theta \subset R^k$. *Функция правдоподобия* (при фиксированных наблюдениях) определяется как следующая функция параметра θ:

$$L(\theta | y^T) = f(\theta | y^T). \tag{93}$$

Принцип максимального правдоподобия (м.п.) предписывает выбор в качестве оценки параметра θ такого значения $\hat{\theta} = (\hat{\theta}_1, \ldots \hat{\theta}_k)$, для которого

$$L(\hat{\theta} | y^T) = \sup_{\theta \subset \Theta} L(\theta | y^T). \tag{94}$$

Верхняя грань может не достигаться, и для таких случаев вводится модифицированная оценка м.п. θ^*, определяемая условием

$$L(\hat{\theta} | y^T) \geq c \sup_{\theta \subset \Theta} L(\theta | y^T),$$

где c - фиксированное число, $0 < c < 1$. Может случится, что для некоторых наблюдений θ^* и θ не существуют. Однако можно показать, что при условиях регулярности на $f(\theta | y^T)$ такими наблюдениями можно пренебречь.

Часто удобно работать не с функцией $L(\theta | y^T)$, а с ее логарифмом: $l(\theta | y^T) = \ln L(\theta | y^T)$. Условие (94) переходит в

$$l(\hat{\theta} | y^T) = \sup_{\theta \subset \Theta} l(\theta | y^T). \tag{95}$$

Если супремум в (95) достигается во внутренней точке множества Θ

и $l\left(\theta|y^T\right)$ дифференцируема по θ, то в точке θ^* дожны обращатся в нуль частные производные функции $l\left(\theta|y^T\right)$. Следовательно, в этом случае $\hat{\theta}$ удовлетворяет уравнениям

$$\frac{\partial}{\partial\theta_i} l\left(\theta|y^T\right) = 0,\ i = 1,\ldots k. \qquad (96)$$

Уравнения (96) называются уравнениями м.п., а всякое их решение - оценкой м.п. Оценку, определяемую условием (95), называют оценкой метода м.п.

Общие свойства оценок м.п. по независимым наблюдениям можно найти, например, в [7]. Нас же будут интересовать оценки м.п. параметров по зависимым наблюдениям, а именно, по с.п. АР(p), СС(q) и АРСС(p,q). Для получения оценок м.п., согластно вышесказанному, нужно прежде всего найти функцию правдоподобия $L\left(\theta|y^T\right)$. Однако, удобные для использования формулы для $L\left(\theta|y^T\right)$ удается получить только в предположении, что входная последовательность $X = \left\{X_t, t = 0, \pm 1, \ldots\right\}$ - гауссовская. Тогда в силу линейности преобразований, осуществляемые формулами (7), (32) и (43), выходная последовательность $Y = \left\{Y_t, t = 0, \pm 1, \ldots\right\}$ тоже будет гауссовской. Если $X_t \sim N\left(0, \sigma_X^2\right)$ и обеспечена стационарность последовательности Y, то вектор $Y^T = \left(Y_1, \ldots, Y_T\right)$ будет гауссовским с нулевым вектором средних значений и корреляционной матрицей $R_T = \left\|R\left(k-j\right)\right\|_{k,j=\overline{1,T}}$, где $R\left(k\right) = R\left(-k\right)$ находятся по формулам (33) для последовательности СС(q) и из уравнений (11) и (46) дял последовательностей АР(p) и АРСС(p,q) соответственно. Следовательно, плотность вектора Y^T будет иметь вид

$$f\left(y^T\right) = \frac{1}{\left(2\pi\right)^{\frac{n}{2}}\left|\det R_t\right|^{\frac{1}{2}}} e^{\frac{1}{2}\left(R_T^{-1}y^T, y^T\right)}, \qquad (97)$$

где $y^T = \left(y_1, \ldots, y_T\right)$, R_T^{-1} - матрица, обратная к корреляционной, символ $\left(\cdot,\cdot\right)$ обозначает скалярное произведение в эвклидовом пространстве размерности T. Практическому использованию плотности (9) при больших T препятствуют трудности обращению матрицы R. Однако м.м.п. удается реализовать для конкретных моделей гауссовских случайных последовательностей, к рассмотрению которых мы и перейдем.

М.м.п. для с.п. АР(p). В соответствии с идеей м.м.п. начнем с рассмотрения функции правдоподобия. Предположим, что на вход авторегрессионной модели (7) подается гауссовский «шум» $X = \{X_t, t = 0, \pm1, \ldots\}$ с $MX_t = 0$ и $DX_t = \sigma_X^2$. Так как модель (7) осуществляет линейное преобразование, то выходной «сигнал» $Y = \{Y_t, t = 0, \pm1, \ldots\}$, как известно, тоже будет гауссовским. Если модель (7) стационарна, то с.п. Y стационарна $MY_t = 0$ и корреляционной функцией, определяемой разностным уравнением вида (11). Для гауссовского вектора $Y^{p+T} = (Y_1, \ldots, Y_p, Y_{p+1}, \ldots, Y_{p+T})$ плотность распределения может быть записана в соответствии с формулой (97)

$$f(y^{p+T}; \beta, \sigma_X) =$$

$$= (2\pi)^{-(p+T)/2} \left|\det R_{p+T}\right|^{-1/2} \cdot \exp\left\{-\frac{1}{2}\left(R_{p+T}^{-1} y^{p+T}, y^{p+T}\right)\right\}, \qquad (98)$$

где $\beta = (\beta_1, \ldots, \beta_p)$-вектор параметров модели (7), $y^{p+T} = (y_1, \ldots, y_{p+T})$, $R_{p+T} = \left\| R_Y(k-j) \right\|_{k,j=\overline{1,p+T}}$ - корреляционная матрица вектора Y^{p+T}.

При использовании плотности (98) можно избежать обращения матрицы R_{p+T}. Для этого воспользуемся представлением

$$f(y^{p+T}; \beta, \sigma_X) = (y^p; \beta, \sigma_X) \cdot (y_{p+1}, \ldots, y_{p+T}; \beta, \sigma_X | y^p), \qquad (99)$$

где первый множитель правой части - безусловная плотность вектора $Y^p = (Y_1, \ldots, Y_p)$, получающаяся из (98) при $T = 0$, а второй - условная плотность вектора $(Y_{p+1}, \ldots, Y_{p+T})$ при фиксированном значении вектора $Y^p = y^p = (y_1, \ldots, y_p)$. Эту условную плотность можно получить, используя плотность распределения вектора $(X_{p+1}, \ldots, X_{p+T})$

$$\varphi(x_{p+1}, \ldots, x_{p+T}; \sigma_X) =$$

$$= \left(2\pi\sigma_X^2\right)^{-T/2} \cdot \exp\left\{\frac{1}{2\sigma_X^2}\sum_{k=p+1}^{p+T} x_X^2\right\}, \qquad (100)$$

Но при фиксированном векторе $Y^p = y^p$ вектор $\left(X_{p+1},\ldots,X_{p+T}\right)$ и $\left(Y_{p+1},\ldots,Y_{p+T}\right)$ связаны преобразованием

$$X_{p+1} = Y_{p+1} - \sum_{j=1}^{p}\beta_j Y_{p+1-j}$$

....

$$X_{p+T} = Y_{p+T} - \sum_{j=1}^{p}\beta_j Y_{p+T-j}$$

с единичным якобианом.

Следовательно,

$$f\left(y_{p+1},\ldots,y_{p+T};\beta,\sigma_X\right) =$$

$$= \varphi\left[y_{p+1} - \sum_{j=1}^{p}\beta_{p+1}y_{p+1-j},\ldots,y_{p+T} - \sum_{j=1}^{p}\beta_{p+1}y_{p+T-j};\sigma_X\right] =$$

$$= \left(2\pi\sigma_X^2\right)^{-T/2} \cdot \exp\left\{\frac{1}{2\sigma_X^2}\sum_{k=1}^{T}\left(y_{p+k} - \sum_{j=1}^{p}\beta_j y_{p+k-j}\right)^2\right\} \qquad (101)$$

Кроме того,

$$f\left(y^p;\beta,\sigma_X\right) =$$

$$= \left(2\pi\right)^{-p/2}\left|\det R_p\right|^{-p/2} \cdot \exp\left\{-\frac{1}{2}\left(R_{p+T}^{-1}y^p, y^p\right)\right\},$$

Учитывая эту формулу и (101), запишем (99) в виде

$$f\left(y^{p+T};\beta,\sigma_X\right) =$$

$$= \left(2\pi\right)^{-(p+T)/2}\sigma_X^{-T}\left|\det R_{p+T}\right|^{-1/2} \cdot \exp\left\{-\frac{S(\beta)}{2\sigma_X^2}\right\}, \qquad (102)$$

В предыдущей формуле

$$S(\beta) = \sigma_X^2\left(R_p^{-1}y^p, y^p\right) +$$

$$+ \sum_{k=1}^{T}\left(y_{p+k} - \sum_{j=1}^{p}\beta_j y_{p+k-j}\right)^2 \qquad (103)$$

Заметим, что $R_p = \sigma_Y^2 r_p$, где $r_p = \left\|r_Y\left(k-j\right)\right\|_{k,j=\overline{1,p}}$ - матрица, составленная из значений нормированной корреляционной функции, не зависящих от σ_X^2, а величина σ_Y^2, согласно формуле (15), пропорциональна σ_X^2. Следовательно $\det R_p = \left(\sigma_X^{2p}/c^p(\beta)\right)\det r_p$ и

$\sigma_X^2\left(R_p^{-1}y^p, y^p\right) = c(\beta)\left(r_p^{-1}y^p, y^p\right)$, где $c(\beta)$ - знаменатель правой части (15). Тогда логарифм функции правдоподобия $l\left(\beta, \sigma_X | y^{p+T}\right) =$ $= \ln L\left(\beta, \sigma_X | y^{p+T}\right) = \ln f\left(y^{p+T}; \beta, \sigma_X\right)$ будет иметь вид

$$l\left(\beta, \sigma_X | y^{p+T}\right) =$$
$$= -(p+T)\ln\left(\sigma_X\sqrt{2\pi}\right) - \frac{1}{2}\ln\left(\det r_p / c^p(\beta)\right) - \frac{S(\beta)}{2\sigma_X^2}. \qquad (104)$$

Наряду с параметрами $\beta_1, \ldots, \beta_p$ будем считать неизвестной также дисперсию σ_X^2 входной последовательности. Тогда уравнения правдоподобия запишутся в виде

$$\frac{\partial l}{\partial \sigma_X} = -\frac{p+T}{\sigma_X} + \frac{S(\beta)}{\sigma_X^3} = 0, \qquad (105)$$

$$\frac{\partial l}{\partial \beta_j} = -\frac{1}{2}\frac{\partial}{\partial \beta_j}\ln\left(\det r_p / c^p(\beta)\right) -$$
$$-\frac{1}{2\sigma_X^2}\frac{\partial}{\partial \beta_j}\left(c(\beta)\left(r_p^{-1}y^p, y^p\right)\right) +$$
$$+\frac{1}{\sigma_X^2}\sum_{k=1}^{T} y_{p+k-j}\left(y_{p+k} - \sum_{m=1}^{p}\beta_m y_{p+k-m}\right) = 0,$$
$$j = \overline{1,p}. \qquad (106)$$

Выражая дисперсию «шума» σ_X^2 из (105) и подставляя ее в систему (106), будем иметь систему относительно неизвестного вектора $\beta = \left(\beta_1, \ldots, \beta_p\right)$. Однако при решении получающейся системы возникают трудности, поскольку в общем случае величины

$$M_j = -\frac{1}{2}\frac{\partial}{\partial \beta_j}\ln\left(\det r_p / c^p(\beta)\right), \ j = \overline{1,p}$$

- сложные функции β. Продемонстрируем эти трудности на простейших случаях.

1. **С.п.АР(1).** В этом случае $c(\beta_1) = 1 - \beta_1^2$, $r_1 = r_Y(0) = 1$ и из (105) и (106) получаем кубическое уравнение относительно β_1

$$\beta_1^3 \frac{T}{T+1}\sum_{k=2}^{T} y_k^2 - \beta_1^2 \frac{T-1}{T+1}\sum_{k=1}^{T} y_k y_{k+1} -$$

$$-\beta_1\left(\frac{T+2}{T+1}\sum_{k=2}^{T} y_k^2+\frac{y_1^2+y_{T+1}^2}{T+1}\right)+\sum_{k=1}^{T} y_k y_{k+1}=0,$$

корни которых в принципе могет быть точно найдены с использованием формул Г.Кардано. Однако если все три корня вещественны, то предстоит еще выяснить, какой из них доставляеи величине $l\left(\beta_1,\sigma_X|y^{T+1}\right)$ наибольшее значение.

2. **С.п.АР(2).** Имеем $c(\beta_1)=\left[(1-\beta_2)^2-\beta_1^2\right](1+\beta_2)/(1-\beta_2)$,

$$r_2=\frac{1}{1-\beta_2}\begin{pmatrix}1-\beta_2 & \beta_1\\ \beta_2 & 1-\beta_2\end{pmatrix}$$

и система (106) получает вид

$$\left[(1-\beta_2)^2-\beta_1\right]\cdot\left\{\beta_1\sum_{k=2}^{T+1} y_k^2+\beta_2\sum_{k=2}^{T} y_k y_{k+1}-\sum_{k=1}^{T} y_k y_{k+1}\right\}+$$

$$\frac{\beta_1}{T+2}S(\beta_1,\beta_2)=0,$$

$$(1+\beta_2)\left[(1-\beta_2)^2-\beta_1^2\right]\cdot$$

$$\cdot\left\{\beta_2\sum_{k=3}^{T} y_k^2+\beta_1\sum_{k=2}^{T} y_k y_{k+1}-\sum_{k=1}^{T} y_k y_{k+1}\right\}+$$

$$\frac{\beta_1^2+2\beta_2(1-\beta_2)}{T+2}S(\beta_1,\beta_2)=0,$$

где

$$S(\beta_1,\beta_2)=\beta_1^2\sum_{k=2}^{T+1} y_k^2+\beta_2^2\sum_{k=3}^{T} y_k^2+2\beta_1\beta_2\sum_{k=2}^{T} y_k y_{k+1}+$$

$$-2\beta_1\sum_{k=1}^{T} y_k y_{k+1}-2\beta_2\sum_{k=1}^{T} y_k y_{k+2}+\sum_{k=1}^{T+2} y_k^2.$$

В этом случае точное решение системы относительно неизвестных β_1 и β_2 уже не может быть найдено.

Итак, для получения оценки вектора β в общем случае нужно либо искать приближенное решение системы (106), либо, отказавшись от ее решения, максимизировать вместо логарифма функции правдоподобия другую, но близкую к нему в каком-то смысле функцию. Ясно, что второй подход приведет к оценкам, которые не будут оценками м.п., но их все-таки естественно называть оценками квазимаксимального правдоподобия.

Оценки квазимаксимального правдоподобия. Примем во внимание, что в логарифме функции правдоподобия (104) слагаемое, содержащее $\det r_p$, не зависит от T и мало по сравнению с остальными двумя слагаемыми, растущими с ростом T. Пренебрегая этими слагаемыми, получим функцию

$$\tilde{l}\left(\beta,\sigma_X|y^{p+T}\right) = -(p+T)\ln\left(\sigma_X\sqrt{2\pi}\right) - \frac{S(\beta)}{2\sigma_X^2}. \tag{107}$$

Крому того заметим, что величина $S(\beta)$, определенная неравенством (103), не только квадратичная форма от наблюдений y^{p+T}, но (как показано в работе [3,с.304]) и от параметров β. Если положить $\beta' = \left(1,\beta_1,\ldots,\beta_p\right)$, то для некоторой матрицы D размера (p+1)x(p+1) справедливо равенство

$$S(\beta) = \left(D\beta',\beta'\right), \tag{108}$$

причем

$$D = \begin{pmatrix} D_{11} & -D_{12} & \cdots & -D_{1,p+1} \\ -D_{12} & D_{22} & \cdots & D_{2,p+1} \\ \vdots & \vdots & & \vdots \\ -D_{1,p+1} & D_{2,p+1} & \cdots & D_{p+1,p+1} \end{pmatrix} \tag{109}$$

а элементы $D_{i,j}$ определяются как суммы

$$D_{i,j} = D_{j,i} = y_i y_j + y_{i+1}y_{j+1} + \ldots + y_{p+T+1-i}y_{p+T+1-j} \tag{110}$$

содержащие $p+T-(i-1)-(j-1)$ слагаемых. Тогда оценки $\tilde{\beta}$ параметров β, полученные путем максимизацмм (107), совпадают с *оценками наименьших квадратов*, полученными минимизацией $S(\beta)$. Дифференцируя (108) по β_j, $j=\overline{1,p}$ и приравнивая к нулю соответствующие производные, получим систему

$$\begin{cases} \beta_1 D_{22} + \beta_2 D_{23} + \ldots + \beta_p D_{2,p+1} = D_{12}, \\ \beta_1 D_{23} + \beta_2 D_{33} + \ldots + \beta_p D_{3,p+1} = D_{13}, \\ \qquad \cdots \\ \beta_1 D_{2,p+1} + \beta_2 D_{3,p+1} + \ldots + \beta_p D_{p+1,p+1} = D_{1,p+1}, \end{cases} \tag{111}$$

которая в очевидных матричныз обозначениях имеет вид

$$D_p\beta = d,$$

откуда получаем оценку *квазиоптимального правдоподобия*

$$\tilde{\beta} = D_p^{-1} d \qquad (112)$$

Приближенные оценки м.п. параметров β можно получить изменяя сами уравнения правдоподобия (106). Сначала производную $\partial l / \partial \beta_j$ запишем в виде

$$\frac{\partial l}{\partial \beta_j} = M_j + \sigma_X^{-2}\left(D_{1,j+1} - \beta_1 D_{2,j+1} - \ldots - \beta_p D_{p+1,j+1}\right), \qquad (113)$$

$$j = \overline{1,p}.$$

Беря математические ожидания от обеих частей этих равенств и учитывая, что $M\left(\partial l / \partial \beta_j\right) = 0$, получаем

$$M_j \sigma_X^2 + (p + T - j) R_Y(j) - \beta_1 (p + T - j - 1) R_Y(j-1) - \ldots$$

$$\ldots - \beta_p (T - j) R_Y(j - p) = 0. \qquad j = \overline{1,p}. \qquad (114)$$

Теперь вспомним разностное уравнение (11) и, умножим его на $p + T$, запишем в виде

$$(p + T) R_Y(j) - \beta_1 (p + T) R_Y(j-1) - \ldots - \beta_p (p + T) R_Y(j - p) = 0$$

и вычтем из (114). Получим

$$M_j \sigma_X^2 = j R_Y(j) - \beta_1 (j + 1) R_Y(j - 1) - \ldots$$

$$\ldots - \beta_p (j + p) R_Y(j - p), \qquad j = \overline{1,p}. \qquad (115)$$

Используя в качестве оценок значений корреляционной функции $R_Y(j - i)$ величины $D_{i+1,j+1} / (p + T - j - i)$, из (115) получим оценку для M_j, а подставляя ее в (113), получим

$$\frac{\partial l}{\partial \beta_j} \approx (p + T) \sigma_X^{-2} \left\{ \frac{D_{1,j+1}}{p + T - j} - \beta_1 \frac{D_{2,j+1}}{p + T - j - 1} - \ldots - \beta_p \frac{D_{p+1,j+1}}{T - j} \right\},$$

что ведет к системе линейных уравнений согластно формуле (111), в которых вместо $D_{i,j}$ стоят величины

$$D_{i,j}^* = (p + T) D_{i,j} / \left[p + T - (i - 1) - (j - 1)\right].$$

В очевидных матричных обозначениях мы будем иметь систему

$$D_p^* \beta = d^*,$$

из которых получаем приближенную оценку м.п.

$$\beta^* = \left[D_p^*\right]^{-1} d^*.$$

Продемонстрируем различия рассмотренных оценок параметров β на частных случаях $p = 1$ и $p = 2$.

1. **С.п.АР(1).** В этом случае оценка квазимаксимального правдоподобия находится по формуле (112), в которой

$$D_1 = D_{22} = \sum_{k=2}^{T} y_k^2 \text{ и } \qquad d = D_{12} = \sum_{k=1}^{T} y_k y_{k+1}.$$

Тогда

$$\tilde{\beta}_1 = D_1^{-1} d = \sum_{k=1}^{T} y_k y_{k+1} \Big/ \sum_{k=2}^{T} y_k^2 .$$

Приближенная оценка м.п. определяется формулой (116), в которой

$$D_1^* = D_{22}^* = \frac{T+1}{T-1} D_{22}, \quad d^* = D_{12}^* = \frac{T+1}{T} D_{12}.$$

Тогда

$$\beta_1^* = \left(D_1^*\right)^{-1} d^* = \frac{T-1}{T} \sum_{k=1}^{T} y_k y_{k+1} \Big/ \sum_{k=2}^{T} y_k^2 .$$

2. **С.п.АР(2).** Для получения оценки квазимаксимального правдоподобия вектора $\beta = (\beta_1, \beta_2)$ нужно иметь вектор $d = (D_{12}, D_{13})$ и матрицу

$$D_2 = \begin{pmatrix} D_{22} & D_{23} \\ D_{32} & D_{33} \end{pmatrix},$$

причем выражения для величин D_{12} и D_{22} указаны в предыдущем примере, а D_{13}, D_{23} и D_{33} имеют следующий вид:

$$D_{13} = \sum_{k=1}^{T} y_k y_{k+2}, \quad D_{23} = \sum_{k=2}^{T} y_k y_{k+1}, \quad D_{33} = \sum_{k=3}^{T} y_k^2 .$$

Используя (112), находим

$$\tilde{\beta}_1 = \frac{D_{12} D_{33} - D_{13} D_{23}}{D_{22} D_{33} - \left(D_{23}\right)^2},$$

$$\tilde{\beta}_2 = \frac{D_{22} D_{13} - D_{12} D_{23}}{D_{22} D_{33} - \left(D_{23}\right)^2}.$$

Приближенная оценка м.п. $\beta^* = (\beta_1^*, \beta_2^*)$ вектора β может быть получена из предыдущих формул, в которых величины $D_{i,j}$ нужно заменить величинами

$$D_{i,j}^* = (T+2) D_{i,j} \Big/ [T + 4 - i - j] \qquad i, j = 1,2,3.$$

Рассмотренные примеры показывают, что для больших выборок различия между оценками, полученными путем разных приближений, малы. То же самое можно сказать об оценках параметров АР(p) и в общем случае.

М.м.п. для с.п. СС(q). Предположим, что на входе модели скользящего среднего (32) имеем гауссовский «шум» $X = \{X_t, t = 0, \pm 1, \ldots\}$ с $MX_t = 0$ и $DX_t = \sigma_X^2$. При этом выходной «сигнал» $Y = \{Y_t, t = 0, \pm 1, \ldots\}$ тоже будет гауссовским и стационарным с $MY_t = 0$ и корреляционной функцией, определяемой формулой (33). Рассмотрим функцию правдоподобия для с.п.Y соответствующую наблюдениям $y^T = (y_{q+1}, \ldots y_{q+T})$. Гауссовский вектор $Y^T = (Y_{q+1}, \ldots Y_{q+T})$ будет иметь плотность распределения вида

$$f(y^T; \alpha, \sigma_X) =$$

$$= (2\pi)^{-T/2} \left|\det R_T\right|^{-1/2} \cdot \exp\left\{-\frac{1}{2}\left(R_T^{-1} y^T, y^T\right)\right\}, \qquad (117)$$

где $\alpha = (\alpha_1, \ldots, \alpha_q)$ - вектор параметров модели (32) $R_T = \left\|R_Y(k-j)\right\|_{k,j=\overline{q,q+T}}$ - корреляционная матрица вектора Y^T с элементами, определяемыми по формуле (33). Хотя формула (33) проста, обращение матрицы R_T при достаточно больших T сопряжено со значительными вычислительными трудностями. Поэтому рассмотрим удобный способ вычисления квадратичной формы $\left(R_T^{-1} y^T, y^T\right)$. Для этого рассмотрим вектор $X^{q+T} = (X_1, \ldots X_{q+T})$, плотность распределения которого имеет вид

$$\varphi(x^{q+T}; \sigma_X) = \left(2\pi\sigma_X^2\right)^{-(q+T)/2} \cdot \exp\left\{-\frac{\left(x^{q+T}, x^{q+T}\right)}{2\sigma_X^2}\right\} \qquad (118)$$

и вектор $Z^{q+T} = (X_1, \ldots, X_q, X_{q+1}, \ldots, X_{q+T}) = (X^q, X^T)$, который связан с X^{q+T} линейным преобразованием в соответствии с (32):

$$Y_{q+T} = X_{q+k} + \alpha_1 X_{q+k-1} + \ldots + \alpha_q X_k, \qquad k = \overline{1, T}.$$

При этом мы можем записать представление

$$X^{q+T} = AX^q + BY^T,$$

где A - матрица размером $(q+T) \times q$, а B - матрица размером $(q+T) \times T$. Элементы этих матриц представляют собой функции параметров $\alpha = (\alpha_1, \ldots, \alpha_q)$. Замечая, что преобразование вектора

X^{q+T} в Z^{q+T} имеет единичный якобиан, можем записать совместную плотность Y^T и X^q

$$f\left(y^T, x^q; \alpha, \sigma_X\right) = \left(2\pi\sigma_X^2\right)^{-(q+T)/2} \cdot \exp\left\{-\frac{S\left(\alpha, x^q\right)}{2\sigma_X^2}\right\},$$

где $S\left(\alpha, x^q\right) = \left(Ax^q + By^T, Ax^q + By^T\right)$.

Вектор предварительных значений «шума» $x^q = \left(x_1, \ldots, x_q\right)$ на практике обычно не наблюдается, и мы найдем для него линейную относительно Y^T оценку, минимизирующую $S\left(\alpha, x^q\right)$. Пусть $\tilde{x}^q$ некоторая оценка (функция наблюдения Y^T) вектора x^q. Запишем вектор

$$Ax^q + By^T = A\left(x^q - \tilde{x}^q\right) + \left(A\tilde{x}^q + By^T\right)$$

и подставим его в выражение для $S\left(\alpha, x^q\right)$. Получим равенство

$$\begin{aligned} S\left(\alpha, x^q\right) &= \left(A\left(x^q - \tilde{x}^q\right), A\left(x^q - \tilde{x}^q\right)\right) + \\ &+ 2\left(x^q - \tilde{x}^q, A'A\tilde{x}^q + A'By^T\right) + \\ &+ \left(A\tilde{x}^q + By^T, A\tilde{x}^q + By^T\right), \end{aligned}$$

где штрих обозначает транспонирование матрицы. Если выбрать $\tilde{x}^q$ так, что

$$A'A\tilde{x}^q = -A'By^T \tag{119}$$

то $S\left(\alpha, x^q\right)$ будет иметь вид

$$S\left(\alpha, x^q\right) = S(\alpha) + \left(A\left(x^q - \tilde{x}^q\right), A\left(x^q - \tilde{x}^q\right)\right) \tag{120}$$

где $S(\alpha) = \left(A\tilde{x}^q + By^T, A\tilde{x}^q + By^T\right)$ - функция только наблюдений y^T, но не предварительных значений «шума».

Из (120) видно, что $S\left(\alpha, x^q\right)$ достигает минимума при $x^q = \tilde{x}^q$, т.е $\tilde{x}^q$ - оценка наименьших квадратов. При этом совместную плотность Y^T и X^q можно представить в виде

$$\begin{aligned} & f\left(y^T, x^q; \alpha, \sigma_X\right) = \\ & = \left(2\pi\sigma_X^2\right)^{-(q+T)/2} \cdot \exp\left\{-\frac{S(\alpha) + \left(A'A\left(x^q - \tilde{x}^q\right), x^q - \tilde{x}^q\right)}{2\sigma_X^2}\right\}. \end{aligned}$$

С другой стороны, ей можно придать вид

$$f\left(y^{T}, x^{q}; \alpha, \sigma_{X}\right) = f\left(y^{T}; \alpha, \sigma_{X}\right) \cdot \varphi\left(x^{q} | y^{T}; \alpha, \sigma_{X}\right),$$

где $\varphi\left(x^{q} | y^{T}; \alpha, \sigma_{X}\right)$ - условная (гауссовская) плотность предварительных значений «шума» при фиксированных наблюдениях Y^{T}. Следовательно

$$\varphi\left(x^{q} | y^{T}; \alpha, \sigma_{X}\right) = \left(2\pi\sigma_{X}^{2}\right)^{-q/2} \left[\det\left(A'A\right)\right]^{-1/2} \cdot$$

$$\cdot \exp\left\{-\frac{\left(A'A\left(x^{q} - \tilde{x}^{q}\right), x^{q} - \tilde{x}^{q}\right)}{2\sigma_{X}^{2}}\right\}, \tag{121}$$

$$f\left(y^{T}; \alpha, \sigma_{X}\right) = \left(2\pi\sigma_{X}^{2}\right)^{-T/2} \left[\det\left(A'A\right)\right]^{-1/2} \cdot \exp\left\{-\frac{S(\alpha)}{2\sigma_{X}^{2}}\right\} \tag{122}$$

Сделаем следующие выводы

1. Из (121) видно, что $\tilde{x}^{q}$ является условным математическим ожиданием вектора X^{q} при заданных при наблюдениях $Y^{T} = y^{T}$ и фиксированном α. Используем обозначение

$$\tilde{X}^{q} = \left[X^{q} | y^{T}, \alpha\right] = \left[X^{q}\right]$$

имеем $\left[X^{q+T}\right] = A\left[X^{q}\right] + BY^{T}$

и

$$S(\alpha) = \left(A\left[x^{q}\right] + BY^{T}, A\left[x^{q}\right] + BY^{T}\right) =$$

$$= \left(\left[x^{q+T}\right], \left[x^{q+T}\right]\right) = \sum_{k=1}^{q+T}\left[x_{k}\right]^{2}. \tag{123}$$

Хотя $\tilde{x}^{q}$ можно получить непосредственно методом наименьших квадратов, на практике его легче вычислять, используя тот факт, что $\tilde{x}^{q} = \left[x^{q}\right]$, и получая $\left[x^{q}\right]$ по методике «прогнозирования» назад описанной в работе [3,разд.7.1.4 и 7.1.5]

Сравнивая (122) и (117), получаем

$$\det\left(A'A\right) = \sigma_{X}^{-2T} \det R_{T}$$

и

$$S(\alpha) = \sigma_{X}^{2}\left(R_{T}^{-1} y^{T}, y^{T}\right).$$

Чтобы вычислить $S(\alpha) = \sum_{k=1}^{q+T}\left[x_{k}\right]^{2}$, можно получить величины $\left[x_{k}\right]$, $k = \overline{1, q+T}$, пользуясь оценками предварительных значений «шума»

$$\left[x^q\right]=\left(\left[x_1\right],\ldots,\left[x_q\right]\right),$$

полученных методом наименьших квадратов (или по методике «прогнозирования назад»), и, вычислив $\left[x_{q+1}\right],\ldots,\left[x_{q+T}\right]$ - реккурентным путем по формулам, получающимся из (32), применяя к обеим частям равенства операции условного математического ожидания при фиксированных наблюдениях y^T

$$\left[x_k\right]=\left[y_k\right]-\alpha_1\left[x_{k-1}\right]-\ldots-\alpha_q\left[x_{k-q}\right],\ k=\overline{q+1,q+T},$$

где $\left[y_k\right]=y_k$.

Наконец, пользуясь (122) и (123), получаем точное выражение для безусловной функции правдоподобия

$$L\left(y^T;\alpha,\sigma_X\right)=$$

$$=\left(2\pi\sigma_X^2\right)^{-T/2}\left[\det\left(A'A\right)\right]^{1/2}\cdot\exp\left\{-\frac{1}{2\sigma_X^2}\sum_{k=1}^{q+T}\left[x_k\right]^2\right\}. \qquad (124)$$

Таким образом, принципиально открыт путь к получению оценок м.п. Если наряду с α неизвестно и σ_X, то уравнения правдоподобия будут иметь вид

$$-\frac{T}{\sigma_X}+\frac{1}{\sigma_X^3}\sum_{k=1}^{q+T}\left[x_k\right]^2=0 \qquad (125)$$

и

$$-\frac{1}{2}\frac{\partial}{\partial\alpha_j}\ln\left[\det\left(A'A\right)\right]-\frac{1}{\sigma_X^2}\sum_{k=1}^{q+T}\left[x_k\right]\frac{\partial\left[x_k\right]}{\partial\alpha_j}=0 \qquad (126)$$

где $j=\overline{1,q}$.

Следует отметить, что $\left[x_k\right]$ - всегда нелинейные функции параметров $\alpha_1,\ldots,\alpha_q$. Если оценки параметра α уже найдены, то подставив их в величины $\left[x_k\right]$, из (125) получаем оценку для σ_X^2

$$\tilde{\sigma}_X^2=\frac{1}{T}\sum_{k=1}^{q+T}\left[x_k\right]^2. \qquad (127)$$

Сами же оценки параметров $\alpha_1,\ldots,\alpha_q$ могут быть найдены как решения системы (126), в которой вместо σ_X^2 следует подставить выражение, найденное из (125). Однако именно решение сисемы (126) представляет наибольшие трудности в силу ее нелинейности относительно α. В отличие от с.п.АР(р) здесь не представляется возможным переход к линейной системе «квазимаксимального правдоподобия». Рассмотрим ситуацию с оцениванием α в простейшем случае $q = 1$.

Модель СС(1) определяется скалярным параметром $\alpha_1 = \alpha$, который и подлежит оценке. Матрица A в этом случае представляет собой вектор-столбец с элементами $1, -\alpha, \alpha^2, \ldots, (-\alpha)^T$. Тогда $A'A = (A, A) = \left[1 - \alpha^{2T+2}\right] / \left(1 - \alpha^2\right)$ и из (125) и (126) получаем уравнение

$$\frac{\alpha}{T}\left[\frac{(T+1)\alpha^{2T+1}}{1-\alpha^{2T+2}} - \frac{1}{1-\alpha^2}\right]\sum_{k=1}^{T+1}\left[x_k\right]^2 -$$

$$-\sum_{k=1}^{T+1}\left[x_k\right]\cdot\partial\left[x_k\right]\Big/\partial\alpha = 0 \qquad (128)$$

Выразим величины $\left[x_k\right]$, $k = \overline{1, T+1}$ через α и наблюдения y_k, $k = \overline{2, T+1}$. Оценку предварительного значения «шума» $\tilde{x}_1 = \left[x_1\right]$ можно найти из равенства (119). Легко убедиться, что в нашем примере матрица B имеет вид

$$B = \begin{pmatrix} 0 & 0 & & \cdots 0 \\ 1 & 0 & 0 & \cdots 0 \\ -\alpha & 1 & 0 & \cdots 0 \\ \alpha^2 & -\alpha & 1 & \cdots 0 \\ -\alpha^3 & \alpha^2 & -\alpha & \cdots 0 \\ (-\alpha)^{T-1} & (-\alpha)^{T-2} & & 1 \end{pmatrix},$$

а уравнение (119) можно записать в виде

$$x_1\sum_{k=0}^{T}\alpha^{2k} = \sum_{k=2}^{T+1}\gamma_k y_k$$

где $\gamma_k = (-1)^k \alpha^{k-1}\left(1 + \alpha^2 + \alpha^4 + \ldots, \alpha^{2(T-k-1)}\right)$, $k = \overline{2, T+1}$.

Тогда

$$\left[x_1\right] = \tilde{x}_1 = \frac{1-\alpha^2}{1-\alpha^{2(T+1)}}\cdot\sum_{k=2}^{T+1}\gamma_k y_k .$$

Величины $\left[x_2\right], \left[x_3\right], \ldots, \left[x_{T-1}\right]$ можно теперь определить из реккурентных уравнений

$$\left[x_k\right] = y_k - \alpha\left[x_{k-1}\right],\ k = \overline{2, T+1}.$$

Решение этих уравнений приводит к формулам

$$[x_k] = y_k - \alpha y_{k-1} + \ldots + (-1)^{k-1}\alpha^{k-1}[x_1], \quad k = \overline{2, T+1}.$$

После подстановки этих значений в (128) получим уравнение, содержащее кроме α лишь наблюдения $y_2, \ldots, y_{T+1}$. Однако точное решение полученного уравнения даже в этом простейшем частном случае мы найти не можем. Поэтому рассмотрим один из подходов к получению приближенных оценок м.п. вектора $\alpha = (\alpha_1, \ldots, \alpha_q)$.

Линеаризация модели. Рассмотрим в качестве «главной части» логарифма функции правдоподобия величину

$$S(\alpha) = \sum_{k=1}^{q+T}[x_k]^2 = \left([x^{q+T}], [x^{q+T}]\right)$$

и будем ее минимизировать. Представляя $[x_k]$ по формуле Тейлора в окрестности точки, соответствующей предполагаемому значению параметра $\alpha^0 = (\alpha_1^0, \ldots, \alpha_q^0)$ получим приближенной равенство

$$[x_k] = [x_{0,k}] - \sum_{j=1}^{q}(\alpha_j - \alpha_j^0)p_{j,k}, \qquad (129)$$

где $[x_{0,k}] = [x_k | y^T, \alpha^0]$, $p_{j,k} = -\left.\dfrac{\partial [x_k]}{\partial \alpha_j}\right|_{\alpha=\alpha^0}$.

В качестве α^0 можно использовать оценку вектора α, предварительно полученную по методу моментов с помощью линейно сходящегося итеративного процесса (см. разд. 3.3.). Если P - матрица с элементами $p_{j,k}$ размера $(q+T)\times q$, то $q+T$ уравнений (129) можно представить в виде

$$[x_0^{q+T}] = P(\alpha - \alpha_0) + [x^{q+T}],$$

где $[x_0^{q+T}]$ и $[x^{q+T}]$ - векторы-столбцы с $q+T$ элементами. Поправка $(\alpha - \alpha_0)$, минимизирующая $S(\alpha)$, находится теперь линейным методом наименьших квадратов. Полагая $\partial S(\alpha)/\partial\alpha = (\partial S(\alpha)/\partial\alpha_1, \ldots, \partial S(\alpha)/\partial\alpha_q)$, получаем

$$\partial S(\alpha)/\partial\alpha = -P'[x_0^{q+T}] + 2P'P(\alpha - \alpha^0).$$

Приравнивая этот вектор к нулевому находим

$$(\alpha - \alpha^0) = (P'P)^{-1}P'[x_0^{q+T}]. \qquad (130)$$

Поскольку $[x_k]$ не является линейной функцией параметров α, однократная поправка не обеспечит получение оценок наименьших квадратов. Поэтому полученные после первой поправки значения используются как новые предполагаемые значения, и процедура повторяется до получения сходящихся результатов. Сходимость оказывается более быстрой, если используются достаточно удачные начальные приближения. Если начальные приближения неудачны, то процесс вычисления параметров может оказаться расходящимся.

М.м.п. для смешанной модели АРСС(p,q). Предположим теперь, что выходной сигнал $Y = \left\{Y_t, t = 0,\pm1,\ldots\right\}$ порождается моделью (43), на входе которой имеем (как и в предыдущих случаях) гауссовский «шум» $X = \left\{X_t, t = 0,\pm1,\ldots\right\}$ с $MX_t = 0$ и $DX_t = \sigma_X^2$. Считаем, что параметры авторегрессии $\beta = \left(\beta_1,\ldots,\beta_p\right)$ обеспечивают стационарность выходного сигнала Y. Тогда Y - гауссовская стационарная с.п. с $MY_t = 0$ и корреляционной функцией, определяемой уравнениями (46). Как в случае чистой с.п.СС(q), уравнения м.п. сложны и не поддаются непосредственному решению. Поэтому на практике можно использовать возможность представления с.п.АРСС(p,q) в виде чистой последовательности СС(∞) [2,с.264], которую в свою очередь заменяют конечной с.п.СС(Q):

$$\sum_{j=0}^{Q}\delta_j x_{t-j},$$

где Q - невелико, а δ_j- параметры, являющиеся функциями параметров $\theta = \left(\alpha_1,..,\alpha_q,\beta_1,..,\beta_p\right)$ исходной последовательности Y.

Для оценки параметра θ в модели (131) теперь можно использовать, например, описанную выше «линеаризацию модели». Иногда для получения оценок, близких к оценкам м.п., можно следовать более прямым путем. Расмотрим пример.

Модель АРСС(2,1) описывается разностным уравнением

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \alpha_1 X_{t-1} \tag{132}$$

в котором оценке подлежит параметр $\theta = \left(\alpha_1,\beta_1,\beta_2\right)$. Полагая $x_1 = x_2 = 0$ из (132) получаем

$$x_3 = y_3 - \beta_1 y_1 - \beta_2 y_2,$$

$$x_4 = y_4 - \beta_1 y_3 - \beta_2 y_2 - \alpha_1 x_3,$$

и т.д. Оценка вектора θ может быть получена по наблюдениям $y^T = (y_1, \ldots, y_T)$ путем минимизации величины

$$S(\theta) = \sum_{t=3}^{T} x_t^2.$$

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Журбенко И.Г. Кожевникова И.А.* Стохастическое моделирование процессов. М.:Изд-во МГУ, 1990 148с.
2. *Андерсон Т.* Стохастический анализ временных рядов. М.:Мир, 1976 757с.
3. *Бокс Дж. Дженникс Г.* Анализ временных рядов. Прогноз и управление. М.:Мир, 1974 Вып.1 406с.
4. *Дианова Р.С.* Моделирование случайных процессов и оценка их параметров. Учеб.пособие. Моск. Авиац. Ин-т. М., 1989 44с.
5. *Харин Ю.С. Степанова М.Д.* Практикум на ЭВМ по математической статистике. Минск: Изд-во Университетское, 304с.
6. *Ивченко Г.И. Медведев Ю.И.* Математическая статистика. М.: Высш.шк. 1984, 248с.
7. *Бендат Дж. Присон А.* Прикладной анализ случайных данных. М.:Мир 1989 541с.

П р и л о ж е н и е 1

ПРОВЕРКА ГИПОТЕЗЫ СЛУЧАЙНОСТИ

Входную с.п. $X = \left(X_1, X_2, \ldots\right)$ часто предролагают «чисто случайной», т.е считают с.в. X_k независимыми и одинаково распределенными. Это предположение, как правило, оправдано и вытекает из самого характера задачи, но иногда оно нуждается в проверке. Математически задачу можно сформулировать следующим образом: проверить гипотезу $H_0 : F(x) = F(x_1) \ldots F(x_T), x = \left(x_1, \ldots, x_T\right)$, где $F(x)$ - некоторая функция распределения. Такую гипотезу назавают *гипотезой случайности*. Критерий согласия для проверки этой гипотезы можно построить [6] исходя из следующих соображений (далее предполагается, что вектор $X = \left(X_1, \ldots, X_T\right)$ имеет непрерывное распределение).

Если гипотеза случайности действительно имеет место, то компоненты вектора X «равноправны» и поэтому данные наблюдения не должны быть ни в каком смысле упорядоченны. Другими словами, ситуацию, соответствующую гипотезе H_0, можно охарактеризовать как «полный хаос» или «полный беспорядок». При отклонениях от H_0 исходные данные имеют тот или иной порядок. Следовательно, критерий проверки H_0 можно построить на основании статистик, измеряющих степень «беспорядка» исходных данных. Одной из таких статистик является число инверсий ввыборке. Эта статистика определяется следующим образом. Построим вариационный ряд $x_{(1)}, \ldots, x_{(T)}$ выборки $x = \left(x_1, \ldots, x_T\right)$. Говорят, что компоненты x_i и x_j образуют инверсию, если $i < j$, но x_i стоит правее x_j в вариационном ряду т.е. наблюдению с меньшим номером соответствуют большее значение.

Если для $i < j$ положить

$$h_{ij} = \begin{cases} 1, \text{если} \quad x_i > x_j, \\ 0, \text{если} \quad x_i \le x_j. \end{cases}$$

то число инверсий t_j, образованных компонентный x_j (в вариационном ряду левее x_i стоит t_i элементов выборки с большими номерами), определится равенством

$$t_j = \sum_{j=i+1}^{T} h_{ij}, i = 1,...,T-1.$$

Тогда общее число инверсий $u_T(x)$ для выборки $x = (x_1,\ldots,x_T)$ будет задано равенство

$$u_T = u_T(x) = t_1+\ldots+t_{T-1}.$$

Статистика u_T является естетственной мерой «беспорядка» среди наблюдений и ее можно использовать для проверки гипотезы H_0. Крайние случаи, когда вариационный ряд имеет вид $x_1 < x_2 <\ldots< x_T$ или $x_T < x_{T-1} <\ldots< x_1$ естетственно рассматривать как свидетельства «полного отсутствия беспорядка», т.е. противоречашее гипотезе H_0. В первом случае статистика u_T принимает минимальное значение, равное 0, а во втором случае она максимальна и равна $(T-1)+(T-2)+\ldots+1 = T(T-1)/2$. Таким образом, слишком малые и слишком большие значения u_T (близкие к $T(T-1)/2$) естетственно рассматривать как критические для гипотезы H_0. Чтобы определить числовые характеристики этого критерия, найдем распределение статистики U_T при гипотезе H_0.

Из соображений симметрии ясно, что при гипотезе H_0 любое из $T!$ относительных расположений элементов выборки в соответствующем вариационном ряду имеет одинаковую вероятность $1/T!$. Введенная с.в. V_i определяется расположением с.в. X_i по отношению к $X_{i+1},\ldots X_T$ в вариационном ряду и не зависит от относительного расположения последних между собой, т.е V_i при любом $i = 1,\ldots,T-2$ не зависит от $V_{i+1},\ldots,V_{T-1}$ взаимно независимы. При этом величина V_i может с одной и той же вероятностью $1/(T-i+1)$ принимать значения $0,1,\ldots T-i$, поэтому ее производящая функция имеет вид

$$\varphi_i(z) = \sum_{r=0}^{T-i} P\{V_i = r\} z^r = \frac{1}{T-i+1}\left(1+z+\ldots+z^{T-1}\right),$$

а производящая функция статистики U_T представлена следующим образом:

$$\Phi_T(z) = \sum_{r=0}^{T-i} P\{U_T = r\} z^r = \prod_{i=1}^{T-1} \varphi_i(z) = \frac{1}{T!}\prod_{i=1}^{T-1}\left(1+z+\ldots+z^r\right).$$

Отсюда имеем

$$MV_i = \varphi_i'(1) = \frac{T-i}{2},$$

$$DV_i = \varphi_i''(1) + MV_i - (MV_i)^2 = \frac{(T-i)(T-i+2)}{12},$$

$$MU_T = \sum_{i=1}^{T-1} MV_i = \frac{T(T-1)}{4},$$

$$DU_T = \sum_{i=1}^{T-1} DV_i = \frac{2T^3 + 3T^2 + 5T}{72}.$$

Итак, среднее значение статистики U_T при нулевой гипотезе совпадает с серединой промежутка $\left[0, T(T-1)/2\right]$, и в критическую область $I_{1\alpha}$ следует включать все целые точки этого промежетка, достаточно удаленные от середины. При заданном уровне значимости α в качестве области принятия нулевой гипотезы можно взять промежуток

$$\left[U_{T:1-\alpha/2}, U_{T:\alpha/2}\right] = I_{0\alpha},$$

где величины $U_{T:\alpha}$ определяются из условия

$$P\left\{U_T > U_{T:\alpha}\right\} = \alpha$$

Разлагая функцию $\Phi_T(z)$ в ряд по степеням z и вычисляя коэффициент при z^r, можно вычислить вероятности $P\left\{U_T = r | H_0\right\}$ при заданном T и любом r и использовать их для нахождения границы $U_{T:\alpha}$.

Приведем небольшую таблицу из работы [7], содержащую 100α - процентные точки функции распределения с.в. U_T.

Таблица

Процентные точки рапределения числа инверсий

T	α					
	0,99	0,975	,095	0,05	0,025	0,01
10	9	11	13	31	33	35
12	16	18	21	44	47	49
14	24	27	30	60	63	66
16	34	38	41	78	81	85
18	45	50	54	98	102	107
20	59	64	69	120	125	130
30	152	162	171	263	272	282
40	290	305	319	460	474	489
50	473	495	514	710	729	751
60	702	731	756	1013	1038	1067
70	977	1014	1045	1369	1400	1437
80	1299	1344	1382	1777	1815	1860
90	1668	1721	1766	2238	2283	2336
100	2083	2145	2198	2751	2804	2866

П р и м е ч а н и е . Значения $U_{T:\alpha}$ такие, что $P\{U_T > U_{T:\alpha}\} = \alpha$, где T - общее число наблюдений.

Пусть имеется следующая последовательность $T = 20$ наблюдений некоторой с.в.

1. 5,5	6. 5,7	11. 6,8	16. 5,4
2. 5,1	7. 5,0	12. 6,6	17. 6,8
3. 5,7	8. 6,5	13. 4,9	18. 5,8
4. 5,2	9. 5,4	14. 5,4	19. 6,9
5. 4,8	10. 5,8	15. 5,9	20. 5,5

Гипотеза заключается в том, что эти наблюдения представляют собой независимые исходы (измерения) с.в. X, и ее надо проверить при заданном уровне значимости $\alpha = 0{,}05$. Подсчитаем число инверсий в последовательности

$$
\begin{array}{llll}
t_1 = 8 & t_6 = 6 & t_{11} = 7 & t_{16} = 0 \\
t_2 = 3 & t_7 = 1 & t_{12} = 6 & t_{17} = 2 \\
t_3 = 8 & t_8 = 8 & t_{13} = 0 & t_{18} = 1 \\
t_4 = 3 & t_9 = 1 & t_{14} = 0 & t_{19} = 1 \\
t_5 = 0 & t_{10} = 4 & t_{15} = 3 &
\end{array}
$$

Общее число инверсий $u_T = 62$. Область принятия имеет вид

$$\{u_{20:0.975} < u_{20} \le u_{20:0.025}\}.$$

Из таблицы при $\alpha = 0{,}05$ находим $u_{20:0.0975} = 64$ и $u_{20:0.0025} = 125$. Следовательно, гипотеза должна быть отвергнута с уровнем значимости 5%, т.к. $u_{20} = 62$ не попадает в промежуток между 64 и 125ю

Для большой «длины» релизации T применяют простой асимптотический вариант этого критерия. Используя производящую функцию $\Phi_T(z)$, можно показать, чтохарактеристическая функция нормированной статистики $U_T^* = \left(U_T - T(T-1)/4\right)\left(6/n^{3/2}\right)$ сходится при $T \to \infty$ и любом конечном t к $e^{-t^2/2}$ - характеристической функции нориального распределения. Это означает, что распределение величины U_T^* слабо сходится к нормальному закону $N(0,1)$ при $T \to \infty$.

Последний результат дает возможность сформулировать следующее правило проверки H_0, когда значение T велико: для заданного уровня значимости α определяют число t_α из условия $\Phi(-t_\alpha) = \alpha/2$; по фактически наблюдающимся данным $x = (x_1, \ldots, x_T)$ вычисляют значение $t = U_T(x)$ числа инверсий в выборке. Если $\left|t - T(T-1)/4\right| 6/T^{3/2} > t_\alpha$, то гипотезу H_0 отвергают ка противоречащую исходным данным. В противном случае признают, что гипотеза независимости и одинаковой распределенности наблюдений согласуется с опытными данными.

Вероятность ошибочно отвергнуть при этом истинную гипотезу H_0

$$P\left\{\left|U_t - \frac{T(T-1)}{4}\right|\frac{6}{T^{3/2}} > t_\alpha | H_0\right\} \xrightarrow[T\to\infty]{} 2\Phi(-t_\alpha) = \alpha .$$

Это правило можно использовать уже при $T > 10$.

Приложение 2

ПРОВЕРКА СТАЦИОНАРНОСТИ

Изложенные в основном тексте настоящего пособия методы статистической обработки временных рядов подразумевают стационарность последних. Стацинарность процесса играет важную роль потому, что методы анализа нестационарных процессов существенно более громозки, чем в стационарном случае. Часто стационарность вытекает из физической природы процесса, когда порождающие его факторы не зависят от времени. Однако далеко не всегда имеется априорная информация, позволяющая считать исследуемый процесс стационарным. Даже если временной ряд получается путем статического моделирования на компьютере, то не всегда удается обеспечить его стационарность с самого начала процесса моделирования. В частности, при моделировании с.п.АРСС с негауссовским шумом начальный «отрезок» выходного процесса может быть далек от стационарности, и возникает необходимость выбора для статистической обработки участка реализации, соответствующего установившемуся (стационарному) режиму. В Любом случае когда нужно убедится в стационарности наблюдений, поступивших для обработки, приходится использовать один из статистических критериев для проверки гипотезы стационарности.

Стационарность в узком смысле означает, как известно, независимость конечномерных распределений от сдвига по временной оси. Это значит, что функция распределния $F_{t_1,\dots,t_n}(x_1,\dots,x_n)$ процесса X_t удовлетворяет условию

$$F_{t_1+\tau,\dots,t_n+\tau}(x_1,\dots,x_n) = F_{t_1,\dots,t_n}(x_1,\dots,x_n), \quad n = 1,2,\dots$$

для любых значений аргументов и таких любых t_k и τ, что и $t_k + \tau, k = \overline{1,n}$ принадлежат допустимому множеству временного параметра t. Очевидно, что в этом случае для любой измеримрй функции $g(x_1,..,x_n)$ случайные величины

$$Y = g\left(X_{t_1},..,X_{t_n}\right) \text{ и } Y_\tau = g\left(X_{t_1+\tau},..,X_{t_n+\tau}\right)$$

одинаково распределены, а моментные функции, если они существуют, не зависят от сдвига по времени.

Если исследователя интересует не сами конечномерные распределения процесса X_t, а лишь его моментные функции, то целесообразно пользоваться определениями стационарности, основанными на свойствах моментных функций.

Стационарным в широком смысле k-го порядка называют процесс X_t, у которого моментные функции до k-го порядка включительно существуют и не зависят от сдвига по времени. При этом чаще всего речь идет о процессах, стационарных в широком смысле 2-го порядка, которые называются просто стационарными в широком смысле. Для них $MX_t = m_t = \text{const}$ и $R(t,s) = M(X_t - m)(X_s - m) = R_1(t-s)$.

Отметим, что если процесс стационарен в узком смысле и обладает моментными функциями до k-го порядка, то он является стационарным и в широком смысле k-го порядка. В случае гауссовских процессов стационарность в широком смысле 2-го порядка будет означать и стационарность в узком смысле, поскольку все конечномерные распределения гауссовского процесса вполне определяются двумя первыми моментными функциями: средним значением и корреляционной функцией.

Перейдем к рассмотрению задачи проверки гипотезы стационарности по существу. Мы используем знакомый из предшествующего раздела критерий инверсий. К сожалению, при построении такого критерия не удается обойтись только одной реализацией. Итак, предположим, что исследуется случайный процесс с дискретным временем, т.е. случайная последовательность $Y = \{Y_t, t = 0, \pm 1, \ldots\}$, относительно которой нужно проверить гипотезу H_0: «с.п. Y стационарна в узком смысле». Рассмотрим T независимых наблюдений последовательности Y, причем k-е наблюдение имеет длину nk, где n- некоторое положительное число. Тогда первое наблюдение представляет собой реализацию с.п. Y длины n

$$y_{1,1}, \ldots, y_{1,n};$$

k-ое наблюдение - реализацию длины kn

$$y_{k,1}, \ldots, y_{k,nk};$$

а последнее наблюдение (с номером T) - это реализация длины nT

$$y_{T,1}, \ldots, y_{T,nT}.$$

С помощью некоторой измеримой функции n переменных $g(x_1, .., x_n)$ построим используя n последниу значения каждой из T полученных реализаций, величины:

$$z_1 = g(y_{1,1}, \ldots, y_{1,n}), \ldots, z_T = g(y_{T,n(T-1)+1}, \ldots, y_{T,nT}).$$

Если в них значения y_{kr} заменить случайными случайными величинами Y_{kr} из последовательности $Y_k = \{Y_{kt}, t = 0, \pm 1, \ldots\}$, являющейся k-ой (по порядку) «копией» исследуемой с.п. Y, то получим случайные величины

$$Z_k = g\left(Y_{k,n(k-1)+1}, \ldots, Y_{k,nk}\right), k = \overline{1,T}.$$

Они будут независимы в совокупности, т.к. порождаются независимыми «копиями» с.п. Y. В случае же справедливости гипотезы H_0, т.е. в случае стационарности в узком смысле с.п. Y, они будут одинаково распределенными. Таким образом, гипотеза H_0 будет подтверждатся или отвергатся одновременно с принятием или отклонением гипотезы случайности для с.в. $Z_1, \ldots, Z_T$ при любой генерирующей их функции $g(x_1, .., x_n)$. Для проверки же гипотезы случайности можно использовать изложенный в прил.1 критерий инверсий.

При описанном выше подходе к проверке гипотезы стационарности функцию $g(x_1, .., x_n)$ цедесообразно выбирать так, чтобы величины z_k представляли собой выборочные моменты. Если существует, в частности, моментная функция r-го порядка и гипотеза H_0 верна, то выбирая функцию

$$g(x_1, .., x_n) = \frac{1}{n} \cdot \sum_{j=1}^{n} x_j^r,$$

в качестве z_k будет иметь оценку начального момента r-го порядка

$$z_k = \frac{1}{n} \cdot \sum_{j=(k-1)n+1}^{kn} y_j^r.$$

При достаточно слабых ограничениях эта оценка состоятельна:

$$Z_k \xrightarrow{p} \alpha_r = MY_t^r, k = \overline{1,T} \text{ при } n \to \infty.$$

Поэтому при истинности гипотезы H_0 величины z_k не должны существенно зависеть от времени (от индекса k). Если же они сильно колеблются в зависимости от k, то это свидетельствует о том, что наблюдения не согласуются гипотезой стационарности и можно не пренебрегать к использованию критерия инверсий. В случае отсутствия явный колебаний в совокупности $\{z_1, .., z_T\}$ придется пользоваться критерием инверсий.

На практике для проверки гипотезы H_0 можно рекомендовать построение критерия инверсий на основе совокупности оценок момента второго порядка

$$z_k = \frac{1}{n} \sum_{j=(k-1)n+1}^{kn} y_i^2, k = \overline{1,T}.$$