

Конспект курса проф. А.С.Шведова по временным рядам

Случайные процессы

Задача состоит в разработке разумно простых моделей, пригодных для прогнозирования, интерпретации и проверки гипотез, связанных с финансово-экономическими временными рядами. На финансовых рынках эти же или близкие модели используются для оценки финансовых инструментов и для выработки стратегий хеджирования (но сами по себе эти вопросы не относятся к курсу по временным рядам). Основной существующий подход заключается в том, что в качестве модели берется случайный процесс. Если модель выбрана правильно, то временной ряд может рассматриваться, как реализация данного случайного процесса.

Пусть Ω - пространство элементарных событий и $T \subset \mathbf{R}$. Пусть при каждом $t \in T$ задана случайная величина (с.в.)

$$X_t : \Omega \rightarrow \mathbf{R}$$

Такой набор случайных величин будем называть *случайным процессом*. Обычно в качестве множества T рассматривается либо

$$T = \{0, 1, 2, 3, \dots\},$$

в этом случае случайный процесс называют процессом с *дискретным временем*, либо

$$T = [0, \infty),$$

тогда говорят о случайном процессе с *непрерывным временем*. Могут рассматриваться и такие случайные процессы, где

$$X_t : \Omega \rightarrow \mathbf{R}^m.$$

Когда это не может вызвать недоразумений, прилагательное «случайный» опускают и случайный процесс называют просто процессом.

Предположим, что некоторый элемент $\omega \in \Omega$ зафиксирован. Тогда можно говорить о наборе чисел $X_t(\omega)$ для всевозможных $t \in T$. Этот набор чисел называется *реализацией* или *траекторией* случайного процесса.

Примеры случайных процессов

Пример Пусть начальный капитал игрока в момент времени 0 составляет N_0 рублей. В каждый из моментов времени 1, 2, ... бросается монета и игрок выигрывает 1 рубль при выпадении герба и проигрывает 1 рубль при выпадении решетки. Пусть N_t - это капитал игрока в момент времени t . Тогда N_t - это случайная величина при каждом целом $t > 0$. Набор этих случайных величин образует случайный процесс

$$N_{t+1} = N_t + \varepsilon_{t+1}$$

$\varepsilon_t = 1$ с вероятностью p и $\varepsilon_t = -1$ с вероятностью $(1-p)$, и с.в. ε_t и ε_s независимы при $t \neq s$.

Пример *Процесс белого шума* Пусть $T = \{0, 1, 2, 3, \dots\}$. Последовательность случайных величин $\{\varepsilon_t\}$ называется процессом белого шума, если выполнены следующие условия. Для всех t

$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2,$$

где $\sigma > 0$ — некоторое число (одно и то же для всех t), а также для всех t и $s \neq 0$

$$E(\varepsilon_t \varepsilon_{t-s}) = 0.$$

Пример. Процесс броуновского движения. Пусть $T = [0, \infty)$. Этот процесс характеризуется следующими свойствами:

(а) Пусть $t_0 < t_1 < \dots < t_n$. Тогда приращения $X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ — независимые в совокупности с.в. (В этом случае говорят, что случайный процесс является процессом с независимыми приращениями).

(б) Распределение вероятностей с.в. $X_t - X_s$, если $t > s$, зависит только от $t - s$ (и не зависит, например, от s) и при любом $x \in R$

$$P(X_t - X_s \leq x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^x \exp\left[-\frac{u^2}{2(t-s)}\right] du,$$

т.е. случайная величина $X_t - X_s \sim N(0, \sqrt{t-s})$.

Рассмотрим процесс броуновского движения. Пусть $X_0 = 0$. Можно доказать, что если $0 < t_1 < t_2 < \dots < t_n < t$, то при любом $x \in R$ и при любых a_1, \dots, a_n из R

$$P(X_t \leq x \mid X_{t_1} = a_1, \dots, X_{t_n} = a_n) = \frac{1}{\sqrt{2\pi(t-t_n)}} \int_{-\infty}^{x-a_n} \exp\left[-\frac{u^2}{2(t-t_n)}\right] du$$

Процесс броуновского движения играет фундаментальную роль в теории и в приложениях. Иногда его называют также *винеровским процессом*.

Замечание Иногда определенный выше случайный процесс называют процессом стандартного броуновского движения, оставляя название «процессом броуновского движения» для случая, когда

$$X_t - X_s \sim N(0, b \sqrt{t-s}),$$

где $b > 0$ — некоторая константа, не зависящая от s и t .

Мартингалы

Случайный процесс X_t называется *мартингалом*, если для любых

$$t_1 < t_2 < \dots < t_n < t$$

справедливо условие $E(X_t \mid X_{t_1} = a_1, \dots, X_{t_n} = a_n) = a_n$.

В рассмотренных примерах процесс броуновского движения является мартингалом. В примере с бросанием монеты случайный процесс N_t будет мартингалом только при $p = 0.5$.

Марковские процессы

Случайный процесс X_t называется *марковским*, если для любых

$$t_1 < t_2 < \dots < t_n < t$$

и для любых $a < b$ справедливо условие

$$P(a < X_t \leq b \mid X_{t_1} = a_1, \dots, X_{t_n} = a_n) = P(a < X_t \leq b \mid X_{t_n} = a_n)$$

И процесс броуновского движения, и процесс, связанный с бросанием монеты, являются марковскими.

Стационарные процессы

Пусть T - одно из рассмотренных ранее множеств, т.е. либо $T = \{0, 1, 2, 3, \dots\}$, либо $T = [0, \infty)$.

Случайный процесс X_t называется *стационарным в узком смысле*, если при любом натуральном n , при всех t_1, t_2, \dots, t_n из T и при любом $h > 0$ таком, что все моменты времени $t_1 + h, t_2 + h, \dots, t_n + h$ принадлежат множеству T , совместное распределение набора с.в.

$$X_{t_1}, \dots, X_{t_n}$$

совпадает с совместным распределением набора с.в.

$$X_{t_1+h}, \dots, X_{t_n+h}.$$

В частности, рассматривая случай $n=1$, получаем, что если случайный процесс X_t является стационарным в узком смысле, то распределение вероятностей с.в. X_t одно и то же при всех t .

Случайный процесс X_t называется *стационарным в широком смысле* или *ковариационно стационарным*, если выполняются следующие условия

1. При любом $t \in T$ существуют конечные $E(X_t)$ и $D(X_t)$ и эти величины, математическое ожидание и дисперсия, не зависят от t : т.е. если $t \in T$ и $s \in T$, то

$$E(X_t) = E(X_s), \quad D(X_t) = D(X_s)$$

2. При любом $t \in T$ и при любом $h > 0$ таком, что $t-h \in T$ существует конечная $Cov(X_t, X_{t-h})$, причем эта величина не зависит от t т.е. если $t \in T$ и $s \in T$ и $h > 0$ таково, что $t+h \in T$ и $s+h \in T$, то

$$Cov(X_t, X_{t-h}) = Cov(X_s, X_{s-h}).$$

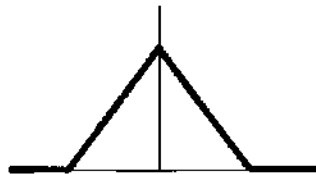
Зависеть от h ковариация может.

Случайный процесс стационарный в узком смысле будет стационарным и в широком смысле, если конечны все $E(X_t)$, $D(X_t)$, $Cov(X_t, X_{t-h})$.

Из стационарности случайного процесса в широком смысле не следует его стационарность в узком смысле. Чтобы показать это, рассмотрим следующий пример. Пусть $T = \{0, 1, 2, 3, \dots\}$; Случайные величины X_t и X_s независимы при $t \neq s$; при любом t

$$E(X_t) = 0, \quad D(X_t) = 1.$$

При этом при четных t с.в. X_t нормальны, а при нечетных t с.в. X_t имеют функцию плотности следующего вида



Стационарным в широком смысле этот случайный процесс является, а стационарным в узком смысле – нет.

Временные ряды

Пусть $T \subset \mathbf{R}$ и T состоит из конечного числа элементов. Элементы множества T будем называть моментами времени. Часто удобно считать, что $T = \{0, 1, 2, 3, \dots, T\}$. Пусть каждому $t \in T$ поставлено в соответствие число x_t . Такой набор чисел $\{x_t\}_{t=0}^T$ называется *временным рядом*.

Модель, нужная для прогнозирования, интерпретации и проверки гипотез, которая строится для данного временного ряда, - это случайный процесс X_t , где t принадлежит некоторому, возможно, большему множеству T . Если такая модель построена, то временной ряд рассматривается, как реализация (траектория) данного случайного процесса.

Стохастические разностные уравнения

Существует большое количество различных видов случайных процессов, некоторые из них будут рассмотрены ниже. Чтобы выбрать подходящий для моделирования данного временного ряда вид случайных процессов, а затем определить конкретные значения параметров, обычно временные ряды раскладываются на следующие компоненты: *тренд*, *сезонная*, *циклическая*, *иррегулярная*.

Пример: Пусть $y_t, t = 1, 2, \dots, 50$, - временной ряд, представленный в виде

$$y_t = T_t + S_t + I_t,$$

где

$$T_t = 1 + 0,1 t$$

$$S_t = 1,6 \sin(t \pi / 2)$$

$$I_t = 0,7 I_{t-1} + \varepsilon_t$$

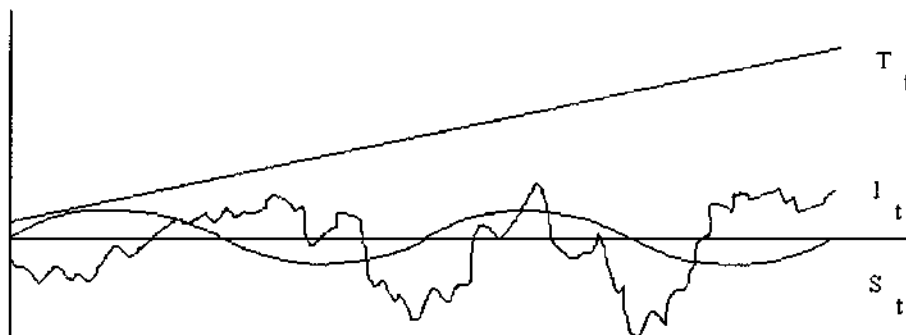
Здесь

T_t - величина тренд-компоненты в период t

S_t - величина сезонной компоненты в период t

I_t - величина иррегулярной компоненты в период t

ε_t - величина случайного возмущения в период t



Каждое из трех выписанных уравнений для T_t , S_t и I_t является *разностным уравнением*.

Примеры задач, где возникают разностные уравнения

Гипотеза о случайном блуждании для цены акции

Пусть изменение цены акции имеет среднее значение 0 и цена акции моделируется разностным уравнением

$$y_{t+1} = y_t + \varepsilon_{t+1}$$

y_t - цена акции в день t ;

ε_{t+1} - случайное возмущение.

Возможны две интерпретации этого уравнения:

либо ε_{t+1} - случайная величина, например $N(0, \sigma)$,

либо ε_{t+1} - значение этой с.в.

В первом случае y_{t+1} – это с.в., и тогда приведенное разностное уравнение называется стохастическим разностным уравнением. Во втором случае y_{t+1} – это число.

Другая запись этого же разностного уравнения

$$\Delta y_{t+1} = \varepsilon_{t+1}.$$

Можно рассмотреть более общее разностное уравнение

$$\Delta y_{t+1} = \alpha_0 + \alpha_1 y_t + \varepsilon_{t+1}.$$

Если гипотеза $\alpha_0 = \alpha_1 = 0$ должна быть отвергнута, то цена акции не является случайным блужданием.

Модель Самуэльсона. Выделение одного уравнения из системы разностных уравнений.

Рассмотрим стохастическую версию классической модели Самуэльсона (1939)

$$\begin{aligned} y_t &= c_t + i_t, \\ c_t &= \alpha y_{t-1} + \varepsilon_{ct}, \quad 0 < \alpha < 1 \\ i_t &= \beta (c_t - c_{t-1}) + \varepsilon_{it}, \quad \beta > 0 \end{aligned}$$

y_t - валовый национальный продукт

c_t - потребление

i_t - инвестиции, в период t .

ε_{ct} , ε_{it} - случайные величины с нулевым средним.

Параметры α и β следует определить.

Второе уравнение утверждает, что затраты на потребление пропорциональны доходам, полученным в предыдущий момент время плюс случайное возмущение. Третье уравнение: затраты на инвестиции пропорциональны увеличению потребления.

Подставляя второе уравнение в третье, получаем:

$$i_t = \beta (\alpha y_{t-1} + \varepsilon_{ct} - c_{t-1}) + \varepsilon_{it} = \alpha \beta y_{t-1} - \beta c_{t-1} + \beta \varepsilon_{ct} - \varepsilon_{it}$$

Но, $c_{t-1} = \alpha y_{t-2} + \varepsilon_{c(t-1)}$. Поэтому

$$i_t = \alpha \beta (y_{t-1} - y_{t-2}) + \beta (\varepsilon_{ct} - \varepsilon_{c(t-1)}) + \varepsilon_{it}$$

Имеем,

$$y_t = c_t + i_t = \alpha (1 + \beta) y_{t-1} - \alpha \beta y_{t-2} + (1 + \beta) \varepsilon_{ct} + \varepsilon_{it} - \beta \varepsilon_{c(t-1)}$$

Модели с одним переменным удобны при прогнозировании.

Форвардные и спот цены. Коррекция ошибки.

Пусть, цена единицы иностранной валюты на рынке спот (с немедленной поставкой) составляет S_t долларов.

Цена с поставкой через 1 период составляет f_t долларов. (В начале периода $t+1$ торговец получает иностранную валюту и платит f_t .) Эта валюта может быть продана за S_{t+1} . Прибыль (или потери) составляют

$$S_{t+1} - f_t$$

за единицу иностранной валюты.

Гипотеза о несмещенности форвардной цены утверждает, что ожидаемый выигрыш должен быть равен нулю:

$$S_{t+1} = f_t + \varepsilon_{t+1},$$

ε_{t+1} - случайное возмущение с нулевым средним.

Предположим, что собраны данные по двум ценам и построено уравнение регрессии

$$S_{t+1} = \alpha_0 + \alpha_1 f_t + \varepsilon_{t+1},$$

Гипотеза о несмещенности форвардной цены принимает вид $\alpha_0 = 0, \alpha_1 = 1$.

Для этих двух цен может быть рассмотрена модель с урегулированием:

$$S_{t+2} = S_{t+1} - \alpha (S_{t+1} - f_t) + \varepsilon_{s(t+2)}, \quad \alpha > 0$$

$$f_{t+1} = f_t + \beta (S_{t+1} - f_t) + \varepsilon_{f(t+1)}, \quad \beta > 0$$

$\varepsilon_{s(t+2)}$ и $\varepsilon_{f(t+1)}$ оба имеют нулевое среднее значение. Из этих уравнений следует. Что при $S_{t+1} - f_t > 0$ цена спот имеет тенденцию к уменьшению, а форвардная цена – тенденцию к увеличению.

Понятие решения разностного уравнения

Рассмотрим разностное уравнение (стохастическое разностное уравнение)

$$y_t = a_0 + \sum_{i=1}^n a_i y_{t-i} + x_t$$

Число n называется *порядком* разностного уравнения, x_t называется вынуждающим процессом. Две интерпретации: 1) x_t, y_t числа; 2) x_t, y_t случайные величины. Важный частный случай, когда $x_t = \sum_{i=0}^n \beta_i \varepsilon_{t-i}$, ε_t – процесс белого шума.

Решением разностного уравнения называется выражение y_t через

- 1) элементы последовательности x_t
- 2) t
- 3) возможно, начальные условия, например известное значение y_0 или y_0, y_1

Пример. Рассмотрим разностное уравнение

$$\Delta y_t = 2, \text{ т.е. } y_t = y_{t-1} + 2$$

Нетрудно проверить, что решением этого уравнения является

$$y_t = 2t + C,$$

где C - произвольная константа

Пример $I_t = 0.7 I_{t-1} + \varepsilon_t$

Проверим, что решением этого уравнения является

$$I_t = \sum_{i=0}^t 0.7^i \varepsilon_{t-i}$$

Действительно, $I_{t-1} = \sum_{i=0}^{t-1} 0.7^i \varepsilon_{t-1-i}$.

Тогда: $0.7 I_{t-1} + \varepsilon_t = \sum_{i=0}^{t-1} 0.7^{-i} \varepsilon_{t-1-i} + \varepsilon_t = \sum_{i=1}^t 0.7^i \varepsilon_{t-i} + \varepsilon_t$.

Замечание. Справедлива, например, запись:

$$I_t = 0.7(0.7 I_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = 0.49 I_{t-2} + 0.7 \varepsilon_{t-1} + \varepsilon_t.$$

Но эта запись не является решением разностного уравнения.

Построение решения разностного уравнения итерациями

Рассмотрим, например, уравнение

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t.$$

Пусть известно значение y_0 . Тогда:

$$y_1 = a_0 + a_1 y_0 + \varepsilon_1.$$

Далее,

$$y_2 = a_0 + a_1 y_1 + \varepsilon_2 = a_0[1 + a_1] + a_1^2 y_0 + a_1 \varepsilon_1 + \varepsilon_2;$$

$$y_3 = a_0 + a_1 y_2 + \varepsilon_3 = a_0[1 + a_1 + a_1^2] + a_1^3 y_0 + a_1^2 \varepsilon_1 + a_1 \varepsilon_2 + \varepsilon_3.$$

Нетрудно убедиться, что при $t > 0$:

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_1^i \varepsilon_{t-i}.$$

Аналогично, если известно не y_0 , а, например, y_{-m} , то

$$y_t = a_0 \sum_{i=0}^{t+m-1} a_1^i + a_1^{t+m} y_{-m} + \sum_{i=0}^{t+m-1} a_1^i \varepsilon_{t-i}.$$

Предположим временно, что $|a_1| < 1$. Тогда

$$a_1^{t+m} \rightarrow 0 \text{ при } m \rightarrow \infty.$$

$$\sum_{i=0}^{\infty} a_1^i = \frac{1}{1 - a_1}.$$

Поэтому решение можно приближенно записать в виде:

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}.$$

Нетрудно непосредственно проверить, что это действительно решение рассмотренного уравнения.

Но и при любом A величина, определяемая выражением

$$y_t = A a_1^t + \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i},$$

является решением того же уравнения.

При известном y_0 решение может быть получено в том же виде и при $|a_1| > 1$. Но при этом $|a_1|^t$ неограниченно возрастает при увеличении t .

Очень интересен случай $a_1 = 1$. Тогда решение принимает вид:

$$y_t = a_0 t + y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}.$$

Т.е. каждое из предыдущих возмущений имеет постоянный неубывающий эффект при всех t .

Альтернативный способ построения решения разностного уравнения

Описанный выше способ построения итерациями решения разностного уравнения первого порядка трудно применим к уравнениям более высокого порядка. Описываемый ниже способ применим к уравнениям любого порядка.

Общее решение разностного уравнения

$$y_t = a_0 + \sum_{i=1}^n a_i y_{t-i} + x_t,$$

которое мы будем называть *неоднородным*, поскольку в нем присутствует свободный член a_0 и вынуждающий процесс x_t , записывается как сумма *частного решения* неоднородного разностного уравнения и общего решения соответствующего *однородного уравнения*

$$y_t = \sum_{i=1}^n a_i y_{t-i}.$$

Методика решения неоднородного разностного уравнения порядка n состоит из следующих шагов:

Шаг 1. Построение однородного разностного уравнения и нахождение n его решений.

Шаг 2. Нахождение частного решения неоднородного уравнения.

Шаг 3. Построение общего решения неоднородного уравнения, как суммы частного решения неоднородного уравнения и линейной комбинации всех решений однородного уравнения.

Шаг 4. Исключение произвольных констант с помощью известных начальных условий.

Пример. Рассмотрим неоднородное разностное уравнение

$$y_t = 0,9y_{t-1} - 0,2y_{t-2} + 3.$$

Соответствующее однородное уравнение:

$$y_t - 0,9y_{t-1} + 0,2y_{t-2} = 0$$

имеет решения

$$y_{1t} = (0,5)^t; y_{2t} = (0,4)^t.$$

(Пока не будем пояснять, как были «угаданы» эти решения). Частным решением неоднородного уравнения является

$$y_t = 10.$$

Таким образом, общее решение имеет вид

$$y_t = A_1(0,5)^t + A_2(0,4)^t + 10.$$

Пусть, например, начальные условия имеют вид:

$$y_0 = 13, y_1 = 11,3.$$

Тогда из системы уравнений

$$13 = A_1 + A_2 + 10$$

$$11,3 = A_1 0,5 + A_2 0,4 + 10$$

находим: $A_1 = 1, A_2 = 2$.

Решение однородных разностных уравнений

Решение однородного разностного уравнения

$$y_t - a_1 y_{t-1} - a_2 y_{t-2} = 0$$

будем искать в виде

$$y_t = A\alpha^t.$$

Тогда

$$0 = A\alpha^t - a_1 A\alpha^{t-1} - a_2 A\alpha^{t-2} = A\alpha^{t-2}(\alpha^2 - a_1\alpha - a_2)$$

Нужно, чтобы

$$\alpha^2 - a_1\alpha - a_2 = 0.$$

Это уравнение называется *характеристическим уравнением*. Корни этого квадратного уравнения имеют вид

$$\alpha_{1,2} = \frac{a_1 \pm \sqrt{a_1^2 + 4a_2}}{2}.$$

Случай 1. $a_1^2 + 4a_2 > 0$.

Корни α_1 и α_2 действительные и различные. Общее решение однородного разностного имеет вид

$$y_t = A_1\alpha_1^t + A_2\alpha_2^t.$$

Случай 2. $a_1^2 + 4a_2 = 0$.

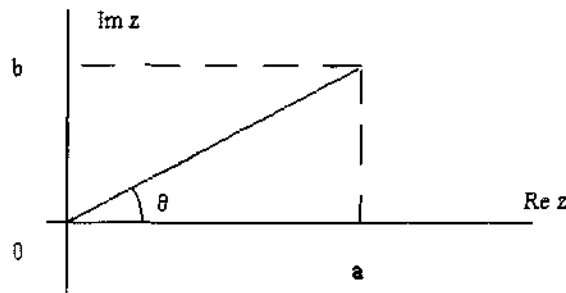
Корни α_1 и α_2 действительные и $\alpha_1 = \alpha_2$. Общее решение однородного разностного имеет вид

$$y_t = A_1\alpha_1^t + A_2 t\alpha_1^t$$

Случай 3. $a_1^2 + 4a_2 < 0$.

Корни α_1 и α_2 комплексно сопряженные

$$\alpha_{1,2} = a \pm ib, i^2 = -1$$



$$\cos\theta = \frac{a}{r}, \sin\theta = \frac{b}{r}, \text{ где } r = \sqrt{a^2 + b^2},$$

$$\alpha_1 = a + ib = r(\cos\theta + i\sin\theta)$$

$$\alpha_2 = a - ib = r(\cos\theta - i\sin\theta)$$

По формуле Муавра:

$$\alpha_1^t = r^t (\cos(t\theta) + i\sin(t\theta))$$

$$\alpha_2^t = r^t (\cos(t\theta) - i\sin(t\theta))$$

Чтобы получить общее действительное решение коэффициенты A_1 и A_2 в выражении

$$y_t = A_1\alpha_1^t + A_2\alpha_2^t$$

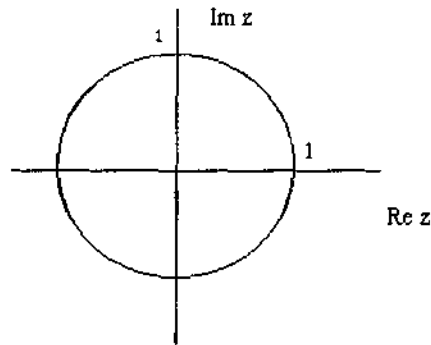
следует взять комплексно сопряженными:

$$A_{1,2} = B_1 (\cos B_2 \pm i\sin B_2)$$

Тогда нетрудно установить, что

$$y_t = 2B_1 r^t \cos(t\theta + B_2).$$

Во всех трех случаях, если корни α_1 и α_2 лежат внутри единичного круга, неограниченного роста решения при увеличении t не происходит.



Условие, что все корни характеристического уравнения лежат строго внутри единичного круга, называется *условием устойчивости*.

Аналогично ищутся решения однородного разностного уравнения n -го порядка

$$y_t - \sum_{i=1}^n a_i y_{t-i} = 0.$$

Условие устойчивости заключается в том, что корни характеристического уравнения

$$\alpha^n - a_1 \alpha^{n-1} - \dots - a_n = 0$$

лежат внутри единичного круга. Но тогда корни уравнения

$$1 - a_1 l - a_2 l^2 - \dots - a_n l^n = 0,$$

которое называется *обратным характеристическим уравнением*, — это величины, обратные корням характеристического уравнения. Поэтому условие устойчивости может быть сформулировано и так, что корни обратного характеристического уравнения лежат строго вне единичного круга.

Задача. Ваш коллега предложил следующую модель для темпа инфляции π_t в США:

$$\pi_t = -0,05 + 0,7\pi_{t-1} + 0,6\pi_{t-2} + \varepsilon_t,$$

ε_t — случайное возмущение с нулевым средним.

Пусть для периодов 0 и 1 темпы инфляции составляли 10% и 11% соответственно.

Найдите однородное, частное и общее решение данного уравнения.

Исходя из того, что в США не ожидается взрывного типа инфляции, объясните, почему вы считаете приведенную модель неудачной.

Операторы запаздывания (лага) и их использование при решении разностных уравнений

Рассмотрим оператор, т.е. функцию, переводящую последовательность случайных величин

$$\dots y_{-2}, y_{-1}, y_0, y_1, y_2, \dots$$

в некоторую другую последовательность случайных величин

$$\dots z_{-2}, z_{-1}, z_0, z_1, z_2, \dots$$

Мы будем широко использовать оператор, обозначаемый L , который определяется следующим способом:

$$L y_t = y_{t-1}$$

В данном случае $z_t = y_{t-1}$. Этот оператор называется *оператором запаздывания* или *оператором лага*. Последовательное применение n раз оператора L приводит к соотношению

$$L^n y_t = y_{t-n}$$

Если a – некоторое число, то через aL^n обозначается оператор, переводящий y_t в $a y_{t-n}$. Оператор aL^0 обозначается просто a . Определим также оператор, являющийся суммой операторов aL^n и bL^m :

$$(aL^n + bL^m) = a y_{t-n} + b y_{t-m}.$$

Отметим, что для константы C $LC=C$. (Т.е. если все y_t равны одной и той же константе C , то после применения оператора L будет получена та же самая последовательность констант.)

Оператор запаздывания в отрицательной степени является оператором опережения: $L^{-n} y_t = y_{t+n}$.

При $|a| < 1$ оператор $\frac{1}{1-aL}$ определяется следующим образом

$$\frac{1}{1-aL} = 1 + aL + a^2 L^2 + a^3 L^3 + \dots$$

Обоснование:

$$(1-aL)(1+aL+a^2 L^2+a^3 L^3+\dots) = 1-aL+aL-a^2 L^2+a^2 L^2-\dots = 1,$$

поскольку $a^n L^n y_t$ стремится к 0 при стремлении n к бесконечности.

При $|a| > 1$ оператор $\frac{1}{1-aL}$ определяется следующим образом:

$$\frac{1}{1-aL} = -(aL)^{-1} - (aL)^{-2} - \dots$$

Обоснование:

$$-(1-aL)((aL)^{-1} + (aL)^{-2} + \dots) = -(aL)^{-1} + 1 - (aL)^{-2} + (aL)^{-1} - (aL)^{-3} - (aL)^{-2} - \dots = 1,$$

поскольку $(aL)^{-n} y_t$ стремится к 0 при стремлении n к бесконечности.

Операторы запаздывания часто бывают удобны при работе с разностными уравнениями. Рассмотрим уравнение

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}$$

Оно может быть записано в виде:

$$A(L)y_t = a_0 + B(L)\varepsilon_t$$

где многочлены $A(L)$ и $B(L)$ имеют вид

$$A(L) = 1 - a_1 L - a_2 L^2 - \dots - a_p L^p, \quad B(L) = 1 + \beta_1 L + \dots + \beta_q L^q.$$

Пример. Рассмотрим стохастическое разностное уравнение

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1},$$

где ε_t – случайный процесс белого шума и $|a_1| < 1$. Запишем его в виде:

$$(1 - a_1 L)y_t = a_0 + (1 + \beta_1 L)\varepsilon_t$$

Или

$$y_t = [a_0 + (1 + \beta_1 L)\varepsilon_t] / (1 - a_1 L),$$

откуда

$$y_t = a_0/(1 - a_1) + \varepsilon_t/(1 - a_1L) + \beta_1 \varepsilon_{t-1}/(1 - a_1L).$$

Пояснение:

$$a_0/(1 - a_1L) = (1 + a_1L + a_1^2L^2 + \dots)a_0.$$

Но $L^i a_0 = a_0$ при любом i по свойству 1. Поэтому

$$a_0/(1 - a_1L) = (1 + a_1 + a_1^2 + \dots)a_0 = a_0/(1 - a_1).$$

Заметим, что при $\beta_1=0$ решение разностного уравнения было построено ранее итерациями, и оно совпадает с полученным решением.

Аналогично можно построить решение и для уравнения n -го порядка. Используя оператор запаздывания, уравнение

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_n y_{t-n} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}$$

можно записать в виде

$$(1 - a_1L - a_2L^2 - \dots - a_nL^n)y_t = a_0 + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}$$

или

$$y_t = \frac{a_0 + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}}{1 - a_1L - a_2L^2 - \dots - a_nL^n}$$

Напомним, условие устойчивости заключается в том, что корни характеристического уравнения

$$\alpha^n - a_1 \alpha^{n-1} - \dots - a_n = 0$$

лежат внутри единичного круга, или, что эквивалентно, все корни обратного характеристического уравнения

$$1 - a_1 l - a_2 l^2 - \dots - a_n l^n = 0$$

лежат вне единичного круга.

Рассмотрим, например, случай $n=2$ и пусть

$$1 - a_1L - a_2L^2 = (1 - b_1L)(1 - b_2L).$$

Условие, что корни обратного характеристического уравнения лежат строго вне единичного круга, в данном случае имеет вид

$$|b_1| < 1, \quad |b_2| < 1.$$

Тогда, например, для уравнения

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$

получаем

$$y_t = \frac{[(a_0 + \varepsilon_t)/(1 - b_1L)]}{1 - b_2L} = \frac{\frac{a_0}{1 - b_1} + \sum_{i=0}^{\infty} b_1^i \varepsilon_{t-i}}{1 - b_2L}.$$

Задание. Построить 25 случайных чисел с распределением $N(0, 1)$ и, считая $y_0 = 0$, построить последовательности

$$y_t = 0.9y_{t-1} + \varepsilon_t$$

$$y_t = 0.5y_{t-1} + \varepsilon_t$$

$$y_t = -0.5y_{t-1} + \varepsilon_t$$

$$y_t = y_{t-1} + \varepsilon_t$$

$$y_t = 1.2y_{t-1} + \varepsilon_t$$

$$y_t = -1.2y_{t-1} + \varepsilon_t$$

Нарисовать их графики.

Условия на коэффициенты характеристического уравнения, гарантирующие, что его корни лежат строго внутри единичного круга

Для линейного разностного уравнения:

$$y_t = a_1 y_{t-1} + \varepsilon_t$$

корни характеристического уравнения $\alpha - a_1 = 0$ лежат строго внутри единичного круга тогда и только тогда, когда $|a_1| < 1$.

Для линейного разностного уравнения

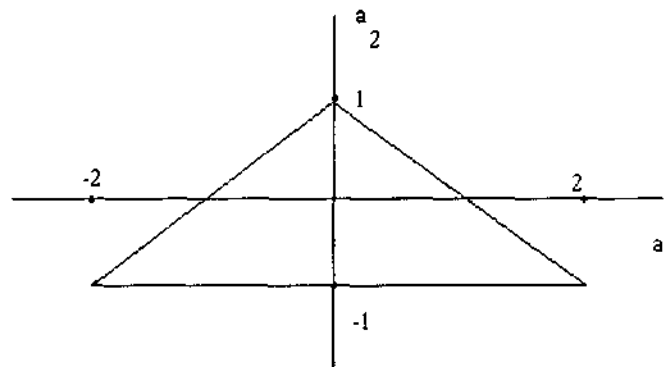
$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$

справедлива следующая теорема.

Все корни характеристического уравнения $\alpha^2 - a_1 \alpha - a_2 = 0$ лежат строго внутри единичного круга тогда и только тогда, когда выполняются условия:

$$a_1 + a_2 < 1, \quad -a_1 + a_2 < 1, \quad a_2 > -1.$$

Т.е. точка (a_1, a_2) должна лежать внутри треугольника:



Моделирование стационарных временных рядов

Экономические данные обычно собираются для дискретных моментов времени $t_0, t_0 + h, t_0 + 2h, t_0 + 3h, \dots$. Мы выбираем такой масштаб времени, что $h = 1$. Но t - это не обязательно физическое время. Например, при игре в рулетку t может обозначать номер испытания.

y_t , т.е. наблюдение, относящееся к моменту времени t , в одних случаях принимает значения из некоторого дискретного множества, в других случаях может принимать произвольные действительные значения.

Основной блок при строительстве дискретных стохастических моделей для временных рядов — это случайный процесс белого шума. Напомним, что последовательность случайных величин $\{\varepsilon_t\}$ называется процессом белого шума, если выполнены следующие условия. Для всех t математические ожидания

$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2,$$

где $\sigma > 0$ — некоторое число (одно и то же для всех t), а также для всех t и $s \neq 0$

$$E(\varepsilon_t \varepsilon_{t-s}) = 0.$$

Нетрудно увидеть, что для случайного процесса белого шума для всех t дисперсия

$$D(\varepsilon_t) = \sigma^2$$

и для всех t и $s \neq 0$ ковариация

$$\text{Cov}(\varepsilon_t, \varepsilon_{t-s}) = 0.$$

Если, кроме того, все случайные величины ε_t нормальные, то говорят, что процесс белого шума нормален.

Рассматривается методология Бокса-Дженкинса построения для временных рядов моделей вида

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i}, \quad \beta_0 = 1,$$

где ε_t - процесс белого шума. Данное уравнение определяет случайный процесс y_t . При соблюдении некоторых дополнительных условий (о которых говорится ниже) такие случайные процессы называют случайными процессами типа ARIMA (autoregressive integrated moving average). Однако не все случайные процессы типа ARIMA стационарны (под стационарностью мы в дальнейшем будем понимать стационарность в широком смысле).

Ближайшие цели:

1. Рассмотрение условий, при которых случайные процессы данного вида стационарны и свойств этих случайных процессов;
2. Развитие математического аппарата, используемого при работе с этими моделями. В частности, введение понятий автокорреляционной и частной автокорреляционной функций;
3. Рассмотрение различных статистик для проверки адекватности модели;

4. Демонстрация того, как построенная модель может быть использована для прогнозирования.

ARMA модели

Если все корни характеристического уравнения

$$\alpha^p - a_1 \alpha^{p-1} - a_2 \alpha^{p-2} - \dots - a_p = 0$$

лежат строго внутри единичного круга, процесс $\{y_t\}$, определяемый стохастическим разностным уравнением

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i}, \quad \beta_0 = 1,$$

называются процессом $ARMA(p, q)$. Условие, что все корни характеристического уравнения лежат строго внутри единичного круга мы называем *условием устойчивости*.

При $q=0$ процесс называется чисто авторегрессионным и обозначается $AR(p)$:

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t,$$

При $p=0$ процесс называется процессом скользящего среднего и обозначается $MA(q)$:

$$y_t = a_0 + \sum_{i=0}^q \beta_i \varepsilon_{t-i}, \quad \beta_0 = 1,$$

Решение разностного уравнения, дающее выражение для y_t через последовательность $\{\varepsilon_t\}$, называется представлением решения в форме скользящего среднего.

Рассмотрим процесс $MA(1)$ с $\beta_0 = 1$, $\beta_1 = 0,5$. Покажем, этот процесс не является процессом белого шума. Первые два условия из определения процесса белого шума выполнены

$$E(y_t) = E(\varepsilon_t + 0,5\varepsilon_{t-1}) = 0,$$

$$D(y_t) = D(\varepsilon_t + 0,5\varepsilon_{t-1}) = D(\varepsilon_t) + 0,25D(\varepsilon_{t-1}) = 1,25\sigma^2.$$

Но не выполнено третье условие

$$\begin{aligned} E(y_t y_{t-1}) &= E((\varepsilon_t + 0,5\varepsilon_{t-1})(\varepsilon_{t-1} + 0,5\varepsilon_{t-2})) = \\ &= E(\varepsilon_t \varepsilon_{t-1} + 0,5\varepsilon_{t-1}^2 + 0,5\varepsilon_t \varepsilon_{t-2} + 0,25\varepsilon_t \varepsilon_{t-2}) = E(0,5\varepsilon_{t-1}^2) = 0,5\sigma^2 \neq 0. \end{aligned}$$

Пример. При бросании доброкачественной монеты вы выигрываете 1 руб. при выпадении герба и проигрываете 1 руб. при выпадении решетки. Случайная величина ε_t — выигрыш при соответствующем броске. $\{\varepsilon_t\}$ — процесс белого шума.

Предположим, что вы ведете записи среднего выигрыша за 4 последних броска:

$$y_t = \frac{1}{4} (\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3}).$$

Это процесс скользящего среднего 3-го порядка.

Упражнение: Найти для этого примера ковариации

$$\text{Cov}(y_t, y_{t-s}).$$

Рассмотрим процесс $AR(1)$

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t, \quad |a_1| < 1.$$

Используя определение оператора $(1 - a_1 L)^{-1}$, нетрудно увидеть, что представление в форме скользящего среднего для этого процесса имеет вид

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}.$$

В данном случае видно, что полученный ряд сходится, т.к. коэффициенты a_1^i стремятся к 0 со скоростью геометрической прогрессии. Это выражение является представлением случайного процесса $AR(1)$ в форме скользящего среднего или в форме $MA(\infty)$.

В общем случае $ARMA(p, q)$ процесса частное решение разностного уравнения записывается в виде

$$y_t = \frac{a_0 + \sum_{i=0}^q \beta_i \varepsilon_{t-i}}{(1 - \sum_{i=1}^p a_i L^i)}.$$

Т.е. и процесс $ARMA(p, q)$ имеет представление в форме бесконечного скользящего среднего

$$y_t = \frac{a_0}{1 - \sum_{i=1}^p a_i} + \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i},$$

но то, что ряд

$$\sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}$$

сходится, необходимо обосновать.

Переход

$$\frac{a_0}{1 - \sum_{i=1}^p a_i} = \frac{a_0}{1 - \sum_{i=1}^p a_i L^i}$$

следует из того, что $L^i a_0 = a_0$ при любом i , и поэтому

$$(1 - \sum_{i=1}^p a_i L^i) a_0 = (1 - \sum_{i=1}^p a_i) a_0.$$

Мы покажем, что условие устойчивости — это необходимое условие для того, чтобы процесс $\{y_t\}$ был стационарным. Под стационарностью понимается стационарность в широком смысле, т.е. ковариационная стационарность.

Исследование на стационарность процесса $AR(1)$

Пусть

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t,$$

ε_t — белый шум.

Предположим, что процесс начинается в момент времени 0, начальное условие y_0 известно. Тогда, как было показано ранее,

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_1^i \varepsilon_{t-i} \quad (*)$$

Беря математическое ожидание, находим

$$E(y_t) = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0,$$

т.е. математическое ожидание зависит от t , и формально данный случайный процесс не является стационарным ни при каком a_1 . Нетрудно увидеть, что аналогичным образом зависят от t и ковариации $Cov(y_t, y_{t-s})$. Однако положение различное для случаев $|a_1| < 1$ и $|a_1| \geq 1$.

Если условие устойчивости разностного уравнения $|a_1| < 1$ выполнено, то из представления (*) видно, что при больших t можно приближенно считать, что

$$y_t = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \quad (**)$$

и рассматривать процесс y_t как стационарный. При нарушении условия устойчивости данный процесс никоим образом нельзя считать стационарным. Будем считать условие устойчивости выполненным и положим

$$\mu = \frac{a_0}{1-a_1}$$

Тогда из представления (**) видно, что

$$E(y_t) = \mu,$$

$$E((y_t - \mu)^2) = E((\varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + \dots)^2) = \sigma^2 (1 + a_1^2 + a_1^4 + \dots) = \frac{\sigma^2}{1-a_1^2}.$$

$$\begin{aligned} E((y_t - \mu)(y_{t-s} - \mu)) &= E((\varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2})(\varepsilon_{t-s} + a_1 \varepsilon_{t-s-1} + a_1^2 \varepsilon_{t-s-2} + \dots)) = \\ &= \sigma^2 a_1^s (1 + a_1^2 + a_1^4 + \dots) = \frac{\sigma^2 a_1^s}{1-a_1^2}. \end{aligned}$$

Контрольный вопрос: Является ли процесс $AR(1)$ процессом белого шума? Почему?

Полученный результат, что процесс $AR(1)$ является стационарным только при больших t имеет прямое отношение к реальной жизни. Методы, предназначенные для работы со стационарными временными рядами, желательно применять к тем рядам, которые начались "давно", даже если известно только относительно небольшое число последних членов этого ряда.

Формально математически случайный процесс $AR(1)$ является стационарным, только если он начался бесконечно давно. Однако обычно о случайных процессах $AR(1)$ говорят, как о стационарных, подразумевая, что они начались достаточно давно, и в этом случае по своим свойствам они близки к стационарным процессам.

Исследование на стационарность процесса $ARMA(2,1)$

Рассмотрим процесс

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t + \beta \varepsilon_{t-1}.$$

Мы рассматриваем случай $a_0 = 0$, чтобы сократить выражения, поскольку a_0 не входит в характеристическое уравнение, и величина a_0 не влияет на выполнение условия устойчивости.

Будем искать представление процесса $ARMA(p, q)$ в виде

$$y_t = \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}.$$

Чтобы этот ряд был решением рассматриваемого разностного уравнения должно выполняться соотношение

$$\begin{aligned} \delta_0 \varepsilon_t + \delta_1 \varepsilon_{t-1} + \delta_2 \varepsilon_{t-2} + \delta_3 \varepsilon_{t-3} + \dots &= a_1 (\delta_0 \varepsilon_{t-1} + \delta_1 \varepsilon_{t-2} + \delta_2 \varepsilon_{t-3} + \dots) + \\ &+ a_2 (\delta_0 \varepsilon_{t-2} + \delta_1 \varepsilon_{t-3} + \delta_2 \varepsilon_{t-4} + \dots) + \varepsilon_t + \beta \varepsilon_{t-1} \end{aligned}$$

Сравнение коэффициентов при $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ приводит к соотношениям

$$\begin{aligned} \delta_0 &= 1 \\ \delta_1 &= a_1 \delta_0 + \beta \\ \delta_2 &= a_1 \delta_1 + a_2 \delta_0 \end{aligned}$$

и вообще при

$$\begin{aligned} i &\geq 2 \\ \delta_i &= a_1 \delta_{i-1} + a_2 \delta_{i-2} \end{aligned}$$

Коэффициенты δ_i удовлетворяют однородному разностному уравнению второго порядка.

Поэтому ряд

$$\sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i} \quad (***)$$

является сходящимся, если корни характеристического уравнения

$$\alpha^2 - a_1 \alpha - a_2 = 0$$

лежат строго внутри единичного круга, так в этом случае коэффициенты ряда δ_i стремятся к нулю со скоростью геометрической прогрессии. Отсюда сле-

дует, в частности, что $\sum_{i=0}^{\infty} \delta_i^2 < \infty$. При нарушении условия устойчивости коэффициенты δ_i не стремятся к 0 при $i \rightarrow \infty$, и ряд (***) сходящимся не является.

Покажем, что процесс $MA(\infty)$: $y_t = \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}$ является стационарным, если $\sum_{i=0}^{\infty} \delta_i^2 < \infty$. Действительно,

$$\text{при любом } t: E(y_t) = 0;$$

$$\text{при любом } t: D(y_t) = E\left[\left(\delta_0 \varepsilon_t + \delta_1 \varepsilon_{t-1} + \delta_2 \varepsilon_{t-2} + \dots\right)^2\right] = \sigma^2 \sum_{i=0}^{\infty} \delta_i^2;$$

$$\text{при любых } t \text{ и } s: Cov(y_t, y_{t-s}) = \sigma^2 \sum_{i=0}^{\infty} \delta_i \delta_{i+s}.$$

Пример. Рассмотрим уравнение

$$y_t = 1.6y_{t-1} - 0.9y_{t-2} + \varepsilon_t + 0.5\varepsilon_{t-1}$$

Тогда

$$\delta_0 = 1$$

$$\delta_1 = 1.6 + 0.5 = 2.1$$

$$\delta_i = 1.6\delta_{i-1} - 0.9\delta_{i-2} \quad \text{при } i \geq 2.$$

Корни характеристического уравнения $\alpha^2 - 1.6\alpha + 0.9 = 0$ комплексно сопряженные и общее решение однородного уравнения имеет вид

$$\delta_i = 0.949' \cdot B_1 \cdot \cos(0.567 \cdot i + B_2)$$

Найденные значения для δ_0 и δ_1 дают

$$B_1 \cdot \cos B_2 = 1$$

$$0.949 \cdot B_1 \cdot \cos(0.567 + B_2) = 2.1$$

Отсюда

$$B_2 = -1.197, \quad B_1 = 2.739$$

и

$$\delta_i = 0.949' \cdot 2.739 \cdot \cos(0.567 \cdot i - 1.197).$$

Первые 10 коэффициентов имеют вид:

j	0	1	2	3	4	5	6	7	8	9	10
δ_j	1.00	2.10	2.46	2.05	1.06	-0.15	1.19	-1.79	-1.76	-1.23	-0.38

Исследование на стационарность процесса $ARMA(p, q)$

Рассмотрение общего случая аналогично исследованию на стационарность процессов $ARMA(2, 1)$. Рассмотрим сначала процесс $AR(p)$,

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t.$$

При соблюдении условия устойчивости, частное решение этого уравнения имеет вид:

$$y_t = \frac{a_0}{1 - \sum_{i=1}^p a_i} + \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i},$$

где неопределенные коэффициенты δ_i определяются как и при рассмотрении процессов $ARMA(2,1)$ и должны при $i \geq p$ удовлетворять однородному разностному уравнению:

$$\delta_t - a_1 \delta_{t-1} - a_2 \delta_{t-2} - \dots - a_p \delta_{t-p} = 0.$$

Если корни характеристического уравнения

$$\alpha^p - a_1 \alpha^{p-1} - \dots - a_p = 0$$

лежат внутри единичного круга, то $\sum_{i=0}^{\infty} \delta_i^2 < \infty$ и, соответственно, процесс y_t является стационарным.

В общем случае $ARMA(p,q)$ процессов

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + x_t,$$

$$x_t = \sum_{i=0}^q \beta_i \varepsilon_{t-i}$$

частное решение имеет вид:

$$y_t = \frac{a_0}{1 - \sum_{i=1}^p a_i} + \frac{\varepsilon_t}{1 - \sum_{i=1}^p a_i L^i} + \frac{\beta_1 \varepsilon_{t-1}}{1 - \sum_{i=1}^p a_i L^i} + \dots + \frac{\beta_q \varepsilon_{t-q}}{1 - \sum_{i=1}^p a_i L^i}.$$

Можно показать (это следует проделать в качестве упражнения), что при выполнении условия устойчивости процесс y_t является стационарным, если он начался бесконечно давно.

Автокорреляционные функции

Для ковариационно-стационарного случайного процесса y_t будем использовать обозначения:

$$\mu = E(y_t), \quad \sigma_y^2 = D(y_t), \quad \gamma_s = \text{Cov}(y_t, y_{t-s}).$$

Величину $\rho_s = \frac{\gamma_s}{\gamma_0}$ называют автокорреляцией между y_t и y_{t-s} . Заметим,

что $\gamma_{-s} = \gamma_s$.

Для процесса AR (1)

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t, \quad |a_1| < 1$$

нами установлено, что

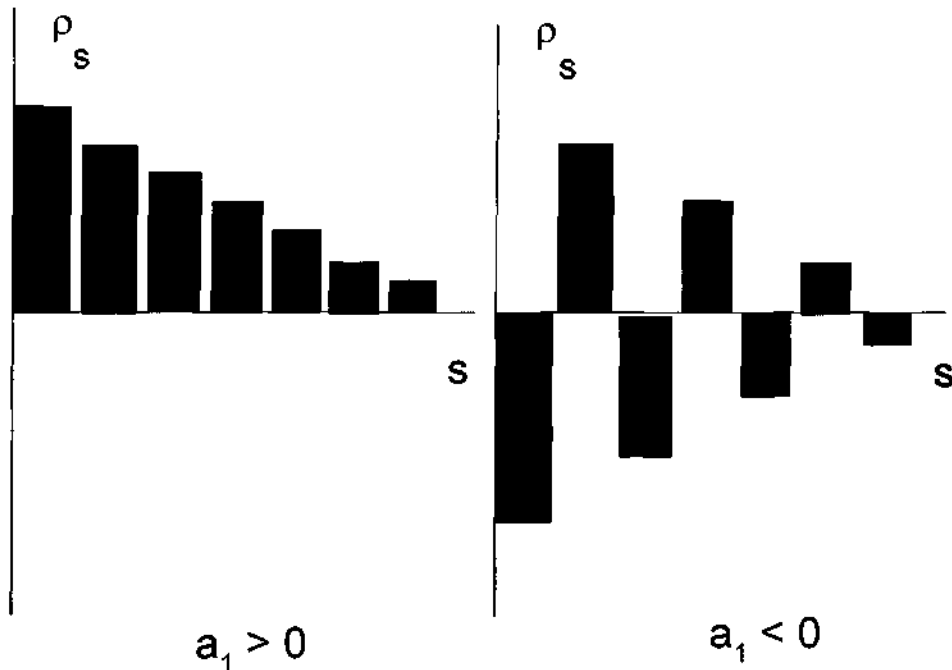
$$\gamma_s = a_1^s \gamma_0, \quad s = 0, 1, 2, \dots$$

Соответственно, автокорреляции

$$\rho_s = a_1^s, \quad s = 0, 1, 2, \dots$$

составляют геометрическую прогрессию.

Для процессов $AR(1)$ коэффициент a_0 не влияет на величины автокорреляций. Нетрудно увидеть, что то же самое верно и для любых процессов $ARMA(p, q)$.



Поэтому при рассмотрении задачи о нахождении автокорреляций для процессов $ARMA(p, q)$ можно ограничиться процессом

$$y_t = \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i}$$

с $a_0 = 0$.

Умножив приведенное уравнение на y_{t-s} и перейдя к математическим ожиданиям, получим

$$E(y_t y_{t-s}) = \sum_{i=1}^p a_i E(y_{t-i} y_{t-s}) + \sum_{i=0}^q \beta_i E(\varepsilon_{t-i} y_{t-s})$$

Данные уравнения, отвечающие различным s , называются *уравнениями Юла-Уокера*.

Заметим, в случае $a_0 = 0$ справедливо $E(y_t) = 0$ и поэтому $E(y_t y_{t-s}) = \gamma_s$.

Рассмотрим процесс $AR(p)$. В этом случае $q = 0$, поэтому из всех математических ожиданий $E(\varepsilon_{t-i} y_{t-s})$ следует рассмотреть только $E(\varepsilon_t y_{t-s})$. Но при $s \geq 1$ из определения белого шума следует, что $E(\varepsilon_t y_{t-s}) = 0$. Из уравнений Юла-Уокера получаем

$$\gamma_s = a_1 \gamma_{s-1} + a_2 \gamma_{s-2} + \dots + a_p \gamma_{s-p}$$

(с учетом того, что $\gamma_{-j} = \gamma_j$). Или, поделив на γ_0 ,

$$\rho_s = a_1 \rho_{s-1} + a_2 \rho_{s-2} + \dots + a_p \rho_{s-p}$$

Т.е. автокорреляции процесса $AR(p)$ доставляют решение однородному разностному уравнению p -го порядка.

Поэтому для процессов $AR(p)$ при соблюдении условия устойчивости $\rho_s \rightarrow 0$ при $s \rightarrow \infty$ примерно со скоростью геометрической прогрессии.

Для процессов $MA(q)$

$$y_t = \sum_{i=0}^q \beta_i \varepsilon_{t-i}$$

нетрудно увидеть, что при $s > q$

$$E(y_t y_{t-s}) = 0$$

т.е. $\rho_s = 0$ при $s > q$. Это является важной отличительной чертой процессов $MA(q)$.

Упражнение. Рассмотрим процесс $ARMA(1, 1)$

$$y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1}$$

Используя уравнения Юла-Уокера, показать, что

$$\rho_1 = \frac{(1 + a_1 \beta_1)(a_1 + \beta_1)}{1 + \beta_1^2 + 2a_1 \beta_1}$$

и $\rho_s = a_1 \rho_{s-1}$ при $s \geq 2$.

При соблюдении условия устойчивости процесс $ARMA(p, q)$ имеет представление $MA(\infty)$. Представление $MA(\infty)$ удобно для нахождения автокорреляционной функции.

Частные автокорреляционные функции

При использовании моделей $ARMA$ важное значение имеют величины $\Phi_{11}, \Phi_{22}, \dots, \Phi_{kk}, \dots$. Величина Φ_{kk} называется *частной автокорреляцией* и определяется как последняя компонента решения следующей системы линейных алгебраических уравнений.

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ & & \dots & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{pmatrix} \begin{pmatrix} \Phi_{k1} \\ \Phi_{k2} \\ \vdots \\ \Phi_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix}$$

Если рассматриваемый процесс является процессом $AR(p)$, то $\Phi_{kk} = 0$ при $k > p$. Действительно, для процессов $AR(p)$ из уравнений Юла-Уокера следует, что при всех $j \geq 1$

$$\rho_j = a_1 \rho_{j-1} + \dots + a_p \rho_{j-p}$$

(мы используем, что $\rho_{-1} = \rho$). Поэтому рассматриваемая система линейных алгебраических уравнений имеет решение

$$\Phi_{k1} = a_1, \dots, \Phi_{kp} = a_p, \Phi_{k,p+1}, \dots, \Phi_{kk} = 0.$$

Покажем, что определитель корреляционной матрицы отличен от 0. Из этого будет следовать, что данное решение единственно.

Действительно, пусть

$$\det \begin{pmatrix} 1 & \gamma_1 & \gamma_2 \dots \gamma_{k-1} \\ \gamma_1 & 1 & \gamma_1 \dots \gamma_{k-2} \\ & & \dots \\ \gamma_{k-1} & \gamma_{k-2} & \gamma_{k-3} \dots 1 \end{pmatrix} = 0.$$

На месте (i,j) этой матрицы стоит величина

$$C_{ij} = \text{Cov}(y_{t-i}, y_{t-j}),$$

и если определитель равен 0, то существует линейная зависимость между столбцами этой матрицы, т.е. существуют числа $\alpha_1, \dots, \alpha_k$ не все равные 0, такие, что при любом $j=1, \dots, k$

$$\sum_{i=1}^k \alpha_i C_{ij} = 0,$$

т.е.

$$\text{Cov}\left(\sum_{i=1}^k \alpha_i y_{t-i}, y_{t-j}\right) = 0.$$

Умножим j -е уравнение на α_j и сложим все уравнения. Тогда

$$\text{Var}\left(\sum_{i=1}^k \alpha_i y_{t-i}\right) = 0.$$

Вспомним, что

$$y_{t-i} = \sum_{j=0}^{\infty} \delta_j \varepsilon_{t-i-j}, \quad \delta_0 \neq 0.$$

Пусть i^* - минимальное из тех i , для которых $\alpha_i \neq 0$. Тогда

$$\sum_{i=1}^k \alpha_i y_{t-i} = \alpha_{i^*} \delta_0 \varepsilon_{t-i^*} + \sum_{i=i^*+1}^k \beta_i \varepsilon_{t-i}.$$

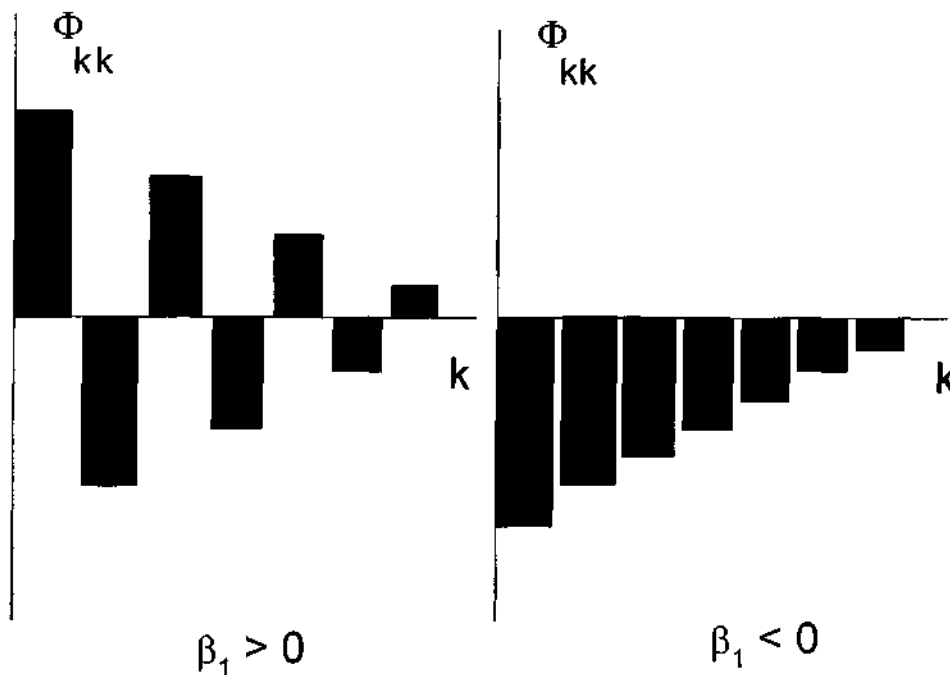
Поэтому

$$\text{Var}\left(\sum_{i=1}^k \alpha_i y_{t-i}\right) \geq (\alpha_{i^*} \delta_0)^2 > 0.$$

Полученное противоречие показывает, что определитель ковариационной (и, следовательно, корреляционной) матрицы не может быть равен 0.

Равенство нулю частных автокорреляций Φ_{kk} при $k > p$ является важной отличительной чертой процессов $AR(p)$.

Приведем частные коррелограммы для процессов $MA(1)$. Они напоминают коррелограммы процессов $AR(1)$.



Приведем без доказательства следующие результаты:

Автокорреляции процесса $ARMA(p, q)$ начинают убывать, начиная с запаздывания q .

Частные автокорреляции процесса $ARMA(p, q)$ начинают убывать, начиная с запаздывания p .

В обоих случаях убывание может быть как монотонным так и немонотонным.

Home task: Consider the stochastic difference equation

$$y_t = a_0 + a_2 y_{t-2} + \varepsilon_t$$

- A. Find the homogeneous solution and determine the stability condition
- B. Find the particular solution using the method of undetermined coefficients

Consider the second-order autoregressive process

$$y_t = a_0 + a_2 y_{t-2} + \varepsilon_t$$

- A. Find
 - (i) $Cov(y_t, y_{t-1})$
 - (ii) $Cov(y_t, y_{t-2})$
 - (iii) The partial autocorrelations Φ_{11}, Φ_{22}

Выборочные автокорреляционные функции и выборочные частные автокорреляционные функции

Пусть исходные данные – это набор чисел y_1, \dots, y_n . Положим

$$\bar{y} = \frac{1}{T} \sum_{i=1}^T y_i$$

и при $s=1, 2, \dots$

$$r_s = \frac{\sum_{t=s+1}^T (y_t - \bar{y})(y_{t-s} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$$

Числа r_1, r_2, \dots называются *выборочными автокорреляциями*. *Выборочные частные автокорреляции* $\Phi_{11}, \Phi_{22}, \dots$ (мы не будем вводить для них нового обозначения) строятся описанным ранее способом, но с использованием r_1, r_2, \dots вместо ρ_1, ρ_2, \dots .

Первые очевидные соображения при идентификации следующие.

Если при s большем некоторого q выборочные автокорреляции r_s становятся близкими к нулю, то подходящей моделью может быть $MA(q)$

Если при s большем некоторого p выборочные частные автокорреляции Φ_{ss} становятся близкими к нулю, то подходящей моделью может быть $AR(p)$

Если и r_s , и Φ_{ss} стремятся к нулю плавно, то подходящей является смешанная модель $ARMA(p, q)$.

Важным является также принцип экономности: предпочтите *меньше* модели с меньшими значениями p и q .

Для уточнения того, что понимать под словами “ r_s близки к нулю” или “ Φ_{ss} близки к нулю” могут быть использованы следующие результаты

Box and Jenkins (1976) показали, что если y_t - это процесс $MA(q)$ и процесс белого шума нормален, то при больших T для $s > q$ с.в. r_s распределена асимптотически нормально со средним равным 0 и приближенно

$$Var(r_s) = \frac{1}{T} \quad \text{при } s=1$$

$$Var(r_s) = \frac{1}{T} \left(1 + 2 \sum_{j=1}^{s-1} r_j^2 \right) \quad \text{при } s > 1$$

(каждой реализации случайного процесса при любом s соответствует свое значение r_s).

Если y_t - это процесс $AR(p)$, то аналогичное утверждение справедливо для частных автокорреляций Φ_{ss} при $s > p$ и дисперсия Φ_{ss} приблизительно равна

$$Var(\Phi_{ss}) = \frac{1}{T}.$$

Процедура Бокса - Дженкинса построения модели ARMA для временного ряда

Данная процедура состоит из 3 этапов.

Этап 1. Идентификация модели: определение чисел p и q .

Этап 2. Оценивание параметров: определение чисел a_1, \dots, a_p и β_1, \dots, β_q .

Этап 3. Проверка соответствия модели и исходных данных.

Приведем описание каждого из этих этапов.

Этап 1 Пусть в качестве кандидатов выбрано несколько пар p и q p^* -это максимальное из всех присутствующих p .

Для каждой такой пары p и q строятся числа a_1, \dots, a_p и β_1, \dots, β_q такие, что величина

$$RSS = \sum_{t=p^*+1}^T \varepsilon_t^2$$

минимальна среди всех отвечающих условию устойчивости наборов $a_1 \dots a_p$ и среди всех наборов β_1, \dots, β_q . Здесь при $t > p^*$

$$\varepsilon_t = y_t - a_0 - a_1 y_{t-1} - \dots - a_p y_{t-p} - \beta_1 \varepsilon_{t-1} - \dots - \beta_q \varepsilon_{t-q}$$

и обычно делается предположение, что $\varepsilon_t = 0$ при $t \leq p^*$

Можно считать $a_0 = 0$, если сразу рассматриваются значения \bar{y} .

Если $q=0$, то определение чисел a_1, \dots, a_p — это задача множественной регрессии. Если $q>0$, то используется тот или иной численный метод нахождения минимума функции нескольких переменных.

Ясно, что модели с большими p и q дают лучшие (меньше RSS), чем модели с меньшими p и q . Но это противоречит принципу экономности.

Для окончательного выбора p и q используется один из следующих критериев, которые позволяют сочетать принцип экономности и минимизацию RSS : информационный критерий Акаике (AIC — Akaike information criterion) и информационный критерий Шварца (BIC — Bayesian information criterion)

$$AIC = T \ln(RSS) - 2n$$

$$BIC = T \ln(RSS) + \ln T n,$$

где $n=p+q$, если строится модель с $a_0 = 0$;
 $n=p+q+1$, если строится модель с $a_0 \neq 0$

Чем меньше значение информационного критерия, тем лучше пара p и q . Окончательно выбирается пара p, q , для которой значение того или иного информационного критерия минимально. (В последующем нами будет разобран пример, когда информационные критерии Акаике и Шварца указывают на разные пары p, q .)

Этап 2 На этом этапе уже описанным способом определяются a_1, \dots, a_p и β_1, \dots, β_q .

Этап 3 При правильном выборе модели построенный ряд из остатков ε_t , должен представлять собой траекторию случайного процесса белого шума. Для проверки этого строятся выборочные автокорреляции $r_k(\varepsilon)$. Величины $r_k(\varepsilon)$ определяются по временному ряду ε_t тем же способом, каким величины r_k определяются по временному ряду y_t . Затем строится статистика Лjungга - Бокса (Ljung and Box (1978))

$$Q(s) = T(T+2) \sum_{k=1}^s \frac{r_k^2(\varepsilon)}{T-k},$$

которая при больших T для случая, когда ε_t — процесс нормального белого шума имеет распределение χ^2 со степенями свободы $2s$.

Статистика Лjungга – Бокса является усовершенствованием применявшейся до этого статистики Бокса – Пирса (Box and Pierce (1970))

$$Q^*(s) = T \sum_{k=1}^s r_k^2(\varepsilon).$$

которая при больших T для случая, когда ε_t – процесс нормального белого шума также имеет распределение χ^2 со степенями свободы $s - p - q$. Статистика Бокса – Пирса иногда применяется и до сих пор, однако статистика Лjungга – Бокса позволяет получать более точные результаты.

Статистики Лjungга – Бокса и Бокса – Пирса называются также статистиками “portmanteau”.

На практике могут быть рассмотрены Q-статистики с несколькими разными s .

При получении неправдоподобно большого значения Q-статистики должна быть отброшена гипотеза о том, что временной ряд ε_t является реализацией случайного процесса белого шума и, следовательно, модель не адекватна.

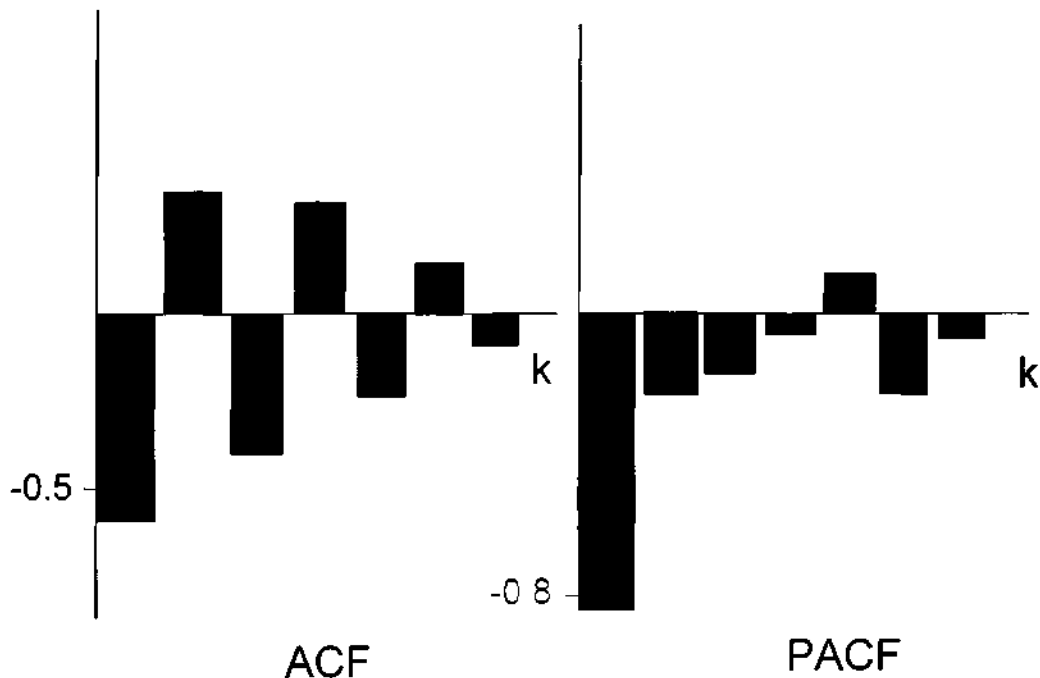
Существуют и другие способы тестирования того, что временной ряд ε_t является траекторией случайного процесса белого шума. Эти методы могут быть использованы вместе или вместо статистики Лjungга – Бокса.

Пример. При помощи 100 стандартных нормальных случайных чисел $\{\varepsilon_t\}$ было построено 100 значений $\{y_t\}$:

$$y_t = -0,7 y_{t-1} + \varepsilon_t - 0,7 \varepsilon_{t-1};$$

y_0 и ε_0 приняты равными 0.

Выборочные автокорреляции и частные автокорреляции имеют следующий вид.



Происхождение данных считается **неизвестным** и сравниваются следующие три модели.

Модель 1 : $y_t = a_1 y_{t-1} + \varepsilon_t$

Модель 2 : $y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1}$

Модель 3 : $y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$

Результаты расчетов таковы.

Модель 1	$a_1 = -0,835$	$Q(8) = 26,19 (0,000)$	$AIC = 507,3$
		$Q(24) = 41,10 (0,001)$	$BIC = 509,9$
Модель 2	$a_1 = -0,679$	$Q(8) = 3,86 (0,695)$	$AIC = 481,4$
	$\beta_1 = -0,676$	$Q(24) = 14,23 (0,892)$	$BIC = 486,6$
Модель 3	$a_1 = -1,16$	$Q(8) = 11,44 (0,057)$	$AIC = 492,5$
	$a_1 = -0,378$	$Q(24) = 22,59 (0,424)$	$BIC = 497,7$

(в скобках – уровни значимости)

Модель 1 должна быть отброшена по значениям статистик $Q(8)$ и $Q(24)$. Предпочтение должно быть отдано модели 2 перед моделью 3 по значениям обоих информационных критериев, а также потому, что в модели 3 значение статистики $Q(8)$ указывает на наличие определенной корреляции между остатками.

Преобразование Бокса-Кокса

Перед построением ARMA модели часто полезно сделать преобразование членов временного ряда (в данном случае предполагается, что все y_t положительны)

$$y_t^* = \begin{cases} (y_t^\lambda - 1) / \lambda, & \text{если выбрано } \lambda \neq 0 \\ \ln(y_t), & \text{если выбрано } \lambda = 0 \end{cases}$$

где число λ должно быть выбрано предварительно. Может оказаться, что для преобразованного временного ряда модель строить удобнее, чем для исходного временного ряда.

Свойства обратимости процессов ARMA (p,q)

Рассмотрим модель

$$A(L) y_t = a_0 + B(L) \varepsilon_t,$$

где

$$\begin{aligned} A(L) &= 1 - a_1 L - a_2 L^2 - \dots - a_p L^p, \\ B(L) &= 1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q. \end{aligned}$$

Если все корни многочлена $B(L)$ лежат вне единичного круга, то имеет место представление

$$\varepsilon_t = B(L)^{-1} (-a_0 + A(L) y_t),$$

и исходная модель может быть представлена в виде сходящегося ряда

$$y_t = \sum_{j=1}^{\infty} \alpha_j y_{t-j} + \varepsilon_t.$$

Такое представление называется AR (∞) представлением.

Об общих множителях многочленов

При построении ARMA моделей

$$A(L)y_t = B(L)\varepsilon_t$$

необходимо следить, чтобы многочлены $A(L)$ и $B(L)$ не имели общих множителей. Рассмотрим, например, процесс

$$y_t = 0,2y_{t-1} + 0,15y_{t-2} + \varepsilon_t + 0,3\varepsilon_{t-1}$$

Заметим, что

$$(1 - 0,2L - 0,15L^2) = (1 - 0,5L)(1 + 0,3L)$$

Поэтому

$$y_t = A^{-1}(L) \cdot B(L)\varepsilon_t = \frac{(1 + 0,3L)}{(1 - 0,5L)(1 + 0,3L)} \varepsilon_t = \frac{\varepsilon_t}{1 - 0,5L}$$

И построенный ARMA (2,1) процесс имеет то же MA(∞) представление, что и AR (1) процесс

$$y_t = 0,5y_{t-1} + \varepsilon_t,$$

т.е. совпадает с ним.

Примеры построения ARMA моделей для конкретных финансовых временных рядов.

Пример На рис. 2.6 показана годовая доходность индекса S&P500. Для этого временного ряда $\bar{y} = 8,21\%$. Первые 12 выборочных автокорреляций и их стандартные ошибки показаны в первых двух столбцах таблицы 2.1. Ни одна из автокорреляций значимо не отличаются от 0. Это вызывает предположение, что моделью для временного ряда может служить процесс белого шума. В столбцах 3 и 4 таблицы 2.1 приведены статистика Бокса-Пирса и Лjung-Бокса

$$Q^*(k) = T \sum_{i=1}^k r_i^2$$

и

$$Q(k) = T(T+2) \sum_{i=1}^k \frac{r_i^2}{T-i},$$

которые асимптотически имеют распределение χ^2 . В квадратных скобках показаны уровни значимости, при которых должна быть отвергнута гипотеза о белом шуме.

В данном случае отбрасывать гипотезу о белом шуме, т.е. о «справедливой» игре нет оснований.

Пример. На рис. 2.7 показана разность между доходностями к погашению для «20 year UK gilts» и «91 day Treasury Bills». Данные приведены поквартально.

Для этого временного ряда $\bar{y} = 1.227$.

В таблице 2.2 показаны выборочные автокорреляции, выборочные частные автокорреляции (до 12 включительно) и соответствующие стандартные ошибки.

Визуально можно определить, что, временной ряд значительно более гладкий, чем реализация процесса белого шума.

Это подтверждается тем, что все выборочные автокорреляции положительны и первые 5 значимо отличны от 0.

Величины «portmanteau» статистик

$$Q^*(12) = 295,2; \quad Q(12) = 305,0$$

Выборочные автокорреляции убывают до $k = 9$, а потом начинают возрастать. Это указывает на возможное присутствие комплексных корней у характеристических уравнений.

Φ_{11} недвусмысленно указывает на $AR(1)$ процесс. Обычный метод наименьших квадратов дает уравнение регрессии

$$y_t = 0,176 + 0,856 y_{t-1} + \varepsilon_t$$

$$(0,098) \quad (0,045)$$

$$\hat{\sigma} = 0,870.$$

В скобках приведены стандартные ошибки.

Для диагностической проверки могут быть использованы найденные при помощи регрессии остатки ε_t . Автокорреляции остатков позволяют найти значение статистик

$$Q^*(12) = 12,26; \quad Q(12) = 13,21.$$

Сравнивая их с распределением χ_{11}^2 видим, что отвергать гипотезу о том, что остатки ε_t являются траекторией процесса белого шума нет оснований

Другой способ проверки адекватности модели состоит в том, чтобы рассмотреть модели с увеличенным числом параметров. В нашем случае это могут быть процессы $AR(2)$ и $ARMA(1,1)$.

Модели оказываются следующие

$$y_t = 0,197 + 0,927 y_{t-1} - 0,079 y_{t-2} + \varepsilon_t, \quad \hat{\sigma} = 0,869;$$

$$(0,101) \quad (0,054) \quad (0,084)$$

$$y_t = 0,213 + 0,831 y_{t-1} + \varepsilon_t - 0,092 \varepsilon_{t-1}, \quad \hat{\sigma} = 0,870.$$

$$(0,104) \quad (0,051) \quad (0,095)$$

В обоих случаях дополнительный параметр оказывается незначимым. Это служит еще одним аргументом в пользу выбранной $AR(1)$ модели.

Пример. Одним из наиболее распространенных в Великобритании индексов акций является Financial Times – Actuaries (FTA) All Share

Таблица 2.3 показывает выборочные автокорреляции и выборочные частные автокорреляции для доходности этого индекса. Используются ежемесячные наблюдения с 1965 по 1990 года.

Хотя значения «portmanteau» статистики $Q(12)$ всего 19,3 (это соответствует уровню значимости 0,08), мы видим, что и r_k и Φ_{kk} для $k = 1$ и $k = 3$ больше, чем две стандартные ошибки. Это указывает на целесообразность построение $ARMA$ модели.

Используем критерии Акаике и Шварца для определения чисел p и q . Выберем два числа P и Q и в качестве параметров модели возьмем p_1 и q_1 такие, что

$$AIC(p_1, q_1) = \min_{\substack{0 \leq p \leq P \\ 0 \leq q \leq Q}} AIC(p, q)$$

или p_2 и q_2 такие, что

$$BIC(p_2, q_2) = \min_{\substack{0 \leq p \leq P \\ 0 \leq q \leq Q}} BIC(p, q)$$

Из таблицы 2.4 видно, что $(p_1, q_1) = (3, 0)$ и $(p_2, q_2) = (0, 1)$. В этом нет противоречия. Вспомним о $AR(\infty)$ представлением $MA(1)$ модели).

Построенные модели имеют вид:

$$y_t = 4.41 + 0.173 y_{t-1} - 0.108 y_{t-2} + 0.144 y_{t-3} + \varepsilon_t \quad \hat{\sigma} = 6.25$$

(0.62) (0.057) (0.057) (0.057)

$$y_t = 5.58 + \varepsilon_t + 0.186 \varepsilon_{t-1} \quad \hat{\sigma} = 6.29$$

(0.35) (0.056)

Заметим, что свободный член в $AR(3)$ модели дает оценку для среднего $4.41 / (1 - 0.173 + 0.108 - 0.144) = 5.58$,

которая в точности совпадает с свободным членом в $MA(1)$ модели. Многочлен первого порядка

$$(1 + 0.186 L)$$

из $MA(1)$ модели может быть обращен

$$(1 + 0.186 L)^{-1} = (1 - 0.186 L + 0.186^2 L^2 - 0.186^3 L^3 + \dots) = (1 - 0.186 L + 0.035 L^2 - 0.006 L^3 + \dots)$$

Использование $MA(1)$ приводит к экспоненциальному убыванию параметров при $y_{t-1}, y_{t-2}, y_{t-3}, \dots$. Поэтому, возможно, $AR(3)$ модель является лучшей.

Обе модели выдерживают диагностическую проверку:

для $MA(1)$	$Q(12) = 15.30$,
для $AR(3)$	$Q(12) = 6.07$.

Это соответствует уровням значимости 0.23 и 0.91 соответственно.

Построение прогнозов

Модель $ARMA(p, q)$

$$A(L) y_t = a_0 + B(L) \varepsilon_t$$

путем преобразования

$$y_t = A(L)^{-1} [a_0 + B(L) \varepsilon_t]$$

приводится к $MA(\infty)$ виду (будем рассматривать случай $a_0 = 0$, к которому можно перейти, заменив все y_t на $y_t - \bar{y}$)

$$y_t = \sum_{r=0}^{\infty} c_r \varepsilon_{t-r}, \quad c_0 = 1.$$

Тогда при $j > 0$

$$y_{t+j} = \sum_{r=0}^{\infty} c_r \varepsilon_{t+j-r}.$$

$E_t y_{t+j}$ -сокращенная запись для условного математического ожидания случайной величины y_{t+j} при условии, что известна вся информация до момента t включительно:

$$E y_{t+j} = E(y_{t+j} | y_t, y_{t-1}, \dots, \varepsilon_t, \varepsilon_{t-1}, \dots)$$

Через $f_{t,j}$ обозначается прогноз величины y_{t+j} , сделанный в момент времени t .
Принимается

$$f_{t,j} = E y_{t+j}$$

Тогда

$$f_{t,j} = \sum_{r=j}^{\infty} c_r \varepsilon_{t+j-r}.$$

Для практического построения прогнозов используется следующий прием.
Имеем

$$y_{t+j} = a_1 y_{t+j-1} + \dots + a_p y_{t+j-p} + \varepsilon_{t+j} + \beta_1 \varepsilon_{t+j-1} + \dots + \beta_q \varepsilon_{t+j-q}.$$

Тогда

$$E_t y_{t+j} = a_1 E_t y_{t+j-1} + \dots + a_p E_t y_{t+j-p} + E_t \varepsilon_{t+j} + \beta_1 E_t \varepsilon_{t+j-1} + \dots + \beta_q E_t \varepsilon_{t+j-q}. \quad (1)$$

Индексы величин, входящих в эту формулу могут быть как больше, так и меньше t , в зависимости от соотношений между j , p и q .

Заметим, что при $i=0, -1, -2, \dots, -(t-1)$

$$E_t y_{t+i} = y_{t+i} \quad (2)$$

$$E_t \varepsilon_{t+i} = \varepsilon_{t+i}$$

При $i = 1, \dots, j$

$$E_t y_{t-i} = f_{t,i} \quad (3)$$

$$E_t \varepsilon_{t+i} = 0.$$

Пример: Рассмотрим процесс $ARMA(2,1)$

$$y_t = 1.8 y_{t-1} - 0.9 y_{t-2} + \varepsilon_t - 0.7 \varepsilon_{t-1}$$

Тогда по формуле (1) с учетом соотношений (2) и (3)

$$f_{t,1} = 1.8 y_t - 0.9 y_{t-1} - 0.7 \varepsilon_t$$

$$f_{t,2} = 1.8 f_{t,1} - 0.9 y_t$$

$$f_{t,j} = 1.8 f_{t,j-1} - 0.9 f_{t,j-2} \text{ при } j=3,4,\dots$$

Для определения значений $\varepsilon_t, \varepsilon_{t-1}, \dots$ используется представление

$$\varepsilon_t = y_t - a_1 y_{t-1} - \dots - a_p y_{t-p} - \beta_1 \varepsilon_{t-1} - \dots - \beta_q \varepsilon_{t-q}$$

Считая

$$\varepsilon_t = 0 \text{ при } t \leq p$$

из этой формулы могут быть получены значения

$$\varepsilon_{p+1}, \dots, \varepsilon_t$$

Если p достаточно велико, а ряд не очень длинный, то могут быть использованы и другие приемы для определения

$$\varepsilon_t \text{ при } t \leq p$$

Например, существует прием, называемый «прогнозированием назад».

Ошибка прогноза определяется как

$$e_{t,j} = y_{t+j} - f_{t,j}.$$

Имеем:

$$e_{t,j} = \sum_{r=0}^{j-1} c_r \varepsilon_{t+j-r}.$$

Отсюда

$$E(e_{t,j}) = 0,$$

$$\text{Var}(e_{t,j}) = \sigma^2 \sum_{r=0}^{j-1} c_r^2 .$$

При условии нормальности белого шума этих данных достаточно, чтобы построить доверительный интервал для прогноза (при фиксированном j).

Упражнение: Доказать формулу для пересмотра прогнозов:

$$f_{t+1,j} = f_{t,j+1} + c_j (y_{t+1} - f_{t,j}).$$

Сезонность

Если используется ,например, квартальные данные , то могут быть использованы «чистые» сезонные модели, например,:

$$y_t = a_4 y_{t-4} + \varepsilon_t, \quad |a_4| < 1$$

или

$$y_t = \varepsilon_t + \beta_4 \varepsilon_{t-4}.$$

Могут использоваться смешанные модели. например:

$$y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_4 \varepsilon_{t-4} .$$

Это был пример аддитивной сезонности. Используется и **мультипликативная** сезонность , например:

$$(1 - a_1 L) y_t = (1 + \beta_1 L) (1 + \beta_4 L^4) \varepsilon_t.$$

Это уравнение совпадает с уравнением

$$y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_4 \varepsilon_{t-4} + \beta_1 \beta_4 \varepsilon_{t-5}$$

Преимуществом такого подхода является то, что оценив только три коэффициента, можно учесть эффекты авторегрессии с запаздыванием 1 и эффекты скользящего среднего с запаздыванием 1, 4 и 5.

Процедура *X-11* для снятия сезонности, используемая в США.

Модели временных рядов, включающие условную гетероскедастичность

Процесс $AR(p)$ имеет вид

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + \varepsilon_t,$$

где ε_t – белый шум, т.е. $E(\varepsilon_t) = 0$,

$$E(\varepsilon_t \varepsilon_{t+s}) = \begin{cases} \sigma^2 & \text{при } s = 0 \\ 0 & \text{при } s \neq 0 \end{cases}$$

Этот процесс рассматривается при условии, что все корни характеристического уравнения

$$z^p - a_1 z^{p-1} - \dots - a_p = 0$$

лежат строго внутри единичного круга. Данное условие означает ковариационную стационарность этого процесса.

Условное математическое ожидание, являющееся прогнозом для соответствующего значения, определяется по формуле:

$$E(y_t | y_{t-1}, y_{t-2}, \dots) = E_{t-1}(y_t) = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p}.$$

$E_{t-1}(y_t)$ изменяется в зависимости от t .

Безусловное математическое ожидание от t не зависит:

$$E(y_t) = \frac{a_0}{1 - a_1 - a_2 - \dots - a_p}.$$

В ряде задач интерес представляет не только прогноз **среднего значения y_t** , но и прогноз дисперсии. Бывает, что процентные ставки в некоторые периоды **облачаются** большей изменчивостью (волатильностью), чем в другие периоды.

Изменения в волатильности важны для понимания финансовых рынков, **поскольку** инвесторы требуют более высокую ожидаемую доходность, как компенсацию за **более** высокий уровень риска, измеряемый волатильностью.

Отсюда возникает идея в качестве ε_t использовать либо случайный процесс белого шума специального вида, условная дисперсия которого меняется в зависимости от времени.

Модели авторегрессии условной гетероскедастичности (ARCH)

Engle (1982). Обозначение $ARCH(m)$, $m \geq 1$ используется для двух разных моделей.

Модель 1. ε_t^2 является $AR(m)$ процессом:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_m \varepsilon_{t-m}^2 + w_t$$

где w_t – новый процесс белого шума:

$$E(w_t) = 0$$

$$E(w_t w_{t+s}) = \begin{cases} \lambda^2 & \text{при } s = 0 \\ 0 & \text{при } s \neq 0 \end{cases}$$

Чтобы величина ε_t^2 была всегда положительной, следует потребовать $\alpha_0 > 0$, и случайная величина w_t не должна принимать значения меньше α . Кроме того, $\alpha_j \geq 0$ при $j = 1, 2, \dots, m$ и корни многочлена

$$z^m - \alpha_1 z^{m-1} - \dots - \alpha_m$$

должны лежать строго внутри единичного круга.

Такую модель называют моделью с аддитивной условной гетероскедастичностью.

Утверждение (без доказательства):

Если $\alpha_j \geq 0$ для всех $j = 1, 2, \dots, m$, то корни многочлена $z^m - \alpha_1 z^{m-1} - \dots - \alpha_m$ лежат строго внутри единичного круга т. и т.т., когда

$$\alpha_1 + \alpha_2 + \dots + \alpha_m < 1.$$

Безусловное математическое ожидание случайной величины ε_t^2 оказывается равным

$$E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_m}.$$

$$E_{t-1}(\varepsilon_t^2) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_m \varepsilon_{t-m}^2$$

Условное математическое ожидание $E_t(\varepsilon_{t+s}^2)$ строится как для процессов авторегрессии.

Модель 2.

$$\varepsilon_t = \sqrt{h_t} \cdot v_t,$$

где $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_m \varepsilon_{t-m}^2$;

$\{v_t\}$ - последовательность независимых одинаково распределенных случайных величин, таких, что $E(v_t) = 0$, $E(v_t^2) = 1$.

Такую модель называют моделью с мультипликативной условной гетероскедастичностью.

Для такой модели условное математическое ожидание

$$E_{t-1}(\varepsilon_t) = \sqrt{h_t} \cdot E(v_t) = 0$$

и условная дисперсия

$$E_{t-1}(\varepsilon_t^2) = h_t,$$

как и в случае модели с аддитивной условной гетероскедастичностью. Точно так же устанавливается, что $E(\varepsilon_t) = 0$. Но для безусловной дисперсии справедливо соотношение

$$E(\varepsilon_t^2) = E(h_t)E(v_t^2) = E(h_t) = \alpha_0 + \alpha_1 E(\varepsilon_{t-1}^2) + \dots + \alpha_m E(\varepsilon_{t-m}^2).$$

Если случайный процесс ε_t стационарный (ниже это будет доказано и даже в более общей ситуации), то

$$E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2) = \dots = E(\varepsilon_{t-m}^2),$$

и из полученного соотношения следует, что

$$E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_m}.$$

Задача. Пусть при $m = 1$ построена модель 2 для временного ряда. При каких α_1 существует эквивалентная ей модель 1, и при каких α_1 – нет?

Решение. Рассмотрим случайный процесс ε_t , построенный по модели 2. Будем считать $m = 1$ и

$$v_t \sim N(0, 1) \text{ при любом } t.$$

Предположим, что для процесса ε_t^2 существует представление в форме модели 1. Тогда должно выполняться соотношение

$$h_t \cdot v_t^2 = h_t + w_t$$

или

$$w_t = h_t(v_t^2 - 1).$$

Напомним, что

$$E(w_t^2) = \lambda^2 .$$

Имеем

$$E(w_t^2) = E(h_t^2) \cdot E((v_t^2 - 1)^2) \quad (*)$$

Для нормально стандартной случайной величины

$$E((v_t^2 - 1)^2) = 2 .$$

Поэтому

$$\begin{aligned} E(h_t^2) &= E((\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^2) = \alpha_0^2 + 2\alpha_0 \alpha_1 E(\varepsilon_{t-1}^2) + \alpha_1^2 E(\varepsilon_{t-1}^4) = \\ &= \alpha_0^2 + 2\alpha_1 \alpha_0 E(\varepsilon_{t-1}^2) + \alpha_1^2 (Var(\varepsilon_{t-1}^2) + (E(\varepsilon_{t-1}^2))^2) . \end{aligned}$$

По предположению случайный процесс ε_t^2 является процессом $AR(1)$. Отсюда следует, что

$$E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1} , \quad Var(\varepsilon_t^2) = \frac{\lambda^2}{1 - \alpha_1^2}$$

Отсюда

$$E(h_t^2) = \alpha_0^2 + 2\alpha_1 \alpha_0 \frac{\alpha_0}{1 - \alpha_1} + \alpha_1^2 \left(\frac{\lambda^2}{1 - \alpha_1^2} + \frac{\alpha_0^2}{(1 - \alpha_1)^2} \right) = \frac{\alpha_0^2}{1 - \alpha_1^2} + \frac{\alpha^2}{(1 - \alpha_1)^2} .$$

Из уравнения (*)

$$\lambda^2 = 2E(h_t^2)$$

получаем

$$\frac{(1 - 3\alpha_1^2)\lambda^2}{1 - \alpha_1^2} = \frac{2\alpha_0^2}{(1 - \alpha_1)^2} .$$

Если $\alpha_1^2 \geq \frac{1}{3}$, то нами получено противоречие, поскольку данное уравнение не имеет действительных решений.

Если $\alpha_1^2 < \frac{1}{3}$, то при условии нормальности случайной величины v_t ответ на поставленный вопрос положителен. В общем случае может не существовать конечного $E((v_t^2 - 1)^2)$.

В дальнейшем, нами будут рассматриваться модели с мультипликативной условной гетероскедастичностью и их обобщения.

Определение параметров ARCH-модели методом максимального правдоподобия

Рассмотрим уравнение регрессии

$$y_t = x_t \cdot \beta - \varepsilon_t .$$

x_t - вектор, состоящий из величин, известных к моменту времени t ; некоторыми из его координат могут быть значения y , относящиеся к более ранним моментам времени. β - вектор той же размерности, что x_t .

$$\varepsilon_t = \sqrt{h_t} v_t ,$$

где $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_m \varepsilon_{t-m}^2$;

Будем считать, что кроме тех свойств, которыми случайные величины v_t обладали ранее, они обладают еще и свойством нормальности.

Пусть временной ряд занумерован следующим образом

$$y_{-m+1}, y_{-m+2}, \dots, y_0, y_1, \dots, y_T$$

Обозначим через Y_t совокупность наблюдений, полученных к моменту времени t :

$$Y_t = (y_t, y_{t-1}, \dots, y_{-m+1}, x_t, x_{t-1}, \dots, x_{-m+1})$$

Условное распределение случайной величины y_t является нормальным со средним значением $x_t \beta$ и с дисперсией h_t :

$$f(y_t | x_t, Y_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} e^{-\frac{(y_t - x_t \beta)^2}{2h_t}},$$

где

$$h_t = \alpha_0 + \alpha_1 (y_{t-1} - x_{t-1} \beta)^2 + \dots + \alpha_m (y_{t-m} - x_{t-m} \beta)^2$$

Тогда функция максимального правдоподобия

$$L(\alpha, \beta) = \sum_{t=1}^T \log f(y_t | x_t, Y_{t-1}; \alpha, \beta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t) - \frac{1}{2} \sum_{t=1}^T (y_t - x_t \beta)^2 / h_t$$

Следует выбрать ту пару (α, β) , при которой $L(\alpha, \beta)$ достигает максимального значения. Созданы специализированные численные методы для максимизации функции L приведенного вида.

Аналогично могут быть рассмотрены случаи, когда случайные величины V_t имеют некоторые другие распределения, отличные от нормального. Например, t -распределение.

Тестирование ARCH – модели

В работе Engle (1982) и в последующих работах развит следующий метод тестирования. Фиксируется некоторое m .

Строится уравнение регрессии

$$y_t = x_t \beta + \varepsilon_t$$

для $t = -m+1, -m+2, \dots, T$ и находятся остатки $\hat{\varepsilon}_t$. Затем строится уравнение регрессии для $\hat{\varepsilon}_t^2$ на константу и m собственных лаговых значений:

$$\begin{aligned} \hat{\varepsilon}_t^2 &= \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \dots + \alpha_m \hat{\varepsilon}_{t-m}^2 + e_t; \\ h_t &= \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \dots + \alpha_m \hat{\varepsilon}_{t-m}^2; \quad t = 1, 2, \dots, T. \end{aligned}$$

Затем строится вектор

$$f_0 = \left(\frac{e_1^2}{h_1} - 1, \dots, \frac{e_T^2}{h_T} - 1 \right)$$

и матрица

$$Z_0 = \left(h_1 \frac{\partial h_1}{\partial \alpha}, \dots, h_T \frac{\partial h_T}{\partial \alpha} \right)$$

Статистика имеет вид

$$\xi = 0,5 \times f_0' \times Z_0 \times (Z_0' \times Z_0)^{-1} \times Z_0' \times f_0,$$

где штрих означает транспонирование.

На реализациях $ARCH(m)$ процесса ξ асимптотически имеет хи-квадрат распределение с m степенями свободы. Получение неправдоподобно большого значения ξ указывает на необходимость увеличения m либо на необходимость другой модификации модели.

Обобщенная авторегрессия условная гетероскедастичность (GARCH)

Как и в предыдущих разделах, наш анализ относится к моделям вида

$$y_t = x_t \beta + \varepsilon_t,$$

когда случайные возмущения ε_t не являются гомоскедастичными. Наше внимание сосредоточено на моделях для случайных возмущений.

Bollerslev (1986). В выражении для случайного возмущения $\varepsilon_t = \sqrt{h_t} v_t$, где $\{v_t\}$ - последовательность независимых одинаково распределенных случайных величин таких, что

$$E(v_t) = 0, \quad E(v_t^2) = 1.$$

Величины h_t имеют вид

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} = \alpha_0 + A(L) \varepsilon_t^2 + B(L) h_t, \quad (*)$$

где

$$\begin{aligned} p \geq 0, \quad q > 0, \\ \alpha_0 > 0, \quad \alpha_i \geq 0, \quad i=1, \dots, q \\ \beta_i \geq 0, \quad i=1, \dots, p \end{aligned}$$

$$A(L) = \sum_{i=1}^q \alpha_i L^i; \quad B(L) = \sum_{i=1}^p \beta_i L^i$$

Такой процесс называется *GARCH (p,q) процессом*. При $p=0$ процесс превращается в *ARCH (q) процесс*; при $p=q=0$ процесс $\{\varepsilon_t\}$ с точностью до константы совпадает с гомоскедастичным белым шумом $\{v_t\}$.

Если все корни многочлена $1-B(z)$ лежат строго вне единичного круга, то

$$h_t = \frac{\alpha_0 + A(L) \varepsilon_t^2}{1 - B(L)} = \frac{\alpha_0}{1 - \sum_{i=1}^p \beta_i} + \sum_{i=1}^{\infty} \delta_i \varepsilon_{t-i}^2 \quad (**)$$

Может возникнуть впечатление из выражения для h_t , что $B(L) h_t$ - это члены "авторегрессии", а $A(L) \varepsilon_t^2$ - это члены "скользящего среднего". Но такой вывод является неверным. Чтобы, понять это, добавим ε_t^2 к левой и правой частям уравнения (*) и проведем некоторые преобразования.

$$h_t + \varepsilon_t^2 = \alpha_0 - \sum_{i=1}^p \beta_i (\varepsilon_{t-i}^2 - h_{t-i}) + \sum_{i=1}^p \beta_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \varepsilon_t^2.$$

Пусть $m = \max(p, q)$; положим $\alpha_i = 0$ при $i > q$ и $\beta_i = 0$ при $i > p$. Введем обозначение

$$\eta_t = \varepsilon_t^2 - h_t = (v_t^2 - 1) h_t$$

Тогда $E(\eta_t) = 0$ и $Cov(\eta_t, \eta_{t+s}) = 0$ при $s \neq 0$. Имеем

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \eta_{t-i} + \eta_t. \quad (***)$$

Т.е. процесс для ε_t^2 напоминает процесс *ARMA (m, p)*. Но при произвольном наборе $\alpha_1, \dots, \alpha_q$ дисперсия случайных величин η_t может не существовать. Последняя запись используется для построения прогнозов $E_t(\varepsilon_{t+s}^2)$ аналогично линейной модели.

Nelson and Cao (1992) заметили, что условия $\alpha_i \geq 0$, $\beta_i \geq 0$ при $i \geq 1$ достаточны, но не необходимы для неотрицательности h_t . Согласно представлению (***) при $\alpha_0 > 0$, $\sum_{i=1}^p \beta_i < 1$ для неотрицательности h_t достаточно, чтобы выполнялось условие $\delta_i \geq 0$ при $i \geq 1$. Например, для процесса $GARCH(1,2)$ имеем

$$\begin{aligned} \sum_{i=1}^2 \delta_i L^i &= (1 - \beta_1 L)^{-1} (\alpha_1 L + \alpha_2 L) = (1 - \beta_1 L + \beta_1^2 L + \dots) (\alpha_1 L + \alpha_2 L) = \\ &= \alpha_1 L + (\beta_1 \alpha_1 + \alpha_2) L^2 + \beta_1 (\beta_1 \alpha_1 + \alpha_2) L^3 + \beta_1^2 (\beta_1 \alpha_1 + \alpha_2) L^4 + \dots \end{aligned}$$

Коэффициенты δ_i неотрицательны, если $\alpha_1 \geq 0$, $0 \leq \beta_1 < 1$ и $\alpha_2 \geq -\alpha_1 \beta_1$, при этом α_2 может быть отрицательным.

Напомним, что $\varepsilon_t = \sqrt{h_t} v_t$, где $\{v_t\}$ - последовательность независимых одинаково распределенных случайных величин таких, что

$$E(v_t) = 0, \quad E(v_t^2) = 1.$$

Величины h_t имеют вид

$$\begin{aligned} h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} = \alpha_0 + A(L) \varepsilon_t^2 + B(L) h_t, \\ A(z) &= \sum_{i=1}^q \alpha_i z^i; \quad B(z) = \sum_{i=1}^p \beta_i z^i. \end{aligned}$$

Теорема. Если $A(1) + B(1) < 1$, то случайный процесс ε_t типа $GARCH(p,q)$ является ковариационно стационарным с $E(\varepsilon_t) = 0$, $Cov(\varepsilon_t, \varepsilon_{t+s}) = 0$ при $s \neq 0$,

$$Var(\varepsilon_t) = \frac{\alpha_0}{(1 - A(1) - B(1))}$$

Доказательство. Из представления $\varepsilon_t = \sqrt{h_t} v_t$, где $\{v_t\}$ - последовательность независимых одинаково распределенных случайных величин таких, что $E(v_t) = 0$, $E(v_t^2) = 1$, следует утверждение о математических ожиданиях, а из него и о ковариациях. В доказательстве нуждается формула для дисперсии.

Имеем

$$\begin{aligned} h_t &= \alpha_0 + \sum_{j=1}^q \alpha_j v_{t-j}^2 h_{t-j} + \sum_{j=1}^p \beta_j h_{t-j} = \alpha_0 + \sum_{j=1}^q \alpha_j v_{t-j}^2 (\alpha_0 + \sum_{i=1}^q \alpha_i v_{t-j-i}^2 h_{t-j-i} + \sum_{i=1}^p \beta_i h_{t-j-i}) + \\ &+ \sum_{j=1}^p \beta_j (\alpha_0 + \sum_{i=1}^q \alpha_i v_{t-j-i}^2 h_{t-j-i} + \sum_{i=1}^p \beta_i h_{t-j-i}) = \alpha_0 \cdot \sum_{i=0}^k M(t, k), \end{aligned}$$

где $M(t, k)$ - это сумма всех слагаемых вида

$$\prod_{i=1}^q \alpha_i^{a_i} \cdot \prod_{j=1}^p \beta_j^{b_j} \cdot \prod_{i=1}^n v_{t-s_i}^2.$$

При $k \geq 1$

$$\sum_{i=1}^q a_i + \sum_{j=1}^p b_j = k; \quad \sum_{i=1}^q a_i = n; \quad 1 \leq s_1 < s_2 < \dots < s_n \leq \max(kq, (k-1)q + p).$$

В частности,

$$M(t, 0) = 1$$

$$M(t, 1) = \sum_{i=1}^q \alpha_i v_{t-i}^2 + \sum_{i=1}^p \beta_i$$

$$M(t,2) = \sum_{j=1}^q \alpha_j v_{t-j}^2 \left(\sum_{i=1}^q \alpha_i v_{t-j-i}^2 + \sum_{i=1}^p \beta_i \right) + \sum_{j=1}^p \beta_j \left(\sum_{i=1}^q \alpha_i v_{t-j-i}^2 + \sum_{i=1}^p \beta_i \right).$$

В общем случае

$$M(t, k+1) = \sum_{i=1}^q \alpha_i v_{t-i}^2 M(t-i, k) + \sum_{i=1}^p \beta_i M(t-i, k)$$

Поскольку случайные величины v_i^2 независимы и одинаково распределены, случайные величины $M(t, k)$ одинаково распределены при всех t . В частности, $E(M(t, k)) = E(M(s, k))$ для всех k, t, s .

Из последних двух формул имеем

$$E(M(t, k+1)) = \left(\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i \right) \cdot E(M(t, k)) = (A(1) + B(1)) \cdot E(M(t, k)) = \dots = (A(1) + B(1))^{k+1} \cdot E(M(t, 0)) = (A(1) + B(1))^{k+1}.$$

Наконец,

$$E(\varepsilon_t^2) = \alpha_0 E\left(\sum_{k=0}^{\infty} M(t, k)\right) = \alpha_0 \sum_{k=0}^{\infty} E(M(t, k)) = \alpha_0 (1 - A(1) - B(1))^{-1}$$

по формуле для суммы геометрической прогрессии. \square

Пример. Рассмотрим 470 наблюдений значений обменного курса доллар фунт. Данные недельные с января 1980 до декабря 1988. Среднее значение \bar{x} в книге Mills не приведено. Значимых SACF и SPACF у этого ряда нет.

Рассмотрим подробнее ряд $x_t = \Delta y_t$. По анализу SACF, SPACF, статистики portmanteau нет оснований отвергать гипотезу, что временной ряд x_t является траекторией случайного процесса белого шума. Кроме ряда x_t рассмотрим ряды:

$$s_t = (x_t - \bar{x})^2 \quad m_t = |x_t - \bar{x}|.$$

Первые 12 SACF r_k для этих рядов, а также статистики portmanteau приведены ниже.

k	Ряд		
	x_t	s_t	m_t
1	-0,064	0,073	0,090
2	0,012	0,053	0,104
3	0,077	0,151	0,181
4	0,054	0,129	0,148
5	-0,021	0,025	0,029
6	0,003	-0,015	0,023
7	0,084	0,054	0,076
8	-0,011	0,021	0,048
9	0,013	0,066	0,107
10	0,042	0,021	0,059
11	-0,060	0,125	0,129
12	0,017	0,021	0,053
s.e.	0,046	0,046	0,046
$Q(12)$	12,35	33,71	54,72

s.e. – асимптотическая стандартная ошибка для r_k . $Q(12)$ – статистика portmanteau с распределением χ_{12}^2 .

$$\bar{x} = -0.000917$$

$$\bar{s} = 0.000785$$

$$\bar{m} = 0.00204$$

Итак, изменения x_t не содержат значимых автокорреляций, но квадраты изменений s_t значимые автокорреляции содержат. Поскольку \bar{x} очень мало, то вместо ряда s_t будем рассматривать ряд x_t^2 . Используя обозначение x_t вместо ε_t , модель можно записать в виде $x_t = \sqrt{h_t} v_t$.

Тестируя на условную гетероскедастичность при $m = 4$ получаем значение $\xi = 18,71$. Эта статистика должна иметь распределение χ_4^2 . Значение 18,71 однозначно показывает недостаточность модели $ARCH(4)$

Строится $GARCH(1,1)$ – модель:

$$x_t^2 = 7.84 \cdot 10^{-4} + 0.9211 x_{t-1}^2 + \eta_t - 0.871 \eta_{t-1}$$

$$(9.04 \cdot 10^{-4}) \quad (0.327) \quad (0.332)$$

(ср. с формулой (***)). В скобках приведены стандартные ошибки, указывающие на значимость коэффициентов при x_{t-1}^2 и η_{t-1} .

Модели ARCH-M (ARCH in mean)

Engle, Lilen, Robbins (1987)

$$y_t = x_t \beta + \delta h_t + \varepsilon_t$$

$$\varepsilon_t = \sqrt{h_t} v_t$$

$$h_t = \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2$$

где, как и раньше, v_t – последовательность независимых одинаково распределенных случайных величин, таких, что

$$E(v_t) = 0, E(v_t^2) = 1$$

При $\delta=0$ данная модель превращается в модель $ARCH(m)$.

Пример. По данным с 1960 до 1984 для США был построен ряд

$$y_t = [(1 + R_t) - (1 + r_{t-1})] - (1 - r_t)$$

где R_t – доходность шестимесячного Treasury bill, r_t – 3-х месячного Treasury bill, измеренные в начале квартала. Приблизительно:

$$y_t = 2R_t - r_{t+1} - r_t$$

Оказалось, что $\bar{y} = 0.142\%$ в квартальном исчислении. В годовом исчислении это соответствует 0.568%

Данная операция соответствует **двукратному займу** в долг на 3 мес. и кредитованию на 6 мес. В течение **рассматриваемого** периода максимальная доходность от такой операции составила 8.2%. **максимальные** потери составили – 3.1% (в годовом исчислении)

Проведенное ARCH-тестирование дало значение $\xi = 10.1$ при $m=1$. Если бы условной гетероскедастичности не было, ξ должно было бы быть значением случайной величины, имеющей χ_1^2 распределение. Но значение 10.1 у такой случайной величины маловероятно.

Временной ряд y_t в период с примерно 1979 до 1982 имел колебания в несколько раз большие, чем до этого. В период с 1961 до 1967 колебания наоборот, были очень низкими. После сравнения ряда моделей, наиболее удачной для данного ряда была признана модель

$$y_t = 0.355 + 0.135 \log h_t + \varepsilon_t$$

$$(4.38\%) \quad (3.36\%)$$

$$\varepsilon_t = \sqrt{h_t} v_t,$$

$$h_t = 0.005 + 1.148 \sum_{i=1}^4 w_i \varepsilon_{t-i}^2$$

$$(2.22\%) \quad (5.56\%)$$

где $w_i = (5 - i)/10$.

Другие примеры использования GARCH-M моделей для анализа финансовых временных рядов описаны в книге Mills'a, раздел 5.2.

Спектральный анализ временных рядов

Ряд Фурье интегрируемой функции

Пусть $f: [0, 2\pi] \rightarrow \mathbf{R}$ интегрируемая функция. Будем считать, что f продолжена до 2π -периодичной функции на всей прямой. Для всех целых $k \geq 0$ определим числа

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx$$

Эти числа называются *коэффициентами Фурье функции f* и впервые были использованы Фурье в 1807 г. при анализе процессов передачи тепла

Теорема Римана - Лебега утверждает, что для любой интегрируемой функции ее коэффициенты Фурье стремятся к 0 при $k \rightarrow \infty$.

Ряд

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

называется *рядом Фурье функции f* . Тригонометрические полиномы

$$S_N(f; x) = \frac{a_0}{2} + \sum_{k=1}^N (a_k \cos(kx) + b_k \sin(kx))$$

называются *частичными суммами ряда Фурье*. Для произвольной интегрируемой функции f и произвольного x из $[0, 2\pi]$ предел последовательности $\{S_N(f; x)\}_{N=1}^{\infty}$ может не существовать. Может и существовать, но не быть равным $f(x)$.

Приведем для примера следующее достаточное условие того, что $\lim_{N \rightarrow \infty} S_N(f; x) = f(x)$.

Теорема (Липшиц, 1864г.) Пусть x из $[0, 2\pi]$ и существуют $M > 0$ и $\alpha > 0$ такие, что при всех h достаточно близких к 0

$$|f(x+h) - f(x)| \leq M|h|^\alpha,$$

тогда $\lim_{N \rightarrow \infty} S_N(f; x) = f(x)$.

Теорема. Пусть даны две последовательности $\{a_k\}$ и $\{b_k\}$, такие, что

$$\sum_{k=1}^{\infty} |a_k| < \infty, \quad \sum_{k=1}^{\infty} |b_k| < \infty.$$

Для каждого $N \in \mathbf{N}$ построим тригонометрический полином

$$S_N(x) = \frac{a_0}{2} + \sum_{k=1}^N (a_k \cos(kx) + b_k \sin(kx)).$$

Тогда существует непрерывная 2π -периодическая функция f такая, что $\forall x \in [0, 2\pi)$ $S_N(x) \rightarrow f(x)$ при $N \rightarrow \infty$ и числа $\{a_k\}$ и $\{b_k\}$ являются коэффициентами Фурье функции f .

Хотя доказательство последней теоремы несложно, мы его опустим.

Заметим, что если функция f четная, то все коэффициенты $b_k = 0$.

Спектральная плотность стационарного случайного процесса и выборочная периодограмма временного ряда

Пусть $\{y_t\}_{t=-\infty}^{\infty}$ стационарный в широком смысле случайный процесс;

$$E(y_t) = \mu; \quad E((y_t - \mu)(y_{t+k} - \mu)) = \gamma_k.$$

Если $\sum_{k=1}^{\infty} |\gamma_k| < \infty$, то функция

$$s(w) = \frac{1}{2\pi} (\gamma_0 + 2 \sum_{k=1}^{\infty} (\gamma_k \cos(kw)))$$

называется *спектральной плотностью (спектром)* случайного процесса $\{y_t\}$.

Задача. Построить спектральные плотности для процессов $MA(1)$ и $AR(1)$

Рассмотрим временной ряд y_1, y_2, \dots, y_T со средним значением $\bar{y} = \frac{1}{T} \sum_{i=1}^T y_i$.

Выборочные автоковариации определяются по формуле

$$g_j = \frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y})(y_{t-j} - \bar{y}) \text{ при } j=0, 1, \dots, T-1.$$

$$g_j = g_{-j} \text{ при } j=-1, -2, \dots, -T+1.$$

Выборочная периодограмма определяется

$$\hat{s}(w) = \frac{1}{2\pi} (g_0 + 2 \sum_{j=1}^{T-1} g_j \cos(jw)).$$

Нетрудно увидеть, что выборочная дисперсия g_0 может быть определена по формуле

$$g_0 = \int_{-\pi}^{\pi} \hat{s}(w) dw = 2 \int_0^{\pi} \hat{s}(w) dw.$$

При определении параметров p и q для модели $ARMA(p, q)$ ранее рассматривались выборочные автокорреляции r_k и выборочные частные автокорреляции $\hat{\Phi}_{kk}$, $k=1, 2, \dots$. При таком подходе в качестве модели $ARMA(p, q)$ для данного временного ряда выбирается тот случайный процесс, чьи автокорреляции ρ_k и частные автокорреляции Φ_{kk} имеют форму сходную с r_k и $\hat{\Phi}_{kk}$. Другой подход, который в настоящее время применяется реже, выбрать в качестве модели случайный процесс $ARMA(p, q)$, чья спектральная плотность $s(w)$ имеет форму сходную с выборочной периодограммой $\hat{s}(w)$ рассматриваемого временного ряда.

Определим величины a_j и b_j , которые можно интерпретировать, как аналог коэффициентов Фурье для временного ряда. Ограничимся случаем нечетного T , для четных T формулы получаются более громоздкими

Пусть T – нечетное целое число, $M=(T-1)/2$ и $w = 2\pi/T$, при $j=1, 2, \dots, M$.

Положим при $j=1, 2, \dots, M$

$$a_j = \frac{2}{T} \sum_{t=1}^T y_t \cos(w_j(t-1)), \quad b_j = \frac{2}{T} \sum_{t=1}^T y_t \sin(w_j(t-1)).$$

Следующая теорема дает выражение членов **временного** ряда через его “коэффициенты Фурье”, аналогичное представлению **периодической** функции через ее частичные суммы ряда Фурье.

Теорема.

При $t=1, 2, \dots, T$,

$$y_t = \bar{y} + \sum_{j=1}^M (a_j \cos(\omega_j(t-1)) + b_j \sin(\omega_j(t-1))).$$

Следующая теорема, во-первых, позволяет выразить через “коэффициенты Фурье” выборочную дисперсию, и, во-вторых, по виду выборочной периодограммы позволяет определить, какие из гармоник являются ведущими в представлении временного ряда, если такие гармоники имеются.

Теорема.

$$\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2 = \frac{1}{2} \sum_{j=1}^M (a_j^2 + b_j^2).$$

и при $j=1, 2, \dots, M$

$$\frac{1}{2} (a_j^2 + b_j^2) = \frac{4\pi}{T} \hat{s}(\omega_j)$$

Доказательство данных теорем можно найти в J. Hamilton Time series analysis, Princeton, Princeton Univ. Press, 1994, Ch. 6.

Определение тех j , при которых $\hat{s}(\omega_j)$ велико, позволяет выявлять циклические и сезонные эффекты (и преобразовывать данные, чтобы эти эффекты устранить)

Можно показать (но это не просто), что выборочная периодограмма не является состоятельной оценкой для спектральной плотности. Это является недостатком данного подхода.

Преимущества данного подхода проявляются, когда временной ряд хорошо аппроксимируется линейной комбинацией небольшого числа синусов и косинусов.