МЕХАНИКО-
МАТЕМАТИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

ФИНАНСОВО-ЭКОНОМИЧЕСКОЕ УПРАВЛЕНИЕ

НОСОВСКИЙ
ГЛЕБ ВЛАДИМИРОВИЧ

МЕХМАТ МГУ

КОНСПЕКТ ПОДГОТОВЛЕН СТУДЕНТАМИ, НЕ ПРОХОДИЛ ПРОФ. РЕДАКТУРУ И МОЖЕТ СОДЕРЖАТЬ ОШИБКИ. СЛЕДИТЕ ЗА ОБНОВЛЕНИЯМИ НА VK.COM/TEACHINMSU.

ЕСЛИ ВЫ ОБНАРУЖИЛИ ОШИБКИ ИЛИ ОПЕЧАТКИ, ТО СООБЩИТЕ ОБ ЭТОМ, НАПИСАВ СООБЩЕСТВУ VK.COM/TEACHINMSU.

БЛАГОДАРИМ ЗА ПОДГОТОВКУ КОНСПЕКТА
СТУДЕНТКУ ФАКУЛЬТЕТА ВМК МГУ
НЕДОЛИВКО ЮЛИЮ НИКОЛАЕВНУ

Содержание

Лекция 1 **4**
Введение . . . 4
Рынок ценных бумаг . . . 5
Эффективные рынки . . . 6
Доходность . . . 7
Требования к модели . . . 7

Лекция 2 **8**
Статистические свойства моделей доходности . . . 8
Классическая теория прогнозирования . . . 10
Дискретный белый шум . . . 10
Процесс авторегрессии $\mathsf{AR}(p)$ . . . 10
Теорема о разложении устойчивого процесса авторегрессии $\mathsf{AR}(p)$ . . . 11

Лекция 3 **14**
Доказательство теоремы о разложении устойчивого процесса авторегрессии $\mathsf{AR}(p)$ . . . 14
Следствия из теоремы . . . 18

Лекция 4 **19**
Уравнения для коэффицентов разложения $\{\delta_i\}$ . . . 19
Примеры . . . 22
Ковариационная функция процессов $\mathsf{AR}(p)$ . . . 23
Примеры . . . 24
Прогнозирование в модели $\mathsf{AR}(p)$ . . . 25

Лекция 5 **26**
Прогнозирование на $s+1$ шаг в модели $\mathsf{AR}(p)$ . . . 26
Модель скользящего среднего $\mathsf{MA}(q)$ . . . 26
Примеры . . . 30

Лекция 6 **32**
Процессы $\mathsf{ARMA}(p,q)$ . . . 32
Теорема о разложении устойчивых процессов $\mathsf{ARMA}(p,q)$ . . . 34
Нестационарные модели . . . 36

Лекция 7 **38**
Примеры . . . 38
Процесс $\mathsf{ARIMA}(p,d,q)$ . . . 39
Недетерминированные тренды в ARIMA . . . 41
Прогнозирование в ARIMA-моделях . . . 43

Лекция 8 **45**
Примеры прогнозирования в модели ARIMA . . . 45

Среднеквадратическая ошибка оптимального в среднеквадратическом прогноза в модели ARIMA . . . 46
Примеры . . . 48
Поведение корреляционной функции в модели ARIMA . . . 49
Дробно-интегрированная модель ARFIMA(p, d, q) . . . 50

Лекция 9 **51**
Применение модели ARIMA для процессов доходности . . . 51
Нелинейность процессов доходности . . . 51

Лекция 10 **57**
Нелинейность процессов доходности . . . 57
Процессы класса ARCH(p) . . . 59
Примеры . . . 60

Лекция 11 **62**
ARCH(1)-модель с условным гауссовским распределением . . . 62
Процессы GARCH(p, q) . . . 66

Лекция 12 **68**
Теорема о достаточных условиях стационарности . . . 68
Эксцесс процесса типа GARCH(p, q) . . . 71
Прогнозирование доходности . . . 72

Лекция 13 **73**
Прогнозирование доходности . . . 73

Лекция 1

Введение

Курс посвящен дискретным моделям финансовой математики, которые предназначены для прогнозирования доходности ценных бумаг.

Впервые теория веротностей в финансах была использована в 1900 г в диссертации «Theorie de la speculation», L. Bachilier. В этой диссертации впервые была введена концепция *винеровского процесса*, строго введенного в 1923 г. Н. Винером.[1]

Бушелье обратил внимание, что цены акций ведут себя очень нерегулярно.

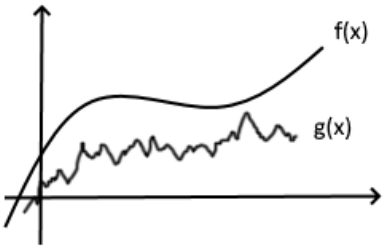

Рис. 1.1. f(x) – гладкая функция, g(c) – не диффирициируемая ни в какой точке функция

Помним, что гладкие функции (рис.1.1, функция $f(x)$) имеют ведут себя как

$$\Delta f \sim \Delta t.$$

Бушелье предложил модель доходностей, для которой

$$\Delta S_t \sim \sqrt{\Delta t}.$$

Такая функция (рис. 1.1, функция $g(x)$) не дифференциируема ни в какой точке. Действительно, предел

$$\lim_{t\to 0} \frac{\Delta S_t}{\Delta t} \to \infty.$$

Такие функции – траектории *броуновского движения.*

На современном языке, предложенная Бушелье модель имеет вид

$$S_t = S_0 + \mu t + \sigma W_t, \tag{1}$$

где S_t – цена акции в момент времени t, W_t – винеровский процесс, μ – коэффицент сноса, σ – коэффицент диффузии.[2]

[1] Долгое время считалось, что впервые понятие винеровского процесса на интуитивном уровне было введено в 1905 г. А. Энштейном.

[2] В финансовой математике σ принято называть *волатильностью*, то есть изменчивостью.

Модель, предложенная Бушелье, была забыта до 1965 г. В этом году Самуэльсон обратил внимание, что для некоторых t уравнение (1) дает отрицательные значения для цены акций, что некорректно. Так, он предложил модификацию этого уравнения (так называемое *уравнение геометрического броуновского движения*):

$$d(\ln S_t) = b \cdot dt + \sigma \cdot dW_t,$$

тогда его решение будет иметь вид

$$S_t = S_0 e^{bt} e^{\sigma W_t}.$$

Или, по-другому,

$$dS_t = S_t(\nu \cdot dt + \sigma \cdot dW_T).$$

тогда

$$S_t = S_0 e^{bt} e^{\sigma W_t - \sigma^2 \frac{t}{2}}.$$

До 1965 г. математика для построения финансовых моделей не применялась, так как приращение цены имело случайное блуждание.

Рынок ценных бумаг

Ценные бумаги бывают следующих видов.

1. *Облигация* (bond) – долговое обязательство, выданное под некоторый процент крупным учереждением (например, государством, центральным банком, банком), которое с математической точки зрения считается абсолютно надежным.

С точки зрения математической модели считается, что облигация *одна*, так как все остальные можно выразить через нее. Цена облигации меняется по закону

$$B_t = B_0 e^{rt},$$

где r – процентная ставка.

2. *Акция* (stock) – паевая ценная бумага, которая обеспечивает некую долю в неком предприятии, вообще говоря, рискованном.

С точки зрения математики, акции – негладкие функции (то есть как в модели Бушелье). Как говорили ранее,

$$d(\ln S_t^i) = b^i \cdot dt + \sigma^i \cdot dW_t, \quad 1 \le i \le n.$$

Различные цены акций в момент времени t

$$S_t^1, \quad S_t^2, \quad \ldots, \quad S_t^n.$$

Математическая модель рынка, иногда называемая (B, S)-рынок, выгядит так

$$(B_t, \quad S_t^1, \quad S_t^2, \quad \ldots, \quad S_t^n).$$

Эти два вида ценных бумаг не зависят от конкретной сделки.

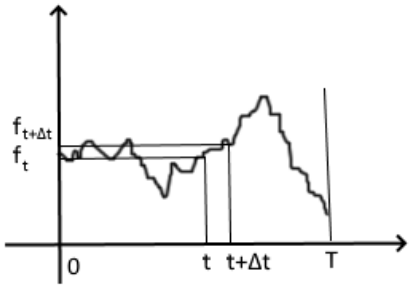

Рис. 1.2. Иллюстрация к определению доходности

3. *Фьючерс* (futures) – двустороннее торговое обязательство, обязывающее одну сторону купить, а вторую продать определенный товар в определенном количестве в определенное время.

Фьючерсы распространены, так как снижают риски, связанные с производством. Человек, приобретший фьючерс, не обязательно что-то покупает или продает, а, возможно, спекулирует.

Разница

$$f_{t+\Delta t} - f_t$$

называется *доходностью* ценной бумаги (рис. 1.2).

4. *Опцион* (option) – одностороннее ценовое обязательство, дающее своему владельцу право купить или продать определенный товар в определенном количестве в определенное время. Своего издателя опцион обязывает эту сделку обеспечить, то есть выплатить разницу, если сделка невыгодна.

Опционы делятся на опционы *купли* (call) и *продажи* (put). Другое деление – *американские* (дает совершить сделку до момента времени T) и *европейские* (дает совершить сделку только в определенный момент T) опционы.

В данном курсе рассматривается задача прогнозирования опционов. За рамки курса выходит задача поддержки, или хэджирования, опциона.

Эта задача имеет решение в рамках математической модели.

Эффективные рынки

Вообще говоря, это не математическое понятие, но понятие эффективного рынка используется в финансовой математике.

Современная финансовая математика делает прогнозы, основываясь на истории цен. *Рынок называется эффективным, если оптимальное поведение на этом рынке зависит только от цен, а другая информация, например, объем или какая-то закрытая информация значения не имеют.*

Доходность

Поведение цен достаточно сложно, так как это нестационарные процессы, поэтому обычно прогнозируют поведение не самих цен, а доходностей.

В рамках курсы будем рассматривать модель, в которой время дискретно, а именно *измеряется в днях*. Цена и доходность фиксируется в момент закрытия биржи.

Обозначим через X_t доходность ценной бумаги в день t. Самый простой способ определить доходность:

$$x_t^* = Z_t - Z_{t-1} + dt,$$

где Z_t – цена ценной бумаги в день t, а dt – дивидент (выплачиваются не каждый день, поэтому обычно равны 0).

Этим определением не пользуются, так как у ценных бумаг цены могут быть в разной валюте. Поэтому используется формула

$$X_t = \ln(Z_t + dt) - \ln(Z_{t-1}).$$

Из разложения логарифма видно, что

$$X_t \sim \frac{Z_t + dt - Z_{t-1}}{Z_{t-1}}.$$

Требования к модели

1. Модель должна приводить к реализациям, «похожим» на реальные процессы, которые моделировались.

2. Гипотезы, лежащие в основе модели, должны допускать проверку методами математической статистики.

3. Модель должна быть по возможности простой.

4. Модель должна предоставлять способ оптимального оценивания доходности при условии, что модель верна.

5. Модель должна быть полезна на практике, то есть применима.

Пункты 1-4 являются обязательными требованиям. Пункт 5 не всегда выполняется в том смысле, в котором применяются математические модели в других областях.

Лекция 2

Статистические свойства моделей доходности

Рассмотрим, какие у процессов доходностей есть статистические свойства, прежде чем займемся их моделированием с целью построения оптимальных прогнозов.

Будем обозначать

$$X_t \text{ – доходность в день } t.$$

Реализация процесса доходности

$$X_t: \ X_1, \quad X_2, \quad X_3, \quad \ldots, \quad X_n, \quad \ldots$$

За время n у нас возникают следующие статистики:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

– выборочное среднее,

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

– выборочная дисперсия,

$$b = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(X_i - \overline{X})^3}{\sigma^3}$$

– выборочный коэффицент ассиметрии,

$$k = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(X_i - \overline{X})^4}{\sigma^4}$$

– выборочный коэффицент эксцесса.

Кроме того, у нас есть соответствующие генеральные характеристики случайной величины ξ:

$$B = B_\xi = \frac{\mathbb{E}(\xi - \mathbb{E}\xi)^3}{\left(\mathbb{E}(\xi - \mathbb{E}\xi)^2\right)^{3/2}}$$

– коэффицент ассиметрии[3] случайной величины ξ,

$$K = K_\xi = \frac{\mathbb{E}(\xi - \mathbb{E}\xi)^4}{\left(\mathbb{E}(\xi - \mathbb{E}\xi)^2\right)^2}$$

[3]Коэффицент ассиметрии используется, чтобы оценить, насколько распределение ассиметрично. Если распределение с.в. ξ симметрично, то

$$B_\xi = 0.$$

– коэффицент эксцесса[4] случайной величины ξ.

Посмотрим, какие значения для этих характеристик наблюдаются в реальных процессах доходности.

Все статистические оценкци коэффицента ассиметрии процессов доходности показывают[5], что

$$b \sim 0.$$

Поэтому в математической модели будем считать, что

$$B = 0.$$

Во всех случаях коэффицент эксцесса процессов доходности

$$k > 3.5,$$

и часто выполнено, что

$$k > 6.$$

Из этого делаем вывод, что у распределение доходности тяжелые хвосты. Поэтому не рассматриваем модели, использующие нормальное распределение.

Остановимся подробнее на одной сложности, возникающей при моделировании доходности, а именно *эффекте понедельника*.

В любой другой день, кроме понедельника, доходность расчитывается как разность между двумя соседними днями, а в понедельник – как разница между ценой в понедельник и ценой в пятницу. Другие процессы, как, например, дешевеют деньги или меняется стоимость облигаций, изменяются по календарным дням, а не по рабочим, поэтому доходность в понедельник немного отличается.

Рассмотрим, как ведет себя корреляционная функция процесса доходности.

Вспомним, что для случайного процесса ξ_t корреляционная функция имеет вид

$$R_{t,\tau} = \mathsf{corr}(\xi_t, \xi_{t+\tau}).$$

Если ξ_t – стационарный процесс, то

$$R_{t,\tau} = R_\tau,$$

то есть не зависит от t.

В случае стационарны процессов для оценки R_τ используется выборочный коэффицент корреляции процесса X_t:

$$r_\tau(X) = \frac{\sum_{i=1}^{n-\tau}(X_i - \overline{X})(X_{i+\tau} - \overline{X})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\frac{n}{n-\tau}.$$

[4]Коэффицент эксцесса используется, чтобы оценить отклонение распределения от «нормальности». Если ξ имеет распределение $\mathcal{N}(\mu, a^2)$, то

$$K_\xi = 3.$$

Если $K_\xi > 3$, говорят, что у распределения с.в. ξ тяжелые хвосты. Если $K_\xi < 3$, у распределения с.в. ξ легкие хвосты.

[5]Для некоторых редких ценных бумаг распределение может быть немного ассиметрично.

Статистика процессов доходности:

$$r_\tau(X) \sim 0, \quad \forall \tau \geq 1,$$

то есть доходности некоррелированы. При этом, например,

$$r_\tau(|X|) \neq 0, \quad r_\tau(X^2) \neq 0.$$

Это означает, что доходности некоррелированы, но зависимы.

Кроме того, при $1 < \tau < 30$ большинство значений ≥ 0.05 и

$$\sum_{i=1}^{30} r_i(X^2) = \begin{cases} 0.48 - 0.49 \text{ для акций,} \\ 1.4 - 6.3 \text{ для остальных ценных бумаг.} \end{cases}$$

Классическая теория прогнозирования

Рассмотрим сначала классическую теорию прогнозирования, так как современные модели являются некой надстройкой над ней. Затем обсудим, что в этих моделях неприменимо.

Дискретный белый шум

Забудем на время об обозначениях, введенных ранее для процессов доходности.

Определение 2.1. Случайная последовательность $(X_i)_{i\geq 0}$ называется *дискретным белым шумом в широком смысле*, если

$$\mathbb{E}X_i = 0, \quad \forall i, \quad \mathbb{E}X_iX_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Здесь X_t – стационарный в широком смысле процесс.[6]

Определение 2.2. Случайная последовательность $(X_i)_{i\geq 0}$ называется *дискретным белым шумом в узком смысле*, если процесс X_i стационарен в узком смысле,

$$\mathbb{E}X_i = 0, \quad \forall i, \quad \mathbb{E}X_i^2 = 1,$$

и X_i и X_j независимы, $i \neq j$.

Процесс авторегрессии $\mathsf{AR}(p)$

Определение 2.3. Случайная последовательность y_t, $t \in \mathbb{N}$ (или $t \in \mathbb{Z}$), называется *процессом авторегрессии порядка* p, $p \in \mathbb{N}$, и обозначается

$$y_t \in \mathsf{AR}(p),$$

[6] То есть процесс, у которого ковариация зависит только от горизонта, а дисперсия и матожидание постоянны.

если $\exists\ \beta_1,\ \beta_2, \ldots, \beta_p \in \mathbb{R}$, $\beta_p \neq 0$, такие, что

$$y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} = u_t, \tag{2}$$

где u_t – дискретный белый шум[7].

Уравнение (2) называется *уравнением процесса* y_t.

Для удобства будем считать, что $t \in \mathbb{Z}$. Обозначим $\beta_0 = 1$. Тогда (2) можно записать в виде

$$\sum_{i=0}^{p} \beta_i y_{t-i} = u_t, \quad t \in \mathbb{Z}. \tag{3}$$

Заметим, что можно записать (2) в виде

$$y_t = -\beta_1 y_{t-1} - \beta_2 y_{t-2} - \ldots - \beta_p y_{t-p} = u_t.$$

Тогда, если бы u_t был равен 0, значение процесса y_t в момент времени t было бы линейной комбинайцией значений в моменты времени $t-1, \ldots, t-p$.

Так как u_t приносит новую стохастическую информацию в процесс y_t, он называется *обновляющим процессом* или *возмущением*.

Определение 2.4. Алгебраическое уравнение

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0 \tag{4}$$

называется *характеристическим уравнением процесса* y_t*, удовлетворяющим уравнению процесса (2).*[8]

Теорема о разложении устойчивого процесса авторегрессии AR(p)

Определение 2.5. Процесс y_t называется *устойчивым*, если все корни характеристического уравнения (4) по модулю меньше 1.

В доказательстве следующей важной теоремы используются основные идеи изучаемой нами науки.

Теорема 2.1. *(о разложении учтойчивого процесса авторегрессии) Если все корни характеристического уравнения (4) по модулю меньше* 1*, то процесс* y_t *представим в виде*[9]

$$y_t \overset{ср.кв.}{=} \sum_{i=0}^{\infty} \delta_i u_{t-i},$$

[7] В узком или широком смысле, будет оговаривать отдельно.

[8] Как мы знаем, алгебраическое уравнение степени p (4) имеет на поле комлпексных чисел p корней.

[9] Другими словами, если все корни характеристического уравнения по модулю меньше 1, процесс авторегрессии представим в виде бесконечного скользящего среднего по значениям процесса белого шума с некоторыми фиксированными коэффицентами.

где $\{\delta_i\}_{i=0}^{\infty}$, $\delta_i \in \mathbb{R}$ *– вещественные коэффиценты, или*

$$\lim_{n\to\infty} \mathbb{E}\left(y_t - \sum_{i=0}^{n} \delta_i u_{t-i}\right)^2 = 0.$$

Доказательство Сформулироем сначала некоторое техническое предположение об ограниченности вторых моментов на $-\infty$.[10]

Будем считать, что $\exists\ N$: $\forall\ n < -N$ выполнено

$$\mathbb{E}y_n^2 < N.$$

- Пусть $p = 1$, $y_t \in \mathsf{AR}(1)$. Запишем уравнение процесса в виде

$$y_t = \rho y_{t-1} + u_t, \quad \rho = -\beta_1, \tag{5}$$

$$y_{t-1} = \rho y_{t-2} + u_{t-1}, \quad \rho = -\beta_1. \tag{6}$$

Теперь подставим (6) в (5). Получим

$$y_t = u_t + \rho u_{t-1} + \rho^2 y_{t-2}.$$

Повторяя эту процедуру, на s-том шаге получим

$$y_t = u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \ldots + \rho^s u_{t-s} + \rho^{s+1} y_{t-(s+1)}.$$

Отсюда получаем, что

$$y_t - \left(u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \ldots + \rho^s u_{t-s}\right) = \rho^{s+1} y_{t-(s+1)}. \tag{7}$$

Возведем (7) в квадрат и возьмем матождание от обеих частей:

$$\mathbb{E}\left(y_t - \sum_{i=0}^{s} \rho^i u_{t-i}\right)^2 = \rho^{2(s+1)} \mathbb{E} y_{t-(s+1)}^2.$$

Запишем характеристическое уравнение (4):

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0,$$

причем $\beta_0 = 1$, $\beta_1 = -\rho$, $p = 1$. Тогда уравнение имеет вид

$$x - \rho = 0,$$

а значит,

$$x = \rho \ \Rightarrow\ |\rho| < 1.$$

[10] Заметим, что, конечно, в реальных процессах время можно брать очень большим отрицательным, но оно не может быть равно $-\infty$. В математической теории для доказательства t может принимать любые значения, поэтому важно, чтобы в $-\infty$ не было никаких проблем.

Тогда

$$\lim_{s\to\infty} \rho^{2(s+1)} \mathbb{E} y_{t-(s+1)}^2 = 0,$$

так как второй момент на бесконечности по предположению ограничен сверху, а $\rho^{2(s+1)} \to 0$, $s \to \infty$. Значит,

$$y_t \overset{\text{ср.кв.}}{=} \sum_{i=0}^{\infty} \rho_i u_{t-i}, \quad \delta_i = \rho^i.$$

Теорема доказана для случая $p = 1$.

- Общий случай $p \geq 1$.

  Аналогично предыдущему случаю, запишем

$$y_t = u_t - \beta_1 y_{t-1} - \ldots - \beta_p y_{t-p},$$

$$y_{t-1} = u_{t-1} - \beta_1 y_{t-2} - \ldots - \beta_p y_{t-p-1}$$

  и подставим второе из уравнений в первое. Получим

$$y_t = u_t - \beta_1 \left(u_{t-1} - \beta_1 y_{t-2} - \ldots - \beta_p y_{t-p-1}\right) -$$

$$-\beta_2 y_{t-2} - \ldots - \beta_p y_{t-p}.$$

  Запишем это в виде

$$y_t = u_t - \beta_1 u_{t-1} - \left(\beta_2 - \beta_1^2\right) y_{t-2} - \left(\beta_3 - \beta_1\beta_2\right) y_{t-3} - \ldots -$$

$$- \left(\beta_p - \beta_1\beta_{p-1}\right) y_{t-p} + \beta_1\beta_p y_{t-p-1}. \tag{8}$$

  Запишем теперь уравнение процесса в точке $t-2$:

$$y_{t-2} = u_{t-2} - \beta_1 y_{t-3} - \ldots - \beta_p y_{t-2-p},$$

  подставляем его в (8). Через s аналогичных шагов получим

$$y_t = \underbrace{u_t + \delta_1^* u_{t-1} + \ldots + \delta_s^* u_{t-s}}_{s+1 \text{ слагаемых}} + \underbrace{a_{s1}^* y_{t-s-1} + \ldots + a_{sp}^* y_{t-s-p}}_{p \text{ слагаемых}}.$$

  Коэффиценты δ_i^* и a_{sk}^* зависят от коэффицентов β_j.

  Чтобы доказать теорему, достаточно показать, что

$$a_{si}^* \to 0, \quad s \to \infty, \quad \forall\, 1 \leq i \leq p.$$

  Покажем это на следующей лекции.

Лекция 3

Доказательство теоремы о разложении устойчивого процесса авторегрессии $\mathsf{AR}(p)$

Продолжим доказательство теоремы 2.1. В прошлый раз закончили на том, что записали

$$y_t = u_t + \delta_1^* u_{t-1} + \ldots + \delta_s^* u_{t-s} + a_{s1}^* y_{t-s-1} + \ldots + a_{sp}^* y_{t-s-p}. \tag{9}$$

Попробуем с помощью некоторых соображений показать, что коэффиценты a_{si}^* стремятся к 0 при $s \to \infty$.

Подставим сначала соотношение

$$y_{t-s-1} = u_{t-s=1} - \beta_1 y_{t-s-2} - \ldots - \beta_p y_{t-s-p-1}$$

в уравнение (9). Получим

$$y_t = u_t + \delta_1^* u_{t-1} + \delta_s^* u_{t-s} + a_{s1}^* u_{t-s-1} + \left(a_{s2}^* - a_{s1}^*\right) y_{t-s-2} + \ldots +$$
$$+ \left(a_{sp}^* - a_{s1}^* \beta_{p-1}\right) y_{t-s-p} - a_{s1}^* \beta_p y_{t-s-p-1}. \tag{10}$$

Соотношение (10) – то же самое, что (9), записанное не для s шагов, а для $s-1$. Воспользовавшись этими соображениями, получим

$$\begin{cases} \delta_{s+1}^* \ = \ a_{s1}^*, \\ s+1, j^* = a_{s,j+1}^* - a_{s1}^* \beta_j, \quad 1 \le j \le p-1, \\ a_{s+1,p}^* = -a_{s1}^* \beta_p. \end{cases} \tag{11}$$

Сформулируем промежуточную лемму, доказательство которой будет завершать доказательство теоремы 2.1.

Лемма 3.1.

$$y_t \overset{ср.кв.}{=} \sum_{i=0}^{\infty} \delta_i^* u_{t-i}, \quad \delta_0 = 1.$$

Доказательство Доказательство будет заключаться в том, что получим коэффицентыдругим способом, который дает возможность доказать их стремление к 0. Получив для них соотношение (11) , докажем, что это одни и те же коэффиценты.

Определим оператор L (оператор *сдвига*) на последовательности ξ_t, $t \in Z$:

$$L\xi_t = \xi_{t-1}.$$

С помощью оператора L уравнение процесса $y_t \in \mathsf{AR}(p)$ можно записать в виде

$$\sum_{i=0}^{p} \beta_i L^i y_t = u_t, \quad \forall\, t \in \mathbb{Z}.$$

Это соотношение можем записать в виде

$$\underbrace{\left(\sum_{i=0}^{p}\beta_i L^i\right)}_{\text{оператор}} y_t = u_t, \quad \forall\, t \in \mathbb{Z}.$$

Не задумываясь пока, существует ли обратный оператор, формально запишем

$$y_t = \left(\sum_{i=0}^{p}\beta_i L^i\right)^{-1} u_t.$$

Будем формально искать оператор $\left(\sum_{i=0}^{p}\beta_i L^i\right)^{-1}$ в виде

$$\left(\sum_{i=0}^{p}\beta_i L^i\right)^{-1} = \sum_{i=0}^{\infty}\delta_i L^i.$$

Определим $\{\delta_i\}_{i=0}^{\infty}$ из соотношения

$$\left(\sum_{i=0}^{p}\beta_i z^i\right)^{-1} = \sum_{i=0}^{\infty}\delta_i z^i, \quad z \in \mathbb{C}.$$

Имеем

$$\frac{1}{1+\beta_1 z+\ldots+\beta_p z^p} = 1 - \frac{\beta_1 z+\ldots+\beta_p z^p}{1+\beta_1 z+\ldots+\beta_p z^p} =$$

$$= 1 - b_1 z - \frac{\left(\beta_2-\beta_1^2\right)z^2+\ldots+\left(\beta_p-\beta_1\beta_{p-1}\right)z^p-\beta_1\beta_p z^{p+1}}{1+\beta_1 z+\ldots+\beta_p z^p}.$$

Продолжая деление, получим некоторое разложение вида

$$\frac{1}{1+\beta_1 z+\ldots+\beta_p z^p} = 1+\delta_1 z+\delta_2 z^2+\ldots+\delta_s z^s + \frac{a_{s1}z^{s+1}+\ldots+a_{sp}z^{s+p}}{1+\beta_1 z+\ldots+\beta_p z^p}. \tag{12}$$

Сделав еще один (явный) шаг деления, наконец получим

$$\frac{1}{1+\beta_1 z+\ldots+\beta_p z^p} = 1+\delta_1 z+\delta_2 z^2+\ldots+a_{s1}z^{s+1}+$$

$$+\frac{\left(a_{s2}-a_{s1}\beta_1 z^{s+2}\right)+\ldots+\left(a_{sp}-a_{s1}\beta_{p-1}z^{s+p}\right)-a_{s1}\beta_p z^{s+p+1}}{1+\beta_1 z+\ldots+\beta_p z^p}$$

Соотнеся полученное соотношение с видом из (12) (для случая, когда вместо s шагов проделаны $s+1$), получим, что

$$1+\delta_1 z+\delta_2 z^2+\ldots+a_{s1}z^{s+1}+\frac{\left(a_{s2}-a_{s1}\beta_1 z^{s+2}\right)+\ldots+\left(a_{sp}-a_{s1}\beta_{p-1}z^{s+p}\right)-a_{s1}\beta_p z^{s+p+1}}{1+\beta_1 z+\ldots+\beta_p z^p} =$$

$$= 1+\delta_1 z+\delta_2 z^2+\ldots+\delta_{s+1}z^{s+1}+\frac{a_{s+1,1}z^{s+2}+\ldots+a_{s+1,p}z^{s+1+p}}{1+\beta_1 z+\ldots+\beta_p z^p}.$$

Следовательно, у нас возникают реккурентные соотношения

$$\begin{cases} \delta_{s+1} \ = \ a_{s1}, \\ s+1, j = a_{s,j+1} - a_{s1}\beta_j, \quad 1 \le j \le p-1, \\ a_{s+1,p} = -a_{s1}\beta_p. \end{cases} \tag{13}$$

Эти соотношения такие же, как и (11). Можно проверить, что начальные значения тоже совпадают и равны

$$\begin{cases} \delta_1 = \delta_1^* = -\beta_1, \\ a_{1j} = a_{1j}^* = -\left(\beta_{j+1} - \beta_1^{j+1}\beta_j\right), \quad 1 \le j \le p-1, \\ a_{1p} = \beta_1\beta_p. \end{cases}$$

Следовательно, достаточно доказать следующее соотношение, сформлуированное в виде леммы.

Лемма 3.2.

$$a_{sj} \to 0, \quad s \to \infty, \quad \forall\ j \in \{1, \dots, p\}.$$

Доказательство Обозначим корни характеристического уравнения

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0 \tag{14}$$

через $x_1, \dots, x_p$. По условию имеем

$$|x_i| < 1, \quad \forall\ 1 \le i \le p.$$

Кроме этого, рассмотрим двойственное к (14) уравнение

$$\sum_{i=0}^{p} \beta_i z^i = 0. \tag{15}$$

Обозначим его корни через $z_1, \dots, z_p$.

Лемма 3.3. *Если $\beta_0 \neq 0$, $\beta_p \neq 0$, то множество корней*

$$\{z_1, \dots, z_p\}$$

связано с множеством

$$\{x_1, \dots, x_p\}$$

следующим образом:

$$x \in \{x_1, \dots, x_p\} \ \Rightarrow \ \frac{1}{x} \in \{z_1, \dots, z_p\},$$

$$z \in \{z_1, \dots, z_p\} \ \Rightarrow \ \frac{1}{z} \in \{x_1, \dots, x_p\}.$$

Доказательство. Пусть
$$x : * \in \{x_1, \ldots, x_p\}.$$
Так как $\beta_p \neq 0$, $x \neq 0$. Тогда
$$\beta_0 x^{*p} + \ldots \beta_{p-1} x^* + \beta_p = 0$$
можно записать в виде
$$x^{*p}\left(\beta_0 + \beta_1 x^{*-1} + \ldots + \beta_p x^{*-p}\right) = 0.$$
Значит,
$$\beta_0 + \beta_1 x^{*-1} + \ldots + \beta_p x^{*-p} = 0,$$
что то же самое, что
$$\beta_0 + \beta_1 \left(\frac{1}{x^*}\right)^1 + \ldots + \beta_p \left(\frac{1}{x^*}\right)^p = 0.$$
Отсюда получаем, что
$$\frac{1}{x^*} \in \{z_1, \ldots, z_p\}.$$
В другую сторону доказывается аналогично.

Лемма доказана.

Заметим, что все корни уравнения (15):
$$|z_i| > 1, \quad 1 \leq i \leq p,$$
так как все корни (14) по модулю меньше единицы.

Рассмотрим круг
$$|z| < \underbrace{\min_{1 \leq i \leq p} |z_p|}_{>1} = \alpha. \tag{16}$$
В этом открытом круге абсолютно сходится ряд
$$\frac{1}{\sum_{i=0}^p \beta_i z^i} = \frac{1}{\prod_{i=1}^p \left(1 - \frac{z}{z_i}\right)} = \prod_{i=1}^p \sum_{j=0}^\infty \left(\frac{z}{z_j}\right)^j.$$
Последнее равенство следует из формулы для суммы бесконечной геометрической прогрессии. Заметим, что из-за ограничений (16) $z/z_j < 1$.

Тогда верно
$$\prod_{i=1}^p \sum_{j=0}^\infty \left(\frac{z}{z_j}\right)^j = \sum_{i=0}^\infty \delta_i z^i.$$
Значит,
$$\frac{1}{\sum_{i=0}^p \beta_i z^i} - (1 + \ldots + \delta_s z^s) = \frac{a_{s1} z^{s+1} + \ldots + a_{sp} z^{s+p}}{1 + \beta_1 z + \ldots + \beta_p z^p} \to 0, \quad s \to \infty, \quad \forall z : |z| < \alpha.$$

Так как знаменатель дроби из соотношения выше ограничен, это означает, что

$$f(z) = a_{s1}z^{s+1} + \ldots + a_{sp}z^{s+p} \to 0, \quad s \to \infty, \quad \forall z : |z| < \alpha, \quad \alpha > 1.$$

По теореме Коши

$$f'(z) = \frac{1}{2\pi i}\int_{C:|z|=1} \frac{f(\omega)}{(\omega - z)^2} d\omega,$$

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_{C:|z|=1} \frac{f(\omega)}{(\omega - z)^{n+1}} d\omega.$$

Тогда

$$a_{si} = \frac{1}{2\pi i}\int_{|\omega|=1} \frac{f(\omega)}{(\omega - a)^{s+1+i}} d\omega \le \sup_{|\omega|\le 1} f(\omega)\frac{2\pi}{2\pi},$$

если взять $a = 0$.

Так как $|f(\omega)| \to 0$, то

$$a_{si} \to 0, \quad s \to \infty, \quad \forall\, 1 \le i \le p.$$

И, наконец, получаем, что

$$\mathbb{E}\left(y_t - \sum_{i=0}^{n} \delta_i u_{t-i}\right)^2 = \mathbb{E}(a_{n1}y_{t-n-1} + \ldots + a_{np}y_{t-n-p})^2 \to 0, \quad n \to \infty,$$

то есть

$$y_t \overset{\text{ср.кв.}}{=} \sum_{i=0}^{\infty} \delta_i u_{t-i} = \sum_{i=0}^{\infty} \delta_i^* u_{t-i}.$$

Теорема доказана.

Следствия из теоремы

Следствие 1. *Если u_t – белый шум в узком смысле, то в условиях теоремы 2.1 y_t не зависит от u_{t+i}, $i \ge 1$.*

Следствие 2. *Если u_t – белый шум в широком смысле, то в условиях теоремы 2.1 y_t не коррелирован с u_{t+i}, $i \ge 1$.*

Следствие 3. *В условиях теоремы 2.1 y_t – стацинарный процесс (в узком или широком смысле в зависимости от того, в узком или широком смысле u_t – белый шум).*

Лекция 4

Уравнения для коэффицентов разложения $\{\delta_i\}$

Для любого z имеем

$$1 = \left(\sum_{r=0}^{p} \beta_r z^r\right)^{-1} \sum_{s=0}^{p} \beta_s z^{p-s} = \sum_{r=0}^{\infty} \delta_r z^r \sum_{r=0}^{p} \beta_s z^r = \sum_{r=0}^{\infty} \sum_{s=0}^{p} \beta_s \delta_r z^{r+s}.$$

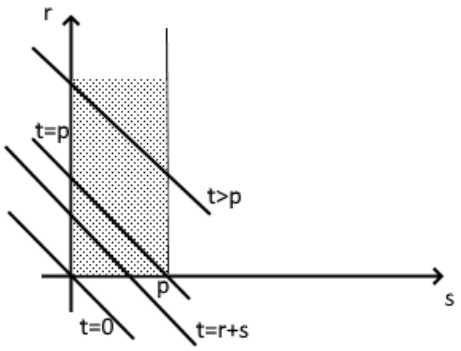

Рис. 4.1. Замена пределов суммирования $t = r + s$.

Сделаем замену переменных. Обозначим $t = r + s$, тогда $r = t - s$. Тогда (рис. 4.1) получим

$$\sum_{r=0}^{\infty} \sum_{s=0}^{p} \beta_s \delta_r z^{r+s} = \sum_{e=0}^{p-1} \left(\sum_{s=0}^{t} \beta_s \delta_{t-s}\right) z^t + \sum_{t=p}^{\infty} \left(\sum_{s=0}^{p} \beta_s \delta_{t-s}\right) z^t \equiv 1,$$

тождество по z.

Следовательно,

$$\begin{cases} 1 = \beta_0 \delta_0 = \delta_0, \\ 0 = \beta_0 \delta_1 + \delta_1 \beta_0 = \delta_1 + \beta_1, \\ \ldots\ldots \\ 0 = \beta_0 \delta_{p-1} + \ldots + \beta_{p-1} \delta_0 \end{cases} \tag{17}$$

Для второй суммы можно условия записать в общем виде

$$0 = \beta_0 \delta_t + \ldots + \beta_p \delta_{t-p}, \quad t = p, p+1, p+2, \ldots \tag{18}$$

Уравнения вида (18) будем называть *однородным разностным уравнением на последовательность* $\{\delta_i\}$. Условия (17) являются его *граничными условиями.*

Докажем теорему о решении однородного разностного уравнения.

Теорема 4.1. *Если все корни $x_1, \ldots, x_p$ зарактористического уравнения*

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0$$

различны, то общее решение однородного разностного уравнения (18) имеет вид

$$\delta_r = \sum_{i=1}^{p} k_i x_i^r, \quad r = 0, 1, 2, \ldots \tag{19}$$

где k_i – неопределенные коэффиценты, которые находятся из граничных условий (17).

Доказательство

- Предположим, что (19) удовлетворяет (18). Пусть

$$\delta_r = \sum_{i=1}^{p} k_i x_i^r.$$

Подставляем в (18):

$$\beta_0 \left(\sum_{i=1}^{p} k_i x_i^t \right) + \ldots + \beta_p \left(\sum_{i=1}^{p} k_i x_i^{t-p} \right) = \sum_{i=1}^{p} k_i \left(\beta_0 x_i^t + \ldots + \beta_p x_i^{t-p} \right) = 0,$$

так как x_i – корень уравнения

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0.$$

Отсюда следует, что

$$\beta_0 x_i^t + \ldots + \beta_p x_i^{t-p} = x_i^{t-p} \left(\beta_0 x_i^p + \ldots + \beta_p \right) = 0.$$

Отсюда $x_i \neq 0$, так как $\beta_p \neq 0$.

- Покажем, что любое решение (18) имеет вид (19) при некоторых значениях коэффицентов k_i. Для этого выпишем уравнения на k_i из тех соображений, что выражение (19) удовлетворяет любым граничным условиям. Так как, если любые граничные условия могут быть удовлетворены выражениям (19) при некоторых k_i, то и любое решение (18) имеет вид (19), поскольку уравнение (18) рекуррентно определяет все δ_i через $\delta_0, \ldots, \delta_{p-1}$.

Напишем систему уравнений на k_i при заданных $\delta_0, \ldots, \delta_{p-1}$:

$$\begin{cases} \delta_0 = \sum_{i=0}^{p} k_i x_i^0 = \sum_{i=0}^{p} k_i, \\ \delta_1 = \sum_{i=0}^{p} k_i x_i, \\ \ldots\ldots \\ \delta_{p-1} = \sum_{i=0}^{p} k_i x_i^{p-1}. \end{cases}$$

Эту систему можно записать в виде

$$A \overrightarrow{k} = \overrightarrow{\delta}, \tag{20}$$

где

$$\overrightarrow{k} = (k_1, \ldots, k_p), \quad \overrightarrow{\delta} = (\delta_0, \ldots, \delta_{p-1}),$$

$$A = \begin{pmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_p \\ \vdots & \ddots & \vdots \\ x_1^{p-1} & \ldots & x_p^{p-1} \end{pmatrix}.$$

Уравнение (20) имеет решение при любой правой части тогда и только тогда, когда определитель матрицы A отличен от нуля.

Докажем, что определитель A не равен 0, от противного. Пусть $\det A = 0$. Тогда

$$\exists\ \mu_0, \ldots, \mu_{p-1} \in \mathbb{R}$$

нетривиальные и такие, что уравнение

$$\mu_0 + \mu_1 z + \ldots + \mu_{p-1} z^{p-1} = 0$$

удовлетворяется при

$$z = x_1, \ldots, z = x_p,$$

то есть имеет $\geq p$ корней. Так как это уравнение $p-1$ степени, по основной теореме алгебры получаем противоречие.

Теорема доказана.

Сформулируем теперь теорему для случая кратных корней характеристического уравнения.[11]

Теорема 4.2. *Если характеристическое уравнение*

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0$$

имеет корни

$$\underbrace{x_1, \ldots, x_m}_{\text{вещественные}}, \quad \underbrace{\alpha_1 e^{\pm i\theta_1}, \ldots, \alpha_n e^{\pm i\theta_n}}_{\text{пары компл.-сопряженных корней}}$$

с кратностями

$$s_1, \ldots, s_m, r_1, \ldots, r_n,$$

где

$$\sum_{i=1}^{m} s_i + 2\sum_{j=1}^{n} r_j = p,$$

то общее решение разностного уравнения (18) имеет вид

$$\delta_t = \sum_{i=1}^{m} P_i^{(s_i-1)}(t) x_i^t + \sum_{j=1}^{n} Q_j^{(r_j-1)}(t) \alpha_j^t e^{i\theta_j t} + \sum_{j=1}^{n} \overline{Q_j}^{(r_j-1)}(t) \alpha_j^t e^{-i\theta_j t},$$

[11] Эта теорема доказывается достаточно громоздно, ее доказательство нам сейчас не требуется.

где

$$\left\{P_i^{(s_i-1)}\right\}_{i=1}^m$$

– многочлены степени $(s_i - 1)$, $1 \le i \le m$, *с неопределенными вещественными коэффицентами, а*

$$\left\{Q_j^{(r_j-1)}\right\}_{j=1}^n$$

– многочлены степени $(r_j - 1)$, $1 \le j \le n$, *с неопределенными комлексными коэффицентами.*

Следствие 1. *Общее количество неопределенных вещественных коэффицентов равно*

$$\sum_{i=1}^m s_i + 2\sum_{j=1}^n r_j = p$$

и равно количеству уравнений в граничных условиях.

Примеры

- $p = 1$. Для этого случая уже знаем, что

  $$\delta_r = (-\beta_1)^r,$$

  характеристическое уравнение имеет вид

  $$x + \beta_1 = 0,$$

  то есть корень

  $$x = \beta_1.$$

- $p = 2$. Пусть корни характеристического уравнения x_1, x_2 и $x_1 \neq x_2$. Характеристическое уравнение

  $$x^2 + \beta_1 x + \beta_2 = 0.$$

  Чтобы найти неопределенные коэффиценты, запишем

  $$\begin{cases} k_1 + k_2 = 1, \\ k_1 x_1 + k_2 x_2 = -\beta_1 = (x_1 + x_2), \end{cases}$$

  где последнее равенство следует из теоремы Виета.

  Отсюда получаем, что

  $$k_1 = \frac{x_1}{x_1 - x_2}, \quad k_2 = -x_2 x_1 - x_2.$$

  Тогда

  $$\delta_r = \frac{x_1^{r+1} - x_2^{r+1}}{x_1 - x_2}. \tag{21}$$

Если x_1, x_2 – комплексные, то они комплексно сопряжены:

$$x_1 = \alpha e^{t\theta}, \quad x_2 = \alpha e^{-i\theta}.$$

Тогда

$$k_1 = \frac{e^{i\theta}}{e^{i\theta} - e^{-i\theta}}, \quad k_2 = \frac{-e^{-i\theta}}{e^{i\theta} - e^{-i\theta}}.$$

$$\delta_r = k_1 x_1^r + k_2 x_2^r = \alpha^r \frac{e^{i\theta(r+1)} - e^{-i\theta(r+1)}}{e^{i\theta} - e^{-i\theta}} = \alpha^r \frac{\sin\theta(r+1)}{\sin\theta}. \tag{22}$$

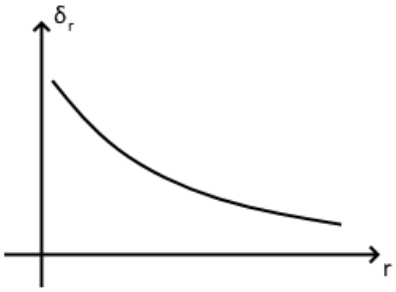

Рис. 4.2. δ_r в случае вещественных корней

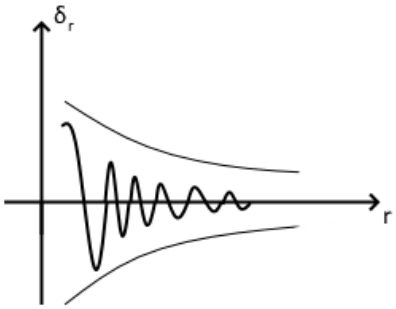

Рис. 4.3. δ_r в случае комплексно-сопряженных корней

Заметим, что функция δ_r в случае вещественных корней (21) ведет себя так, как показано на рис. 4.2, а в случае комплесно-сопряженных корней (22) – как на рис. 4.3.

Ковариационная функция процессов $\mathsf{AR}(p)$

Ковариационная функция имеет вид

$$\sigma(s) = \mathbb{E} y_t y_{t+s}, \quad y_t \in \mathsf{AR}(p).$$

Получим уравнения Юла-Уолкера на ковариационную функцию процесса $\mathsf{AR}(p)$. Запишем разложение процесса в точке $t - s$

$$y_{t-s} = \sum_{r=0}^{\infty} \delta_r u_{t-r-s}$$

и уравнение процесса в точке t

$$\sum_{r=0}^{p} \beta_r y_{t-r} = u_t.$$

Умножим эти два уравнения. Получим

$$\sum_{r=0}^{p} \beta_r y_{t-r} y_{t-s} = \sum_{q=0}^{\infty} \delta_q u_t u_{t-s-q}.$$

Возьмем матожидание от обеих частей, пользуясь тем, что

$$\mathbb{E} y_t y_{t+s} = \sigma(s), \quad E y_t^2 = \sigma^2.$$

Получим

$$\begin{cases} \sum_{r=0}^{p} \beta_r \sigma(-r) = \sigma^2, & s = 0, \\ \sum_{r=0}^{p} \beta_r \sigma(s-r) = 0, & s \geq 1. \end{cases} \tag{23}$$

Уравнения (23) и есть уравнения Юла-Уолкера для ковариационной функции. На самом деле, (23) – однородное разностное уравнение для функции σ (раньше писали для коэффицентов).

В случае, когда все корни характеристического уравнения

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0$$

различны: $x_1, \ldots, x_p$, общее решение (23) имеет вид

$$\sigma(r) = \sum_{i=1}^{p} k_i x_i^r.$$

Граничные условия имеют вид

$$\begin{cases} \sum \beta_r \sigma(-r) = \sigma^2, \\ \sigma(-h) = \sigma(h), \quad 1 \leq h \leq p-1. \end{cases}$$

Примеры

- $p = 1$. В этом случае

$$\sigma(h) = \frac{(-\beta_1)^h \sigma^2}{1 - \beta_1^2}, \quad h = 0, 1, 2, \ldots, \quad \sigma^2 = \mathbb{E} u_t^2.$$

- $p = 2$, причем $x_1 \neq x_2$. Тогда

$$\sigma(h) = \frac{\sigma^2}{(x_1 - x_2)(1 - x_1 x_2)} \left(\frac{x_1^{h+1}}{1 - x_1^2} - \frac{x_2^{h+1}}{1 - x_2^2} \right).$$

  Если $x1$, $x2$ комплексно сопряжены:

$$x_1 = \alpha e^{i\theta}, \quad x_2 = \alpha e^{-i\theta},$$

  то

$$\sigma(h) = \frac{\sigma^2 \alpha^h \left[\sin\theta(h+1) - \alpha^2 \sin\theta(h-1) \right]}{(1 - \alpha^1) \sin\theta \left[1 - 2\alpha^2 \cos 2\theta + \alpha^4 \right]}.$$

Прогнозирование в модели $\mathsf{AR}(p)$

Теорема 4.3. *Пусть y_t – устойчивый процесс $\mathsf{AR}(p)$, удовлетворяющий условиям теоремы 2.1 о разложении. Тогда прогноз*

$$\widetilde{y}_t = -\beta_1 y_{t-1} - \ldots - \beta_p y_{t-p}$$

имеет наименьшую среднеквадратичную ошибку среди всех остальных прогнозов, измеримых относительно σ-алгебры $\mathcal{F}^y_{t-1}$, порожденной процессом y до момента $t-1$ включительно.

Доказательство Если есть σ-алгебра событий $\mathcal{F}$ и случайная величина ξ, тогда наименьшая в среднеквадратическом оценка для ξ среди всех случайных величин, измеримых относительно $\mathcal{F}$ – это $\mathbb{E}\left(\xi|\mathcal{F}\right)$. То есть

$$\mathbb{E}\left(\xi - \mathbb{E}\left(\xi|\mathcal{F}\right)\right)^2 \leq \mathbb{E}(\xi-\eta)^2, \quad \forall\, \eta \sim \mathbb{F}.$$

Поэтому для того, чтобы доказать теорему, достаточно доказать, что

$$\mathbb{E}\left(y_t|\mathcal{F}^y_{t-1}\right) = \widetilde{y}_t.$$

Действительно,[12]

$$\mathbb{E}\left(y_t|\mathcal{F}^y_{t-1}\right) = \mathbb{E}\left(-\underbrace{\beta_1 y_{t-1}}_{\sim\mathcal{F}^y_{t-1}} - \ldots - \underbrace{\beta_p y_{t-p}}_{\sim\mathcal{F}^y_{t-1}} + u_t|y_{t-1}, y_{t-2}, \ldots\right) =$$

$$= -\beta_1 y_{t-1} - \ldots - \beta_p y_{t-p} + \mathbb{E}(u_t|\mathcal{F}^y_{t-1}).$$

Так как будем считать u_t белым шумом в узком смысле, он не зависит от $\mathcal{F}^y_{t-1}$, а значит,

$$\mathbb{E}\left(y_t|\mathcal{F}^y_{t-1}\right) = -\beta_1 y_{t-1} - \ldots - \beta_p y_{t-p} + \underbrace{\mathbb{E}(u_t)}_{=0} = \widetilde{y}_t.$$

Теорема доказана.

[12] Напомним, что обозначение $X \sim \mathcal{F}$ означает *измеримость* случайной величины X относительно σ-алгебры $\mathcal{F}$.

Лекция 5

Прогнозирование на $s+1$ шаг в модели $\mathsf{AR}(p)$

В прошлый раз сформулировали теорему 4.3 о том, как выглядит прогноз на один шаг.

Теорема 5.1. *В тех же условиях,*[13] *оптимальной в среднеквадратическом прогоз в модели* $\mathsf{AR}(p)$ *на* $s+1$ *шагов вперед от момента* $t-s-1$ *до момента* t *равен*

$$\mathbb{E}\left(y_t|\mathcal{F}^y_{t-s-1}\right) = a_{s1}y_{t-s-1} + \ldots + a_{sp}y_{t-s-p},$$

где коэффициенты $a_{s1},\ldots,a_{sp}$ *были определены при доказательстве теоремы 2.1 о разложении устойчивого* $\mathsf{AR}(p)$.

Доказательство Вычислим условное матожидание, используя представление из доказательства теоремы 2.1 о разложении:

$$\mathbb{E}\left(y_t|\mathcal{F}^y_{t-s-1}\right) = \mathbb{E}\left(\underbrace{u_t + \delta_1 u_{t-1} + \ldots + \delta_s u_{t-s}}_{\text{не зависят от } \mathcal{F}^y_{t-s-1}} + \underbrace{a_{s1}y_{t-s-1} + \ldots + a_{sp}y_{t-s-p}}_{\sim\mathcal{F}^y_{t-s-1}}|\mathcal{F}^y_{t-s-1}\right) =$$

$$= \underbrace{\mathbb{E}(u_t)}_{=0} + \delta_1 \underbrace{\mathbb{E}(u_{t-1})}_{=0} + \ldots + \delta_s \underbrace{\mathbb{E}(u_{t-s})}_{=0} + a_{s1}y_{t-s-1} + \ldots + a_{sp}y_{t-s-p}.$$

Теорема доказана.

Модель скользящего среднего $\mathsf{MA}(q)$

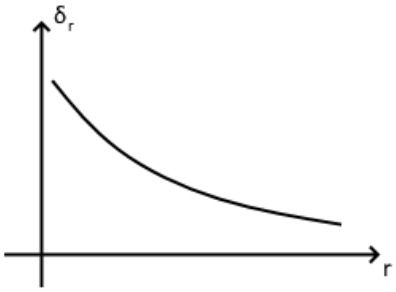

Рис. 5.1. δ_r в модели $\mathsf{AR}(p)$.

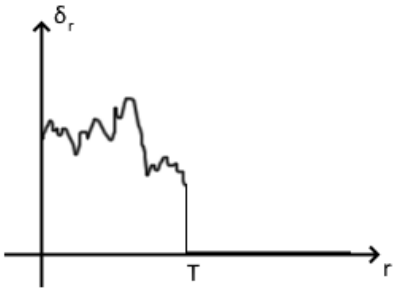

Рис. 5.2. δ_r в модели $\mathsf{MA}(q)$.

Остается вопрос – для всех ли процессов подходит такая модель? Вспомним, что ковариационная функция процесса $\mathsf{AR}(p)$, которую мы вычислили, сразу начинает экспоненциально убывать (рис. 5.1).

[13] Имеются в виду условия теоремы 4.3, то есть y_t – устойчивый процесс $\mathsf{AR}(p)$, удовлетворяющий условиям теоремы 2.1 о разложении.

Это является проблемой в случаях, когда процесс имеет конечную «память», и его ковариационная функция на конечном растоянии может вести себя произвольным образом (а потом, как правило, экспоненциально убывает).

Чтобы расширить границы применимости модели авторегрессии, ее модифицируют с помощью *модели скользящего среднего* $\mathsf{MA}(q)$ *порядка* q, ковариационная функция которой ведет себя произвольно на отрезке от 0 до q (рис. 5.2). Затем эти две модели объединяют.

Определение 5.1. Процесс $y_t \in \mathsf{MA}(q)$, если

$$\exists\ \alpha_0, \dots, \alpha_q \in \mathbb{R}, \quad \alpha_0 = 1, \quad \alpha_q \neq 0$$

такие,что

$$y_t = \sum_{i=0}^{q} \alpha_i u_{t-i}, \quad t \in \mathbb{Z}$$

где u_t – белый шум.

Лемма 5.1. *Математическое ожидание процесса* $y_t \in \mathsf{MA}(q)$ *имеет вид*

$$\mathbb{E} y_t = 0,$$

а ковариационная функция

$$\sigma(t-s) = \mathbb{E} y_t y_s = \begin{cases} \sigma^2 \sum_{i=0}^{q-|t-s|} \alpha_i \alpha_{t+|t-s|}, & \text{если } 0 \leq |t-s| \leq q, \\ 0, \quad \text{если } |t-s| > q, \end{cases}$$

где

$$\sigma^2 = \mathbb{E} u_t^2.$$

Доказательство данной леммы производится непосредственным вычислением и предлагается в качестве упражнения.

Определение 5.2. *Производящей функцией* для $\sigma(h)$ называется

$$\Sigma(z) = \sum_{h=-\infty}^{\infty} \sigma(h) z^h, \quad z \in \mathbb{C}. \tag{24}$$

Следующая лемма покажет, что ряд (24) сходится.

Лемма 5.2. *Для (24) верно*

$$\Sigma(z) = \sigma^2 P(z) P(z^{-1}),$$

где

$$P(z) = \sum_{i=0}^{q} \alpha_i z^{q-i}.$$

Доказательство Вычислим

$$\sum_{k=-\infty}^{\infty} \sigma(k) z^k = \sum_{k=-q}^{q} \sigma(k) z^k = \sigma^2 \left(\sum_{i=0}^{q-|k|} \alpha_i \alpha_{i+|k|} \right) z^k.$$

Здесь воспользовались результатом леммы 5.1, что при больших значениях аргумента ковариационная функция равна 0.

С другой стороны,

$$P(z)P(z^{-1}) = \sum_{i=0}^{q} \alpha_i z^{q-i} \sum_{j=0}^{q} \alpha_j z^{j-q} = \sum_{i=0}^{q} \alpha_i z^{q-i} \sum_{k=-i}^{q-i} \alpha_j z^{k+j-q}.$$

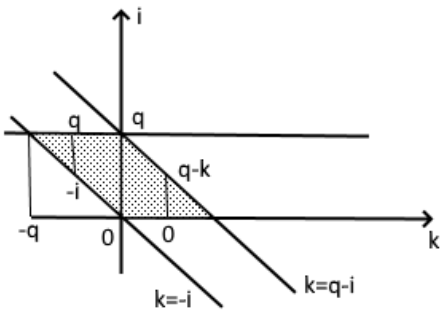

Рис. 5.3. Замена переменных $k = j - i$.

Сделаем замену $k = j - i$ (рис. 5.3). Тогда

$$P(z)P(z^{-1}) = \sum_{i=0}^{q} \sum_{k=-i}^{q-i} \alpha_i \alpha_{k+i} z^k.$$

Теперь поменяем порядок суммирования. Получим, что

$$P(z)P(z^{-1}) = \sum_{k=-q}^{0} \sum_{i=-k}^{q} \alpha_i \alpha_{i+k} z^k + \sum_{k=1}^{q} \sum_{i=0}^{q-k} \alpha_i \alpha_{i+k} z^k.$$

Сделав замену $i' = i + k$, получим

$$P(z)P(z^{-1}) = \sum_{k=-q}^{0} \sum_{i'=0}^{q+k} \alpha_{i'} \alpha_{i'+k} z^k + \sum_{k=1}^{q} \sum_{i=0}^{q-k} \alpha_i \alpha_{i+k} z^k =$$

$$= \sum_{k=-q}^{q} \sum_{i=0}^{q-|k|} \alpha_i \alpha_{i+|k|} z^k.$$

Лемма доказана.

Теперь покажем, что для любой ковариационной функции стационарного процесса, которая обращается в ноль при значениях аргумента, больших q, можно подобрать процесс скользящего среднего, который будет иметь эту ковариационную функцию.

Лемма 5.3. *Для $\forall\ \sigma(k)$ – ковариационной функции стационарного процесса, такой, что*

$$\sigma(k) = 0, \quad |k| > q, \quad \sigma(q) \neq 0,$$

$\exists\ y_t \in \mathsf{MA}(q)$ такой, что

$$\sigma_y(k) = \sigma(k),$$

то есть ковариационная функция процесса y_t совпадает с заданной ковариационной функцией $\sigma(k)$.

Доказательство Построим искомый процесс y_t. Пусть такой процесс имеется и его уравнение

$$y_t = \sum_{i=0}^{q} {}_i u_{t-i}, \quad {}_0 = 1, \quad \alpha_q \neq 0.$$

Будем искать $\alpha_0, \ldots, \alpha_q$.

Рассмотрим многочлен

$$P(z) = \sum_{i=0}^{q} \alpha_i z^{q-i}.$$

Обозначим его корни

$$z_1, \ldots, z_q \in \mathbb{C}.$$

Имеем, во-первых,

$$z_j \neq 0, \quad \forall\, 1 \leq j \leq q, \text{ так как } \alpha_q \neq 0,$$

и, во-вторых, z_j либо вещественны, либо образуют пары комплексно-сопряженных корней, так как коэффиценты многочлена $P(z)$ вещественны.

Далее, по лемме 5.2, если $\sigma(k)$

$$\sum_{k=-q}^{q} \sigma(k) z^k = \sigma^2 P(z) P(z^{-1})$$

Следовательно, уравнение

$$\sum_{i=0}^{2q} \sigma(i-q) z^i = z^q \sigma^2 P(z) P(z^{-1}) = 0$$

имеет корни

$$\{z_1, \ldots, z_q, 1/z_1, \ldots, 1/z_q\}. \tag{25}$$

Теперь,

$$P(z) = \sum_{i=0}^{q} \alpha_i z^{(q-i)} = \prod_{i=1}^{q} (z - z_i).$$

Если бы знали, в каком порядке расположены корни (25), могли бы подставить их в произведение и получить все α_i.

Обозначим (25) через

$$X = \{x_1, \ldots, x_{2q}\}.$$

Заметим, что, если $x_i \in X$, то $1/x_i \in X$, если $|x_i| \neq 1$, и кратности x_i и $1/x_i$ в множестве X совпадают.

Если $|x_i| = 1$, то он имеет четную кратность в множестве X.

Остановимся на последнем утверждении подробнее. Действительно, если $|x_i| = 1$, то

$$x_i = e^{i\lambda}.$$

Тогда

$$\sum_{k=-q}^{q} \sigma(k) x_i^k = \sum_{k=-q}^{q} \sigma(k) e^{i\lambda k}$$

– спектральная плотность стационарного процесса, для которого $\sigma(k)$ –ковариационная функция. Из теории стационарных случайных процессов известно, что спектральная плотность – вещественная неотрицательная функция.

Корень в минимуме всегда четной кратности.

Отсюда следует, что множество $X = \{x_1, \ldots, x_{2q}\}$ можно перенумеровать так, чтобы

$$|x_i| < 1 \;\Rightarrow\; x_{q+i} = \frac{1}{x_i}, \quad 1 \leq i \leq q,$$

$$|x_i| = 1 \;\Rightarrow\; x_{q+i} = x_i, \quad 1 \leq i \leq q.$$

После такой перенумерации строим многочлен

$$P^*(z) = \prod_{j=1}^{q} (z - x_i) = \sum_{j=0}^{q} \alpha_j^* z^{q-j}$$

и получаем коэффиценты

$$\alpha_0^*, \ldots, \alpha_q^*.$$

Строим процесс $y_t^* \in \mathsf{MA}(q)$ с уравнением

$$y_t^* = \sum_{i=0}^{q} \alpha_i^* u_{t-i}.$$

По построению ковариационная функция этого процесса будет совпадать с $\sigma(k)$ – заданной ковариационной функцией.

Лемма доказана.

Примеры

- Рассмотрим $\mathsf{MA}(1)$, $q = 1$. Уравнение процесса имеет вид

$$y_t = u_r - \theta u_{t-1}.$$

Тогда

$$\sigma(0) = \sigma^2(1+\theta), \quad \sigma^2 = \mathbb{E}u_t^2,$$
$$\sigma(1) = -\sigma^2\theta.$$

При этом корреляция

$$\rho_{t-s} = \mathsf{corr}(y_t, y_s) = \begin{cases} -\theta/(1+\theta^2), & |t-s| = 1, \\ 0, & |t-s| > 1. \end{cases}$$

Если известна функция

$$\rho_{t-s} = \mathsf{corr}(y_t, y_s) = \begin{cases} \rho_1, & h = 1, \\ 0, & h > 1, \end{cases}$$

то θ определяется из уравнения

$$\theta^2\rho_1 + \theta + \rho_1 = 0.$$

Корни уравнения имеют вид

$$\theta_{1,2} = \frac{-1 \pm \sqrt{1-4\rho_1^2}}{2\rho_1}.$$

При этом θ_1, θ_2 – вещественны, если

$$D = 1 - 4\rho_1^2 \geq 0,$$

то есть

$$|\rho_1| \leq \frac{1}{2}.$$

Таким образом, коэффиценты вещественны, если коэффицент корреляции не превосходит по модулю 1/2.

Лекция 6

Процессы $\mathsf{ARMA}(p, q)$

Определение 6.1. Процесс $y_t \in \mathsf{ARMA}(p, q)$, если

$$\exists\ \beta_0, \ldots, \beta_p, \quad \beta_0 = 1, \quad \beta_p \neq 0, \quad \beta_i \in \mathbb{R},$$

$$\alpha_0, \ldots, \alpha_p, \quad \alpha_0 = 1, \quad \alpha_q \neq 0, \quad \alpha_j \in \mathbb{R}$$

такие, что $\forall\ t \in \mathbb{Z}$ выполнено

$$\sum_{s=0}^{p} \beta_s y_{t-s} = \sum_{j=0}^{q} \alpha_j u_{t-j}, \tag{26}$$

где u_t – дискретный белый шум. Уравнение (26) называется уравнением процесса, (p, q) – порядок процесса $\mathsf{ARMA}(p, q)$.

Процессу y_t соответствует два характеристических уравнения

$$\sum_{s=0}^{p} \beta_s x^{p-s} = 0, \tag{27}$$

$$\sum_{j=0}^{q} \alpha_q z^{q-j} = 0. \tag{28}$$

Обозначим корни уравнения (27) через $x_1, \ldots, x_p$, а корни уравнения (28) через $z_1, \ldots, z_q$. Если все[14]

$$|x_i| < 1, \quad 1 \leq i \leq p,$$

то процесс y_t называется *устойчивым.*

Предположение Будем всегда считать, что $\exists\ N$ такое, что $\forall\, n < -N$ выполнено

$$\mathbb{E} y_n^2 < N.$$

Теорема 6.1. *Если процесс $y_t \in \mathsf{ARMA}(p, q)$ устойчив и пара его характеристических уравнений (27), (28) имеет m попарно совпадающих корней:*

$$\exists\ 1 \leq i_1 < i_2 < \ldots < i_m \leq p,$$

такие, что

$$\begin{array}{ccc} x_{i_1} & = & z_{j_1}, \\ & \ldots\ldots & \\ x_{1_m} & = & x_{j_m}, \end{array}$$

где $j_1, \ldots, j_m$ – несовпадающие индексы от 1 до q. Тогда

$$y_t \in \mathsf{ARMA}(p - m, q - m).$$

[14] Обратим внимание, что в данном случае делается условие только на корни авторегрессионной части.

Доказательство Обозначим оператор сдвига

$$L\xi_t = \xi_{t-1}.$$

Уравнение (26) можно записать в виде

$$\left(\sum_{s=0}^{p} \beta_s L^s\right) y_t = \left(\sum_{j=0}^{q} \alpha_j L^j\right) u_t.$$

Тогда можно записать

$$\left(\prod_{i=1}^{p} (1 - x_i L)\right) y_t = \left(\prod_{j=1}^{q} (1 - z_j L)\right) u_t, \tag{29}$$

так как $\beta_0 = 1$, $\alpha_0 = 1$, где корни многочлена

$$\sum_{i=0}^{p} \beta_i x^i = 0$$

это числа

$$1/x_1, \ldots, 1/x_p,$$

а корни многочлена

$$\sum_{j=0}^{q} \alpha_j z^j = 0$$

– числа

$$1/z_1, \ldots, 1/z_q.$$

Можем переписать (29) в виде

$$\underbrace{\left(\prod_{i=i_1,\ldots,i_m} (1 - x_i L)\right)}_{A} \left(\prod_{\substack{i \neq i_1,\ldots,i_m \\ 1 \le i \le p}} (1 - x_i L)\right) y_t =$$

$$= \underbrace{\left(\prod_{j=j_1,\ldots,j_m} (1 - z_j L)\right)}_{B} \left(\prod_{\substack{j \neq j_1,\ldots,j_m \\ 1 \le j \le q}} (1 - z_j L)\right) u_t. \tag{30}$$

По условию $A = B$. Так как все $|x_i| < 1$, то оператор A имеет обратный (см. теорему 2.1 о разложении). Значит,

$$\exists\ A^{-1} = B^{-1}.$$

Тогда можем сократить A и B в (30). Получим, что

$$\left(\prod_{\substack{i \neq i_1,\ldots,i_m \\ 1 \le i \le p}} (1 - x_i L)\right) y_t = \left(\prod_{\substack{j \neq j_1,\ldots,j_m \\ 1 \le j \le q}} (1 - z_j L)\right) u_t.$$

Раскроем скобки. Получим

$$\left(\sum_{i=0}^{p-m} \beta_i^* L^i\right) y_t = \left(\sum_{j=0}^{q-m} \alpha_j^* L^j\right) u_t,$$

то есть

$$\sum_{i=0}^{p-m} \beta_i^* y_{t-i} = \sum_{j=0}^{q-m} \alpha_j^* u_{t-j}.$$

Значит, $y_t \in \mathsf{ARMA}(p-m, q-m)$.

Теорема доказана.

Теорема о разложении устойчивых процессов $\mathsf{ARMA}(p,q)$

Теорема 6.2. *(о разложении устойчивых процессов $\mathsf{ARMA}(p,q)$) Если $y_t \in \mathsf{ARMA}(p,q)$ устойчив, то есть все $|x_i| < 1, \quad 1 \le i \le p$, то $\exists\ \{\gamma_i\}_{i=0}^{\infty}, \quad \gamma_i \in \mathbb{R}$ такие, что*

$$y_t \overset{ср.кв.}{=} \sum_{i=0}^{\infty} \gamma_i u_{t-i},$$

где γ_i можно определить из соотношений

$$\begin{cases} \gamma_0 = \delta_0 \alpha_0 = 1, \\ \gamma_1 = \delta_0 \alpha_1 + \delta_1 \alpha_0 = \alpha_1 + \delta_1, \\ \ldots\ldots \\ \gamma_{q-1} = \delta_0 \alpha_{q-1} + \delta_1 \alpha_{q-2} + \ldots + \delta_{q-1} \alpha_0, \end{cases}$$

и однородного разностного уравнения

$$\gamma_r = \delta_{r-q} \alpha_q + \ldots + \delta_r \alpha_0, \quad r \ge q,$$

где коэффиценты δ определяются по корням характеристического уравнения (27):

$$\sum_{i=0}^{p} \beta_i x^{p-i} = 0,$$

как в теореме 2.1 о разложении устойчивого процесса $\mathsf{AR}(p)$.

Доказательство Обозначим

$$g_t = \sum_{j=0}^{q} \alpha_j u_{t-j}.$$

Тогда (26) примет вид

$$\sum_{i=0}^{p} \beta_i y_{t-i} = g_t, \quad t \in \mathbb{Z}.$$

Это уравнение устойчивого процесса $\mathsf{AR}(p)$, где вместо белого шума стоит процесс g_t с ограниченной дисперсией и нулевым матожиданием, так как

$$\mathbb{E}g_t = \sum \alpha_j \mathbb{E}u_{t-j} = 0,$$

$$\mathbb{E}g_t^2 = \sum \alpha_j^2 \sigma^2 < \infty,$$

где $\sigma^2 = \mathbb{E}u_t^2$.

К такому процессу применима теорема о разложении устойчивого процесса $\mathsf{AR}(p)$, а значит,

$$y_t \overset{\text{ср.кв.}}{=} \sum_{s=0}^{\infty} \delta_s g_{t-s} = \sum_{s=0}^{\delta}{}_s \sum_{j=0}^{q} \alpha_j u_{t-s-j}.$$

Сделаем замену $r = s + j$.

Получим, что

$$y_t \overset{\text{ср.кв.}}{=} \sum_{r=0}^{\infty} \underbrace{\left(\sum_{j=0}^{q \wedge r} \delta_{r-j} \alpha_j \right)}_{\gamma_r} u_{t-r} = \sum_{r=0}^{\infty} \gamma_r u_{t-r},$$

где

$$\gamma_r = \sum_{j=0}^{q \wedge r} \delta_{r-j} \alpha_j.$$

Теорема доказана.

Введем ковариационную функцию

$$\sigma(r) = \mathbb{E}y_t y_{t-r}.$$

Теорема 6.3. *Если y_t – устойчивый процесс $\mathsf{ARMA}(p, q)$, то*

$$\sum_{s=0}^{p} \beta_s \sigma(h - s) = 0, \quad h \geq q + 1.$$

Доказательство Имеем

$$\sum_{s=0}^{p} \beta_s y_{t-s} = \sum_{j=0}^{q} \alpha_j u_{t-j}.$$

Домножим обе части равенства на y_{t-h} и возьмем матожидание. Учитывая, что

$$\mathbb{E}y_t = 0, \quad \mathbb{E}y_{t-s} y_{t-h} = \sigma(h - s),$$

получим, что

$$\sum_{s=0}^{p} \beta_s \delta(h - s) = \sum_{j=0}^{q} \sum_{r=0}^{\infty} \alpha_j \gamma_r \mathbb{E}u_{t-j} u_{t-h-r} = \sigma^2 \sum_{j=0}^{q} \sum_{r=0}^{\infty} \alpha_j \gamma_r \delta_{j,r+h}, \quad h \geq 0, \tag{31}$$

где

$$\delta_{i,j} = \begin{cases} 1, & i = j, \\ 0, & \textit{иначе}. \end{cases}$$

При $h > q$ (31) $= 0$, так как в этих условиях индекс $j \leq q$, а $r + h > q$, так как $r \geq 0$. Теорема доказана.

Получаем однородное разностное уравнение на ковариационную функцию $\sigma(\cdot)$ процесса $\mathsf{ARMA}(p, q)$:

$$\sum_{s=0}^{p} \beta_s \sigma(h - s) = 0, \quad h \geq q + 1.$$

Граничными условиями для этого уравнения являются соотношения

$$\sum_{s=0}^{p} \beta_s \sigma(h - s) = \sigma^2 \sum_{j=0}^{q} \sum_{r=0}^{\infty} \alpha_j \gamma_r \delta_{j,r+h}, \quad h = 1, 2, \ldots, q,$$

где

$$\delta_{i,j} = \begin{cases} 1, & i = j, \\ 0, & \textit{иначе}. \end{cases}$$

Так, видим, что для процессов $\mathsf{ARMA}(p, q)$ возникает ситуация, похожая на ситуацию для процессов $\mathsf{AR}(p)$. Справедлива теорема о разложении, разложение имеет такую же структуру – бесконечно скользящее среднее по значениям белого шума, вычисляется ковариационная функция.

Такие процессы могут иметь существенно более широкое применение, так как на конечном интервале их ковариационная функция может вести себя произвольно.

Позже в курсе покажем, почему эти процессы недостаточны и их нужно еще модифицировать, чтобы удовлетворить запросам финансовой математики.

Нестационарные модели

Обсудим теперь, что делать, если процесс нестационарен.

- Рассмотрим случай непостоянной дисперсии. Пусть процесс представим в виде

  $$X_t = \mu_t + \epsilon_t,$$

  и матожидание $\mathbb{E}X_t = \mu_t$. Пусть

  $$\mathbb{D}(\epsilon_t) = \mathbb{D}(X_t) = \mathbb{E}(X_t - \mu_t)^2 = h^2(\mu_t)\sigma^2.$$

  Рассмотрим функцию от X_t

  $$g(X_t) = \int_a^x \frac{dx}{h(x)},$$

  тогда

  $$g'(x) = \frac{1}{h(x)}.$$

По формуле Тейлора имеем

$$g(X_t) = g(\mu_t) + (X_t - \mu_t)g'(\mu_t) + o(|X_t - \mu_t|).$$

Значит,

Тогда

$$\mathbb{D}(g(X_t)) \sim \mathbb{D}\left[g(\mu_t) + (X_t - \mu_t)g'(\mu_t)\right] =$$
$$= \left[g'(\mu_t)\right]^2 \mathbb{D}(X_t) = \left[g'(\mu_t)\right]^2 h^2(\mu_t)\sigma^2 = \sigma^2,$$

так как

$$g'(\mu_t) = \frac{1}{h(\mu_t)}.$$

Таким образом, дисперсию можно сделать приблизительно постоянной.

Типичным примером, например, является случай, когда разброс пропорционален достигнутому уровню процесса, то есть

$$h(\mu_t) = c\mu_t.$$

Тогда

$$g'(\mu_t) = \frac{1}{c\mu_t},$$

а значит,

$$g(x) = \frac{1}{c}\ln x.$$

В этом случае процесс надо логарифмировать

$$X_t \ \to \ \ln(X_t).$$

Тогда дисперсия стабилизируется.

Лекция 7

Примеры

- Рассмотрим теперь случай, когда у процесса нестационарно среднее значение. Представим процесс в виде
$$X_t = \mu_t + \epsilon_t. \tag{32}$$
Будем считать, что[15]
$$\mu_t = \sum_{j=0}^{d} \beta_j t^j,$$
а $\epsilon_t \in \mathsf{ARMA}(p,q)$ с уравнением
$$\phi(L)\epsilon_t = \theta(L)u_t,$$
где L – оператор сдвига, u_t – дискретный белый шум,
$$\phi(L) = \phi_0 + \phi_1 L + \ldots + \phi_p L^p, \quad \phi_p \neq 0,$$
$$\theta(L) = \theta_0 + \theta_1 L + \ldots + \theta_q L^q, \quad \theta_q \neq 0.$$
Основная идея заключается в том, чтобы преобразовать представление процесса (32). Так как μ_t – полином степени d, от него можно было бы избавиться, $d+1$ раз продифференциировав (32), но в нашем случае величины дискретны. Воспользуемся оператор разностного дифференцирования
$$\nabla X_t = (1 - LX_t),$$
$$\nabla = 1 - L,$$
$$\nabla^d X_t = (1-L)^d X_t.$$

Лемма 7.1. *Пусть ϵ_t – устойчивый процесс $\mathsf{ARMA}(p,q)$. Тогда процесс*
$$y_t = \nabla^d X_t$$
удовлетворяет уравнению типа $\mathsf{ARMA}(p, d+q)$
$$y_t = \psi_0 + \frac{\nabla^d \theta(L)}{\phi(L)} u_t,$$
где дробь имеет смысл обратного, или
$$\phi(L)y_t = \psi_1 + \nabla^d \theta(L) u_t,$$
где
$$\psi_0 = d!\beta_d, \quad \psi_1 = \psi_0 + \sum_{i=0}^{p} \phi_i = \phi(1)\beta_d d!.$$

[15] Здесь исходим из соображений, что любую функцию на некотором интервале можно приблизить полиномом с заданной степенью точности.

Доказательство Имеем

$$X_t = \sum_{j=0}^{d} \beta_j t^j + \epsilon_t,$$

$$\epsilon_t = [\phi(L)]^{-1}\,\theta(L)u_t$$

в силу устойчивости ϵ_t по теореме о разложении. Тогда

$$X_t = \sum_{j=0}^{d} \beta_j t^j + [\phi(L)]^{-1}\,\theta(L)u_t.$$

Применив к обеим частям d раз оператор ∇, получим, что

$$\nabla^d X_t = \nabla^d \left(\sum_{j=0}^{d} \beta_j t^j\right) + \nabla^d\,[\phi(L)]^{-1}\,\theta(L)u_t.$$

Для второго слагаемого имеем

$$\nabla^d[\phi(L)]^{-1}\theta(L)u_t = [\phi(L)]^{-1}\nabla^d\theta(L)u_t = \frac{\nabla^d}{\theta(L)}\phi(L)u_t.$$

Для первого слагаемого

$$\nabla^d \left(\sum_{j=0}^{d} \beta_j t^j\right) = d!\beta_d,$$

доказательство можно осуществить по индукции.

Получается, что

$$\phi(L)y_t = \psi_1 + \nabla^d\theta(L)u_t.$$

С учетом того, что

$$\phi(L)\psi_0 = (\phi_0 + \phi_1 + \ldots + \phi_p)\,\psi_0 = \psi_1,$$

получим, что

$$\phi(L)y_t = \psi_1 + \nabla^d\theta(L)u_t.$$

Лемма доказана.

Следствие 1. *Если $\psi_0 = 0$ (то есть $\sum \phi_i = 0$), то*

$$\nabla^d X_t \in \mathsf{ARMA}(p, d+q).$$

Процесс $\mathsf{ARIMA}(p, d, q)$

Определение 7.1. Случайный процесс $X_t \in \mathsf{ARIMA}(p, d, q)$, если для него справедливо

$$\phi(L)\nabla^d X_T = a + \theta(L)u_t,$$

где a – константа, u_t – дискретный белый шум,

$$\phi(L) = \phi_0 + \phi_1 L + \ldots + \phi_p L^p, \quad \phi_p \neq 0, \quad \phi_0 = 1, \quad \phi_i \in \mathbb{R}.$$

$$\theta(L) = \theta_0 + \theta_1 L + \ldots + \theta_q L^q, \quad \theta_q \neq 0, \quad \theta_0 = 1, \quad \theta_j \in \mathbb{R}.$$

Здесь $t \geq T_0$ ($t_0 \geq 0$).[16]

[16] Это ограничение возникает из-за дифференцирования.

Предположение Все корни характеристического уравнения

$$\sum_{i=0}^{p} \phi_i x^{p-i} = 0$$

по модулю < 1.

В зависимости от значения константы a у нас возникает два принципиально разных случая.

- $a = 0$. Тогда

$$\phi(L) \underbrace{\nabla^d X_T}_{y_t} = \theta(L) u_t,$$

то есть

$$\phi(L) \nabla^d X_T = \theta(L) u_t,$$

y_t – устойчивый $\mathsf{ARMA}(p, q)$. Тогда

$$X_t = [\nabla^d]^{-1} y_t = \left[(1-L)^d\right]^{-1} y_t = \left[(1-L)^{-1}\right]^d y_t = S^d y_t,$$

где

$$S = (1 - L) = 1 + L + L^2 + L^3 + \ldots$$

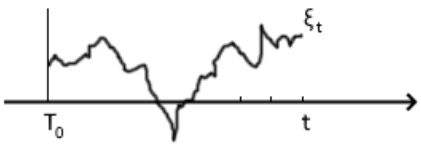

Рис. 7.1. Вид процесса ξ_t

Заметим, что

$$S\xi_t = \xi_t + \xi_{t-1} + \ldots + \xi_{T_0},$$

где T_0 – начальный момент времени (7.1). Итак, X_t – интегрированный процесс $\mathsf{ARMA}(p, q)$.

- $a \neq 0$.

Лемма 7.2. *Пусть $a \neq 0$. Тогда процесс X_t имеет полиномиальный тренд порядка d, то есть $\mathbb{E}X_t$ – полином степени d от t.*

Доказательство Запишем

$$\phi(L) y_t = a + \theta(L) u_t.$$

Тогда

$$y_t = [\phi(L)]^{-1} a + \phi^{-1}(L) \theta(L) u_t,$$

$$\mathbb{E}y_t = [\phi(L)]^{-1}a,$$

так как

$$\mathbb{E}\phi^{-1}(L)\theta(L)u_t = 0.$$

Но

$$[\phi(L)]^{-1}a = \frac{a}{1+\phi_1+\ldots+\phi_p},$$

так как

$$\phi(L)\left(\frac{a}{\sum\phi_i}\right) = \sum\phi_i\left(\frac{a}{\sum\phi_i}\right) = a.$$

Предполагаем, что $\sum\phi_i \neq 0$.

В итоге получаем, что

$$\mathbb{E}y_t = A = \frac{a}{\sum\phi_i},$$

$$\mathbb{E}X_t = S^d A = \sum_{i=0}^{d}\beta_i t^i$$

– некий многочлен. Здесь A – константа.

Лемма доказана.

Недетерминированные тренды в ARIMA

Обсудим теперь, как по процессу понять, есть ли у него тренд. Рассмотрим несколько примеров.

Примеры

- Рассмотрим случайное блуждание

  $$X_t = X_{t-1} + u_t,$$

  где $a = 0$, $u_t \sim \mathcal{N}(0, \sigma^2)$.

  X_t является процессом $\mathsf{ARIMA}(0,1,0)$.

- Рассмотрим процесс

  $$X_t = 2 + X_{t-1} + u_t,$$

  где $a \neq 0$, u_t – дискретный белый шум, $u_t \sim \mathcal{N}(0, \sigma^2)$.

  Процесс $X_t \in \mathsf{ARIMA}(0,1,0)$.

Во втором случае у процесса есть тренд, так как $a \neq 0$. Если мы посмотрим на реализации этих процессов, они могу оказаться очень похожими.

Теорема 7.1. *(закон арксинуса) Пусть $\xi_1, \ldots, \xi_n, \ldots$ – н.о.р.с.в,*

$$\mathbb{E}\xi_i = 0, \quad \mathbb{P}(\xi_i > 0) = \frac{1}{2}.$$

Обозначим

$$S_k = \sum_{i=1}^{k} \xi_i,$$

$$T_n = \#\{k:\ 1 \le k \le n:\ S_k > 0\}.$$

Тогда

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{T_n}{n} < x\right) = \frac{2}{\pi}\arcsin\sqrt{x} = F(x).$$

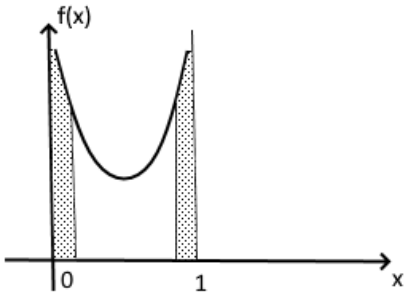

Рис. 7.2. график плотности f(x)

Если изобразим график плотности

$$f(x) = F'(x) = \frac{1}{\sqrt{x(1-x)}},$$

обнаружим, что большая масса распределения сосредоточена по краям (рис. **??**).

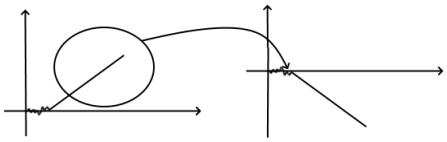

Рис. 7.3. Процесс с недетерминированным трендом.

Как раз такой процесс рассматривался в первом примере, когда каждый раз мы прибавляем симметрично распределенную случайную величину. Такой процесс типично будет выглядеть следующим образом (рис. 7.3)

В каждый момент времени будет казаться, что у наблюдаемого процесса есть тренд, но на самом деле он может быть как положительным, и отрицательным. Когда процесс меняет знак, участок текущего роста превосходит предыдущие по длине и поэтому кажется, что тренд определен.

Этот парадокс называется *недетерминированным трендом.*

Прогнозирование в ARIMA-моделях

Пусть процесс $X_t \in \mathsf{ARIMA}(p, d, q)$. Тогда имеет место уравнение

$$\phi(L)\nabla^d X_t = a + \theta(L)u_t.$$

Разложим на множители оператор

$$\phi(L)\nabla^d = \phi(L)(1-L)^d = \alpha(L) = \alpha_0 + \alpha_1 L + \ldots + \alpha_{p+d}L^{p+d}.$$

Обозначим $f_{T,h}$ оптимальный в среднеквадратическом смысле прогноз для X_{T+h} в момент времени t.

$$f_{T,h} = \mathbb{E}\left(X_{T+h}|\mathcal{F}_t^X\right),$$

где $\mathcal{F}_t^X$ – σ-алгебра событий, доступных в моменты времени t.

Теорема 7.2. *Пусть все корни характеристического уравнения*

$$\sum_{i=0}^{q} \theta_i x^{q-i} = 0$$

по модулю < 1. *Тогда*

$$f_{T,h} = \alpha_1 f_{T,h-1} + \alpha_2 f_{T,h-2} + \ldots + \alpha_{h-1} f_{T,1} + \alpha_h X_t + \ldots +$$

$$+\alpha_{p+d} X_{T-p-d+h} + G + a, \quad \textit{если } h \le p+d,$$

и

$$f_{T,h} = \alpha_1 f_{T,h-1} + \ldots + \alpha_{p+d} f_{T,h-p-d} + G + a, \quad \textit{если } h > p+d,$$

где

$$G = \theta_n u_T + \theta_{n+1} u_{t-1} + \ldots + \theta_q u_{T+h-q},$$

если $h < q$, *и*

$$G = 0$$

иначе.

Доказательство Вычислим

$$f_{T,h} = \mathbb{E}(X_{T+h}|X_T, X_{T-1}, \ldots) =$$

$$= \mathbb{E}\left(\alpha_1 X_{T+h-1} + \alpha_2 X_{T+h-2} + \ldots + \alpha_{p+d} X_{T+h-p-d} + a + \theta(L)u_{T+h}|\mathcal{F}_T^X\right),$$

Поскольку уравнение процесса имеет в точке $T+h$ вид

$$X_{T+h} - \alpha_2 X_{T+h-2} - \ldots - \alpha_{p+d} X_{T+h-p-d} = a + \theta(L)u_{T+h}.$$

Имеем

$$\mathbb{E}\left(a|\mathcal{F}_T^X\right),$$

$$\mathbb{E}\left(\theta(L)u_{T+h}|\mathcal{F}_T^X\right) = \mathbb{E}\left(\theta_0 u_{T+h} + \theta_1 u_{T+h-1} + \ldots + \theta_q u_{T+h-q}|\mathcal{F}_T^X\right). \tag{33}$$

Здесь получим, что первая часть слагаемых под знаком матожидания будет равна 0 при $i < h$, так как в этом случае $T + h - i > T$, а значит, u_{T+h-i} не зависит от $\mathcal{F}_T^X$ и

$$\mathbb{E}\left(u_{T+h-i}|\mathcal{F}_T^X\right) = \mathbb{E}\left(u_{T+h-i}\right) = 0.$$

Вторая часть слагаемых равна $\theta_i u_{T+h-i}$, если $i \geq h$, так как тогда $T + h - i \leq T$ и в силу предположения

$$u_{T+h-i} \sim \mathcal{F}_T^X$$

и, значит,

$$\mathbb{E}\left(\theta_i u_{T+h-i}|\mathcal{F}_T^X\right) = \theta_i u_{T+h-i}.$$

Значит, выражение (33) равно G.

$$\mathbb{E}(\alpha_i X_{T+h-i}|\mathcal{F}_t) = \begin{cases} \alpha_i \mathbb{E}(X_{T+h-i}|\mathcal{F}_t) = \alpha_i f_{T,h-i}, & \text{если } i < h, \\ \alpha_i \mathbb{E}(X_{T+h-i}|\mathcal{F}_t) = \alpha_i X_{T+h-i}, & \text{если } i \geq h. \end{cases}$$

Теорема доказана.

Лекция 8

Примеры прогнозирования в модели ARIMA

- Рассмотрим процесс
$$(1-\rho L)X_t = a + u_t.$$
Это процесс ARIMA$(1,0,0)$. Здесь
$$\alpha(L) = (1-\rho L),$$
то
$$\alpha_0 = 1, \quad \alpha_1 = \rho.$$
Запишем уравнение процесса для момента $T+h$:
$$X_{T+h} = \rho X_{T+h-1} + a + u_{T+h}.$$
Вычислим
$$\begin{aligned} f_{T,h} &= \rho f_{T,h-1} + a = \rho^2 f_{T,h-2} + \rho a + a = \ldots = \\ &= \rho^h f_{T,0} + a(1+\rho+\rho^2+\ldots+\rho^{h-1}) = \\ &= \rho^h X_t + a(1+\rho+\rho^2+\ldots+\rho^{h-1}). \end{aligned}$$
Если $|\rho| < 1$ и $h \to \infty$, то
$$\lim_{h\to\infty} f_{T,h} = a(1+\rho+\rho^2+\ldots) = \frac{a}{1-\rho}.$$
Следовательно,
$$\mathbb{E}X_t = \frac{a}{1-\rho}.$$

- Рассмотрим процесс ARIMA$(0,1,1)$ с уравнением
$$\nabla X_t = (1-)u_t,$$
$$\alpha(L) = 1 - L.$$
Тогда по теореме
$$X_{T+h} = X_{T+h-1} + u_{T+h} - \rho u_{T+h-1}$$
Если $h = 1$, то
$$f_{T,1} = X_T - \rho u_T.$$
При $h = 2$
$$f_{T,2} = f_{T,1}.$$
При $h > 1$ получаем
$$f_{T,h} = f_{T,h-1},$$
то есть прогноз постоянный.

Из уравнения процесса

$$u_T = (1-L)(1-\rho L)^{-1}X_T.$$

Тогда

$$f_{T,h} = f_{T,1} = X_T - \rho u_T = X_T - \rho(1-L)(1-\rho L)^{-1}X_T =$$
$$= X_T - [(\rho-1)+(1-\rho L)](1-\rho L)^{-1}X_T = X_T - X_t + (1-\rho)(1-\rho L)^{-1}X_T =$$
$$= (1-\rho)(1-\rho L)^{-1}X_T = (1-\rho)(1+++\rho^2L^2+\ldots)X_T =$$
$$= (1-\rho)(X_T + \rho X_{T-1} + \rho^2 X_{T-2} + \ldots).$$

Здесь последняя сумма продолжается до момента времени T_0.

- Рассмотрим процесс $\mathsf{ARIMA}(0,2,1)$. Уравнение процесса имеет вид

$$\nabla^2 X_T = (1-\theta_1 L - \theta_2 L^2)u_t,$$

$$\alpha(L) = (1-L^2) = (1-2L+L^2).$$

Тогда в момент $T+h$

$$X_{T+h} = 2X_{T+h-1} - X_{T+h-2} + u_{T+h} - \theta_1 u_{T+h-1} - \theta_2 u_{T+h-2}.$$

По теореме получим, что при $h=1$

$$f_{T,1} = 2X_T - X_{T-1} - \theta_1 u_T - \theta_2 U_{T-1},$$

при $h=2$

$$f_{T,2} = 2f_{T,1} - X_T - \theta_2 u_T,$$

а при $h \geq 3$

$$f_{T,h} = 2f_{T,h-1} - f_{T,1}.$$

Последнюю формулу получили, найдя уравнение для прямой $f_{T,h}$ по двум точкам $(1, f_{T,1})$ и $(2, f_{T,2})$.

Среднеквадратическая ошибка оптимального в среднеквадратическом прогноза в модели ARIMA

Пусть процесс $X_t \in \mathsf{ARIMA}(p,d,q)$. Его уравнение

$$\phi(L)\nabla^d X_t = a + \theta(L)u_t,$$

или

$$\alpha(L)X_t = a + \theta(L)u_t.$$

Обозначим[17]

$$\alpha^{-1}(L)\theta(L) = \psi(L),$$
$$\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \ldots$$

[17]Напомним, что так как у нас есть начальный момент времени, все операторы обратимы.

Теорема 8.1. *Пусть все корни характеристического уравнения*

$$\sum_{i=0}^{q} \theta_i x^{q-i} = 0$$

по модулю < 1. *Тогда дисперсия ошибки прогноза*

$$f_{T,h} = \mathbb{E}(X_{T+h}|\mathcal{F}_T^X)$$

дается формулой

$$\mathbb{D}(\underbrace{f_{T,h} - X_{T+h}}_{\text{ошибка прогноза}}) = \mathbb{D}(e_{T,h}) = \mathbb{E}(e_{T,h}^2) = \mathbb{E}(f_{T,h} - X_{T+h})^2 =$$

$$= \sigma^2(1 + \psi_1^2 + \phi_2^2 + \ldots + \psi_{n-1}^2),$$

где

$$\sigma^2 = \mathbb{E}u_t^2.$$

Доказательство Запишем

$$\alpha(L)X_{T+h} = a + \theta(L)u_{T+h}.$$

Следовательно,

$$X_{T+h} = [\alpha(L)]^{-1}a + [\alpha(L)]^{-1}\theta(L)u_{T+h}$$

Обозначим

$$[\alpha(L)]^{-1}a = A = \text{const}.$$

Тогда

$$X_{T+h} = A + \psi(L)u_{T+h},$$

или

$$X_{T+h} = A + u_{T+h} + \psi_1 u_{T+h-1} + \psi_2 u_{T+h-2} + \ldots \tag{34}$$

С другой стороны,

$$f_{T,h} = \mathbb{E}(X_{T+h}|\mathcal{F}_T^X) = \mathbb{E}(A + u_{T+h} + \psi_1 u_{T+h-1} + \ldots|\mathcal{F}_T^X) =$$
$$= A + \mathbb{E}(u_{T+h} + \psi_1 u_{T+h-1} + \ldots + \psi_n u_{T+1}|\mathcal{F}_T^X) +$$
$$+\mathbb{E}(\psi_n u_T + \psi_{n+1}u_{T-1} + \ldots + |\mathcal{F}_T^X).$$

Так как слагаемые под знаком первого матожидания независимы от $\mathcal{F}_T^X$, а под знаком второго все случайные величины измеримы относительно $\mathcal{F}_T^X$, получим, что

$$f_{T,h} = A + \underbrace{\mathbb{E}u_{T+h}}_{=0} + \psi_1 \underbrace{\mathbb{E}u_{T+h-1}}_{=0} + \ldots + \psi_n \underbrace{\mathbb{E}u_{T+1}}_{=0} + \psi_n u_T + \psi_{n+1}u_{T-1} + \ldots =$$
$$= A + \psi_n u_T + \psi_{n+1}u_{T-1} + \ldots. \tag{35}$$

Вычтем из (34) формулу (35). Получим

$$X_{T+h} - f_{T,h} = u_{T+h} + \psi_1 u_{T+h-1} + \ldots + \psi_{n-1}u_{T+1}.$$

Возведем в квадрат и возьмем матожидание. С учетом того, что

$$\mathbb{E}\left(X_{T+h} - f_{T,h}\right) = \mathbb{E}\left(X_{T+h} - \mathbb{E}(X_{T+h}|\mathcal{F}_T^X)\right) = 0,$$

получим, что

$$\mathbb{E}\left(X_{T+h} - f_{T,h}\right)^2 = \sigma^2(1 + \psi_1^2 + \ldots + \psi_{n-1}^2).$$

Теорема доказана.

Примеры

Рассмотрим несколько примеров.

- Рассмотрим процесс $\mathsf{ARIMA}(1,0,0)$ – случайное блуждание. Уравнение процесса имеет вид
$$X_T = \rho X_{t-1} + a + u_t.$$
Выше было показано, что
$$f_{T,h} = \rho^h X_T + a(1 + \rho + \rho^2 + \ldots + \rho^{h-1}).$$
В этом случае
$$\alpha(L) = 1 - \rho L, \quad \theta(L) = 1.$$
Значит,
$$\psi(L) = (1 - \rho L)^{-1} = 1 + \rho L + \rho^2 L^2 + \rho^3 L^3 + \ldots.$$
Из этого представления
$$\psi_i = \rho^i.$$
Значит,
$$\mathbb{D}(e_{T,h}) = \sigma^2(1 + \rho^2 + \rho^4 + \ldots + \rho^{2(h-1)}) = \sigma^2 \frac{1 - \rho^{2h}}{1 - \rho^2}.$$
Если $h \to \infty$, то
$$\mathbb{D}(e_{T,h}) \underset{h\to\infty}{\to} \frac{\sigma^2}{1 - \rho^2}.$$

- Рассмотрим процесс $\mathsf{ARIMA}(0,1,1)$ с уравнением
$$\nabla X_T = (1 - \rho L) u_t.$$
Здесь
$$\alpha(L) = 1 - L, \quad \theta(L) = (1 - \rho L).$$
Тогда
$$\psi(L) = (1 - L)^{-1}(1 - \rho L) = (1 - \rho L)(1 + L + L^2 + L^3 + \ldots) =$$
$$= 1 + (1 - \rho)L + (1 - \rho)L^2 + \ldots$$
Получаем, что
$$\psi_i = 1 - \rho.$$
Тогда
$$\mathbb{D}(e_{T,h}) = \sigma^2 \left(1 + (1 - \rho^2)^2 (h - 1)\right).$$
В этом случае при $h \to \infty$ получаем
$$\mathbb{D}(e_{T,h}) \underset{h\to\infty}{\to} \infty.$$

- Рассмотрим процесс $\mathsf{ARIMA}(0,2,2)$, разобранный выше. Уравнение процесса имеет вид
$$\nabla^2 X_T = (1-\theta_1 L - \theta_2 L^2)u_t.$$
Здесь
$$\alpha(L) = (1-L)^2, \quad \theta(L) = 1-\theta_1 L - \theta_2 L^2.$$
Тогда
$$\psi(L) = [(1-L)^2]^{-1}(1-\theta_1 L - \theta_2 L^2) =$$
$$= (1+L+L^2+\ldots)(1+L+L^2+\ldots)(1-\theta_1 L - \theta_2 L^2) =$$
$$= 1+\psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \ldots$$
по определению. Найдем ψ_i. Сначала,
$$(1+L+L^2+\ldots)(1+L+L^2+\ldots)(1-\theta_1 L - \theta_2 L^2) = \sum_{i=0}^{\infty}(i+1)L^i.$$
Тогда если
$$\left(\sum_{i=0}^{\infty}(i+1)L^i\right)(1-\theta_1 L - \theta_2 L^2) = \sum \psi_i L^i,$$
то для коэффицентов верно
$$\psi_i = (i+1) - i\theta_1 - (i-1)\theta_2 = 1+\theta_2 + i(1-\theta_1-\theta_2).$$
Отсюда получим, что
$$\mathbb{D}(e_{T,h}) = \sigma^2\Big(1+(h-1)(1+\theta_2)^2 + \frac{h(h-1)(2h-1)(1-\theta_1^2-\theta_2^2)}{6} +$$
$$+h(h-1)(1+\theta_2)(1-\theta_1-\theta_2)\Big).$$
В этом примере
$$\lim_{h\to\infty}\mathbb{D}(e_{T,h}) = \infty.$$

Поведение корреляционной функции в модели ARIMA

Вспомним, что в модели $\mathsf{ARMA}(p,q)$ корреляционная функция может вести себя произвольным образом, но затем экспоненциально стремится к нулю. Так как модель $\mathsf{ARIMA}(p,d,q)$ является обобщением модели ARMA, ее корреляционная функция тоже может вести себя таким образом. Но, вообще говоря, корреляционная функция ARIMA не обязательно экспоненциально угасает.

Пример Пусть дан процесс $\mathsf{ARMA}(1,1)$ с уравнением
$$X_T - \phi X_{t-1} = u_t - \theta u_{t-1}.$$
Выше было показано, что функции
$$\rho_h = \mathsf{corr}(X_t, X_{t+h})$$

имеют следующий вид:

$$\begin{cases} \rho_1 = (1-\phi\theta)(\phi-\theta)/(1-\theta^2-2\phi\theta), \\ \rho_h = \phi\rho_{h-1}, \quad h > 1. \end{cases}$$

Теперь, устремим $\phi \to 1$. Получим

$$\mathsf{ARMA}(1,1) \to \mathsf{ARIMA}(0,1,1)$$

с уравнением

$$X_T - X_{t-1} = u_t - \theta u_{t-1}.$$

Тогда при $\phi \to$

$$\rho_1 \to \frac{(1-\theta)^2}{(1-\theta)^2} = 1,$$
$$\rho_h = 1,$$

а значит,

$$\mathsf{corr}(X_t, X_{t+h}) \equiv 1,$$

то есть ковариационная функция процесса $\mathsf{ARIMA}(0,1,1)$ не стремится к 0.

Дробно-интегрированная модель $\mathsf{ARFIMA}(p,d,q)$

В модели ARFIMA считаем $0 < d < 1$.

Определим $\nabla^d X$ для нецелых d следующим образом:

$$\nabla^d = (1_L)^d = \sum_{k=0}^{\infty} C_d^k (-L)^k = 1 - dL - \frac{d(1-d)}{2}L^2 - \frac{d(1-d)(2-d)}{3!}L^3 - \dots$$

Эта формула позволяет определить ∇^d при нецелых d.

При $0 < d < 1/2$ корреляции

$$\rho_h = \mathsf{corr}(X_t, X_{t+h})$$

убывают как $1/h$, так называемый «процесс с долгой памятью». При $d > 1/2$ процесс нестационарен, а его дисперсия бесконечна.

То есть, если имеем процесс с долго убывающими (не экспоненциально) корреляциями, можем воспользоваться моделью ARFIMA.

Лекция 9

Применение модели ARIMA для процессов доходности

Рассмотри теперь вопрос о применимости моделей ARMA для задач финансовой математики. Заметим, что белый шум, используемый в уравнении процесса, определялся ранее как в узком, так и в широком смысле.

Определение 9.1. Случайный процесс X_t называется *линейным*, если он представим в виде

$$X_t \overset{\text{ср.кв.}}{=} \sum_{i=0}^{\infty} \delta_i u_{t-i},$$

где $\{\delta_i\}_{i=0}^{\infty} \in \mathbb{R}$, u_t – белый шум в узком смысле.

Теорема 9.1. *Стационарный в широком смысле процесс X_t, такой, что*

$$\lim_{n\to\infty} \mathbb{E}(X_{t+n}|X_t) = \mathbb{E}X_t,$$

допускает представление

$$X_t = c + \sum_{i=0}^{\infty} \delta_i u_{t-i}, \tag{36}$$

где u_t – белый шум в широком смысле. [18]

Заметим, что класс стационарных в широком смысле процессов довольно широк, поэтому искать в нем нужный нам процесс – задача довольно трудоемкая. В этом классе есть гораздо более узкий класс линейных процессов. Возникает вопрос – при каких условиях процесс X_t, с которым работаем, будет находиться в этом классе линейных процессов? Если окажется, что процесс нелинеен, придется искать какое-то другое небольшое подмножество, в котором будем оценивать его параметры.

Далее покажем, что процесы доходности не являются линейными, и поэтому классическая теория авторегрессии для процессов доходности неприменима и необходимо ее модифицировать.

Нелинейность процессов доходности

Пусть X_t – линейный процесс. Тогда $\exists\ \{\delta_i\}_{i=0}^{\infty}$ таких, что

$$X_t - \mu = \sum_{i=0}^{\infty} \delta_i u_{t-i},$$

где u_t – белый шум в узком смысле, а

$$\mu = \mathbb{E}X_t.$$

[18] Данная теорема показывает, что модель авторегрессии ARMA никак не будет ограничивать класс стационарных в широком смысле процессов, так как все процессы ARMA допускают представление (36).

Будем считать, что

$$\delta_0 = 1, \quad \mathbb{E}u_t^4 = \lambda, \ \mathbb{E}u_t^2 = 1, \mathbb{E}X_t^4 < \infty.$$

Обозначим

$$\rho_X(\tau) = \frac{\mathsf{cov}(X_t, X_{t+\tau})}{\sqrt{\mathbb{D}X_t \mathbb{D}X_{t+\tau}}} = \frac{\mathbb{E}(X_t - \mu)(X_{t+\tau} - \mu)}{\sum_{i=0}^{\infty} \delta_i^2}$$

– коэффицент корреляции X_t,

$$K = \frac{\mathbb{E}(X_T - \mu)^4}{(\mathbb{E}(X_t - \mu)^2)^2}$$

– эксцесс процесса X_t, а

$$S_t = (X_t - \mu)^2$$

– центрированный квадрат.

Рассмотрим коэффицент корреляции квадрата процесса X_t

$$\rho_S(\tau) = \frac{\mathsf{cov}(S_t, S_{t+\tau})}{\sqrt{\mathbb{D}S_t \mathbb{D}S_{t+\tau}}}.$$

Теорема 9.2. *(об автокорреляциях линейного процесса) В сделанных предположения справедлива формула*

$$\rho_S(\tau) = \frac{2}{K-1}\rho_X^2(\tau) + \frac{K-3}{K-1}a_\tau,$$

где

$$a_\tau = \frac{\sum_{i=0}^{\infty} \delta_i^2 \delta_{i+\tau}^2}{\sum_{i=0}^{\infty} \delta_i^4}.$$

Доказательство Имеем

$$S_t = (X_t - \mu)^2 = \sum_{i=0}^{\infty} \delta_i^2 u_{t-i}^2 + 2\sum \sum_{0 \le i < j} \delta_i \delta_j u_{t-i} u_{t-j}.$$

Тогда

$$S_{t+\tau} = \sum_{k=0}^{\infty} \delta_k^2 u_{t+\tau-k}^2 + 2\sum \sum_{0 \le k < l} \delta_k \delta_l u_{t+\tau-k} u_{t+\tau-l}.$$

Отсюда

$$\mathbb{E}S_t = \sum_{i=0}^{\infty} \delta_i^2.$$

Теперь докажем следующую лемму.

Лемма 9.1.

$$\mathsf{cov}(S_t, S_{t+\tau}) = (\lambda - 3)\sum_{i=0}^{\infty} \delta_i^2 \delta_{t+\tau}^2 2 \left(\sum_{i=0}^{\infty} \delta_i \delta_{i+\tau}\right)^2.$$

Доказательство Обозначим

$$f(i,j,k,l) = \mathbb{E}u_{t-i}u_{t-j}u_{t-k}t_{t-l}.$$

Тогда

$$f(i,j,k,l) = \begin{cases} \lambda, & \text{если } i=j=k=l, \\ 1, & \text{если } \{i,j,k,l\} \text{ распадается на 2 набора разных индексов}, \\ 0, & \text{иначе.} \end{cases}$$

Если $i<j$ и $k<l$, то

$$f(i,j,k,l) = \begin{cases} 1, & \text{если } i=k, \quad j=l, \\ 0, & \text{иначе.} \end{cases}$$

Имеем

$$\mathbb{E}S_tS_{t+\tau} = \left(\sum_{i=0}^{\infty}\delta_i^2u_{t-i}^2 + 2\sum\sum_{i<j}\delta_i\delta_ju_{t-i}u_{t-j}\right)\times$$

$$\times\left(\sum_{k=0}^{\infty}\delta_k^2u_{t+\tau-k}^2 + 2\sum\sum_{k<l}\delta_k\delta_lu_{t+\tau-k}u_{t+\tau-l}\right) =$$

$$= \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\delta_i^2\delta_k^2\underbrace{f(i,j,k-\tau,k-,tau)}_{=\lambda,\text{ если } i=k-\tau,\quad 1\text{ иначе}} + \tag{37}$$

$$+2\sum_{i}\sum_{k<l}\delta_i^2\delta_k\delta_l\underbrace{f(i,i,k-\tau,l-\tau)}_{=0} + \tag{38}$$

$$+2\sum_{i<j}\sum_{k}\delta_i\delta_j\delta_k^2\underbrace{f(i,j,k-\tau,k-\tau)}_{=0} + \tag{39}$$

$$+4\sum_{i<j}\sum_{k<l}\delta_i\delta_j\delta_k\delta_l\underbrace{f(i,j,k-\tau,l-\tau)}_{=1\text{ при } i=k-\tau,\, j=l-\tau}. \tag{40}$$

Имеем, что (37) равно

$$\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 - \sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 + \left[\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 + \sum_{i=0}^{\infty}\sum_{k=0,k+\tau}^{\infty}\delta_i^2\delta_k^2\right] =$$

$$= (\lambda-1)\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 + \left(\sum_{i=0}^{\infty}\delta_i^2\right)\left(\sum_{k=0}^{\infty}\delta_k^2\right) = (\lambda-1)\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 + \left(\sum_{i=0}^{\infty}\delta_i^2\right)^2.$$

Слагаемые (38) и (39) равны 0. Слагаемое (40) равно

$$4\sum_{i<j}\delta_i\delta_j\delta_{i+\tau}\delta_{j+\tau} = 2\sum_{i<j}\delta_i\delta_j\delta_{i+\tau}\delta_{j+\tau} + 2\underbrace{\sum_{j<i}}_{\text{замена } i\leftrightarrow j}\delta_i\delta_j\delta_{i+\tau}\delta_{j+\tau} =$$

$$= 2\sum_{i,j=0}^{\infty}\delta_i\delta_j\delta_{i+\tau}\delta_{j+\tau} - 2\sum_{i=j}^{\infty}\delta_i\delta_j\delta_{i+\tau}\delta_{j+\tau} = 2\left(\sum_i \delta_i\delta_{i+\tau}\right)\left(\sum_j \delta_j\delta_{j+\tau}\right) - 2\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 =$$

$$= 2\left(\sum_i \delta_i\delta_{i+\tau}\right)^2 - 2\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2.$$

Тогда получим, что

$$\mathbb{E}S_tS_{t+\tau} = (\lambda-1)\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 + \left(\sum_{i=0}^{\infty}\delta_i^2\right) +$$

$$+2\left(\sum_i \delta_i\delta_{i+\tau}\right)^2 - 2\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 = (\lambda-3)\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 + 2\left(\sum \delta_i\delta_{i+\tau}\right)^2 + \left(\sum \delta_i^2\right)^2.$$

И, наконец, можем вычислить

$$\mathsf{cov}(S_t, S_{t+\tau}) = \mathbb{E}S_tS_{t+\tau} - (\mathbb{E}S_t)(\mathbb{E}S_{t+\tau}) =$$

$$= (\lambda-3)\sum \delta_i^2\delta_{i+\tau}^2 + 2\left(\sum \delta_i\delta_{i+\tau}\right)^2 + \left(\sum \delta_i^2\right)^2 - \underbrace{\left(\sum \delta_i^2\right)^2}_{=(\mathbb{E}S_t)^2=\mathbb{E}S_t\mathbb{E}S_{t+\tau}} =$$

$$= (\lambda-3)\sum \delta_i^2\delta_{i+\tau}^2 + 2\left(\sum \delta_i\delta_{i+\tau}\right)^2.$$

Лемма доказана.

Вернемся к доказательству теорему. Получим, что

$$\rho_S(\tau) = \frac{\mathsf{cov}(S_t, S_{t+\tau})}{\sqrt{2\mathbb{D}S_t\mathbb{D}S_{t+\tau}}} = \frac{(\lambda-3)\sum \delta_i^2\delta_{i+\tau}^2 + 2\left(\sum \delta_i\delta_{i+\tau}\right)^2}{(\lambda-3)\sum \delta_i^4 + 2\left(\sum \delta_i^2\right)^2}.$$

Докажем следующую лемму.

Лемма 9.2.

$$(\lambda-3) = \frac{(K-3)(\sum_i \delta_i^2)}{\sum_i \delta_i^4}.$$

Доказательство Вычислим

$$K = \frac{\mathbb{E}(X_t-\mu)^4}{\left(\mathbb{E}(X_t-\mu)^2\right)^2}.$$

Для знаменателя этой дроби верно

$$\left(\mathbb{E}\underbrace{(X_t-\mu)^2}_{=S_t}\right)^2 = (\mathbb{E}S_t)^2 = \left(\sum_{i=0}^{\infty}\delta_i^2\right)^2.$$

Распишем числитель:

$$\mathbb{E}(X_t-\mu)^4=\mathbb{E}\left(\sum_{i=0}^{\infty}\delta_i u_{t-i}\right)^4=\sum_{i,j,k,l=0}\infty\delta_i\delta_j\delta_k\delta_l\underbrace{f(i,j,k,l)}_{=\mathbb{E}u_{t-i}u_{t-j}u_{t-k}u_{t-l}}=$$

$$=\underbrace{\lambda\sum_{i=0}^{\infty}\delta_i^4}_{(1*)}+\underbrace{3\sum_{i=0}^{\infty}\sum_{m=0}^{\infty}\delta_i^2\delta_m^2}_{(2*)}-\underbrace{3\sum_{i=0}^{\infty}\delta_i^4}_{(3*)}.$$

Такой переход обусловлен следующими соображениями. Для слагаемого (1*) можно считать при $i=j=k=l$

$$\lambda=f(i,j,k,l),$$

для разности (2*)-(3*) можем заметить, что если $\{i,j,k,l\}$ разбивается одним из трех способов

$$\begin{cases}(i=j)\neq(k=l),\\(i=k)\neq(j=l),\\(i=l)\neq(j=k),\end{cases}$$

то

$$f(i,j,k,l)=1.$$

Слагаемое (3*) возникает, так как в сумму (2*) трижды выходит случай $i=j=k=l$ и это надо компенсировать.

Следовательно,

$$\mathbb{E}(X_t-\mu)^4=(\lambda-3)\sum_{i=0}^{\infty}\delta_i^4+3\left(\sum_{i=0}^{\infty}\delta_i^2\right)^2.$$

Тогда

$$K=\frac{(\lambda-3)\sum\delta_i^4+3\left(\sum\delta_i^2\right)^2}{\left(\sum\delta_i^2\right)^2}=\frac{\lambda-3)\sum\delta_i^4}{\left(\sum\delta_i^2\right)^2}+3.$$

Окончательно получим, что

$$(K-3)=\frac{\lambda-3)\sum\delta_i^4}{\left(\sum\delta_i^2\right)^2}.$$

Лемма доказана.

Лемма 9.3.

$$\rho_X(\tau)=\frac{\sum_{i=0}^{\infty}\delta_i\delta_{i+\tau}}{\sum_{i=0}^{\infty}\delta_i^2}.$$

Доказательство Распишем

$$\rho_X(\tau)=\frac{\mathsf{cov}(X_t,X_{t+\tau})}{\sqrt{\mathbb{D}X_t\mathbb{D}X_{t+\tau}}}=\frac{\mathbb{E}(X_t-\mu)(X_{t+\tau}-\mu)}{\mathbb{D}X_t}=\frac{\mathbb{E}\left(\sum_i\delta_i u_{t-i}\right)\left(\sum_j\delta_j u_{t+\tau-j}\right)}{\mathbb{E}S_t}=$$

$$= \frac{\sum_i \sum_j \delta_i \delta_j \mathbb{E} u_{t-i} u_{t+\tau-j}}{\sum \delta_i^2} = \frac{\sum_{i=0}^{\infty} \delta_i \delta_{i+\tau}}{\sum \delta_i^2}.$$

Лемма доказана.

Теперь, согласно леммам 9.1-9.3 имеем

$$\rho_S(t) = \frac{(\lambda - 3) \sum_i \delta_i^2 \delta_{i+\tau}^2 + 2 \left(\sum_i \delta_i \delta_{i+\tau}\right)^2}{(\lambda - 3) \sum_i \delta_i^4 + 2 \left(\sum_i \delta_i^2\right)^2}.$$

Воспользовавшись выражением для $(\lambda - 3)$ по лемме 9.2, получим

$$\rho_S(t) = \frac{(K-3) \frac{\left(\sum \delta_i^2\right)^2}{\sum \delta_i^4} \sum_i \delta_i^2 \delta_{i+\tau}^2 + 2 \left(\sum_i \delta_i \delta_{i+\tau}\right)^2}{(K-3) \frac{\left(\sum \delta_i^2\right)^2}{\sum \delta_i^4} \sum_i \delta_i^4 + 2 \left(\sum_i \delta_i^2\right)^2} =$$

$$= \frac{(K-3) \frac{\sum \delta_i^2 \delta_{i+\tau}^2}{\sum \delta_i^4} + \frac{2\left(\sum_i \delta_i \delta_{i+\tau}\right)^2}{\left(\sum \delta_i^2\right)^2}}{(K-3) + 2}.$$

По лемме 9.3 получим, что

$$\rho_S(t) = \frac{(K-3) a_\tau + 2 \left[\rho_X(\tau)\right]^2}{K-1} = \frac{2}{K-1} \rho_X^2(\tau) + \frac{K-3}{K-1} a_\tau.$$

Теорема доказана.

Лекция 10

Нелинейность процессов доходности

В прошлый раз доказали теорему о том, что корреляция квадрата линейного процесса выражается через корреляцию самого линейного процесса по формуле

$$\rho_S(t) = \frac{(K-3)a_\tau + 2\left[\rho_X(\tau)\right]^2}{K-1} = \frac{2}{K-1}\rho_X^2(\tau) + \frac{K-3}{K-1}a_\tau,$$

где

$$S_t = (X_t - \mu)^2.$$

Следствие 1. *Если u_t – гауссовский белый шум, то*

$$\rho_S(\tau) = \rho_X^2(\tau).$$

Доказательство Если $u_t \sim \mathcal{N}(\cdot,\cdot)$, то $X_t \sim \mathcal{N}(\cdot,\cdot)$. Тогда $K = 3$, а значит,

$$\frac{2}{K-1} = 1, \quad \frac{K-3}{K-1} = 0.$$

Следствие доказано.

Пусть X_t – линейный процесс с гауссовским белым шумом. Тогда

$$\sum_{\tau=1}^{30} \rho_S(\tau) = \sum_{\tau=1}^{30} \rho_X^2(\tau).$$

Но статистика процессов доходности этому противоречит. Для процессов доходности

$$\sum_{\tau=1}^{30} \rho_X^2(\tau) < 0.33,$$

$$\sum_{\tau=1}^{30} \rho_S(\tau) = \begin{cases} 0.48 - 4.99 \text{ для акций,} \\ 1.42 - 6.26 \text{ для других ценных бумаг.} \end{cases}$$

Отсюда видно, что процессы доходности не могу смоделированы с помощью класса процессов с гауссовским белом шумом.

Покажем, что они вообще не могут быть смоделированы с помощью линейных процессов.

Для этого докажем следующую лемму.

Лемма 10.1. *Коэффициенты*

$$a_\tau = \frac{\sum \delta_i^2 \delta_{i+\tau}^2}{\sum \delta_i^4}$$

удовлетворяют неравенствам

$$a_\tau \geq 0, \quad \forall\, \tau > 0, \quad \sum_{\tau=1}^{\infty} a_\tau \leq \frac{\theta}{(1-\theta)^2},$$

где

$$\theta = \frac{\mathbb{D}X_t - \mathbb{D}u_t}{\mathbb{D}X_t} = 1 - \frac{1}{\mathbb{D}X_t}.$$

Доказательство Заметим, что можно записать

$$\sum_{\tau=1}^{\infty}\sum_{i=0}^{\infty}\delta_i^2\delta_{i+\tau}^2 = \frac{1}{2}\left[\left(\sum\delta_i^2\right)^2 - \sum\delta_i^4\right].$$

Отсюда следует, что

$$\sum_{\tau=1}^{\infty} a_\tau = \frac{1}{2}\frac{\left(\sum\delta_i^2\right)^2 - \sum\delta_i^4}{\sum\delta_i^4}.$$

Кроме того,

$$\sum\delta_i^2 = \frac{1}{1-\theta}, \quad \sum\delta_i^4 \geq 2,$$

так как $\delta_0 = 1$.

Отсюда

$$\sum_{\tau=1}^{\infty} a_\tau \leq \frac{1}{2}\left[\left(\frac{1}{1-\theta}\right)^2 - 1\right] \leq \frac{\theta}{(1-\theta)^2}.$$

Лемма доказана.

Лемма 10.2. *При $K \geq 3$ справедлива оценка*

$$0 \leq \sum_{\tau=1}^{N}\rho_S(\tau) \leq \max\left\{\sum_{\tau=1}^{N}\rho_X^2(\tau), \frac{\theta}{(1-\theta)^2}\right\}.$$

Доказательство По теореме 9.2 имеем

$$\rho_S(\tau) = \frac{2}{K-1}\rho_X^2(\tau) + \frac{K-3}{K-1}a_\tau.$$

По лемме 10.1,

$$\sum_{\tau=1}^{\infty} a_\tau \leq \frac{\theta}{(1-\theta)^2}, \quad a_\tau \geq 0.$$

Следовательно,

$$\rho_S(\tau) \geq 0, \quad \forall\,\tau \geq 1,$$

так как $K \geq 3$.

Имеем

$$\sum_{\tau=1}^{N}\rho_X(\tau) = \alpha\sum_{\tau=1}^{N}\rho_X^2(\tau) + \beta\sum_{tau=1}^{N} a_\tau,$$

где

$$\alpha = \frac{2}{K-1} \geq 0, \quad \beta = \frac{K-3}{K-1} \geq 0,$$

и $\alpha + \beta = 1$. Знаем, что, если

$$C = \alpha A + \beta B, \quad \alpha, \beta \geq 0, \quad \alpha + \beta = 1,$$

то

$$C \in [A, B] \text{ или } C \in [B, A]$$

в зависимости от того, какое из этих чисел больше.

Значит,

$$\sum_{\tau=1}^{N} \rho_S(\tau) \leq \max\{\sum_{\tau=1}^{N} \rho_X^2(\tau),\ \sum_{\tau=1}^{N} a_\tau\},$$

и, так как

$$a_\tau \geq 0, \quad \forall\, \tau,$$

и

$$\sum_{\tau=1}^{\infty} \leq \frac{\theta}{(1-\theta)^2},$$

окончательно получаем

$$\sum_{\tau=1}^{N} \rho_S(\tau) \leq \max\{\sum_{\tau=1}^{N} \rho_X^2(\tau),\ \frac{\theta}{(1-\theta)^2}\}.$$

Лемма доказана.

Обратимся теперь к поставленному вопросу. Статистика процессов доходности:

$$K > 3, \quad \theta \leq 0.05,$$

$$|\rho_X(1)| \leq 0.15,$$

$$\rho_X(\tau)| \leq 0.1, \quad \forall\, \tau > 1.$$

Если бы процесс доходности был линейным, то отсюда следовало бы, что

$$\rho_S(1) \leq 0.06,$$

$$\sum_{\tau=1}^{30} \rho_S(\tau) \leq 0.33.$$

Это противоречит статистике процессов доходности. Таким образом, мы показали, что процессы доходности не являются линейными.

Процессы класса $\mathsf{ARCH}(p)$

Класс ARCH – модели авторегрессии с условной неоднородностью, то есть стационарные модели с непостоянной условной дисперсией, которая ведет себя по авторегрессионному закону.

Определение 10.1. Пусть X_t – случайная последовательность такая, что

$$\mathbb{E}X_t = 0, \quad \mathbb{E}X_t^2 < N, \quad \forall\, t.$$

Говорят, что $X_t \in \mathsf{ARCH}(p)$, где $p \in \mathbb{N}$, если

$$\exists\ \alpha_0, \alpha_1, \ldots, \alpha_p \in \mathbb{R}, \quad \alpha_p \neq 0,$$

такие, что

$$\mathbb{E}(X_t|\mathcal{F}_{t-1}^X) = 0, \quad \mathbb{E}(X_t^2|\mathcal{F}_{t-1}^X) = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2,$$

где

$$\mathcal{F}_s^X = \sigma(X_s, X_{s-1}, X_{s-2}, \ldots).$$

Примеры

- Рассмотрим белый шум в широком смысле ϵ_t такой, что, во-первых,

$$\mathbb{E}(\epsilon_t|\mathcal{F}^{\epsilon}_{t-1}) = 0,$$

а во-вторых,

$$\epsilon_t^2 \in \mathsf{AR}(p),$$

то есть

$$\epsilon_t^2 = \alpha + 0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + u_t,$$

где u_t – белый шум в узком смысле, u_t не зависит от $\mathcal{F}^{\epsilon}_{t-1}$. Тогда

$$\mathbb{E}(\epsilon_t^2|\mathcal{F}^{\epsilon}_{t-1}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2.$$

Поскольку

$$\mathbb{E}(\epsilon_{t-i}^2|\mathcal{F}^{\epsilon}_{t-1}) = \epsilon_{t-i}^2,$$

так как $\epsilon \sim \mathcal{F}^{\epsilon}_{t-1}$, а

$$\mathbb{E}(u_t|\mathcal{F}^{\epsilon}_{t-1}) = 0,$$

так как u_t независимо от $\mathcal{F}^{\epsilon}_{t-1}$. Таким образом,

$$\epsilon_t \in \mathsf{ARCH}(p).$$

- Рассмотрим случайный процесс X_t такой, что

$$\mathbb{E}X_t = \mu = \text{const},$$

а

$$X_t - \mu = v_t \cdot u_t,$$

где u_t – белый шум в узком смысле, не зависящий от прошлого, то есть от $\mathcal{F}^{X}_{t-1}$, а $v_t \sim \mathcal{F}^{X}_{t-1}$, то есть измерим относительно данной σ-алгебры, и выражается с помощью следующей формулы:

$$v_t = \left\{ \alpha_0 t + \sum_{i=1}^{p} \alpha_i (X_{t-i} - \mu)^2 \right\}^{1/2}.$$

Тогда, во-первых,

$$\mathbb{E}(X_t - \mu|\mathcal{F}^{X}_{t-1}) = \mathbb{E}(v_t u_t|\mathcal{F}^{X}_{t-1}) = v_t \mathbb{E}(u_t|\mathcal{F}^{X}_{t-1}) = v_t \mathbb{E}u_t = 0,$$

и, во-вторых,

$$\mathbb{E}((X_t - \mu)^2|\mathcal{F}^{X}_{t-1}) = \mathbb{E}(v_t^2 u_t^2|\mathcal{F}^{X}_{t-1}) = v_t^2 \mathbb{E}(u_t^2|\mathcal{F}^{X}_{t-1}) = v_t \mathbb{E}u_t^2 = v_t^2,$$

так как $\mathbb{E}u_t^2 = 1$.

Отсюда следует, что если

$$y_t = X_t - \mu,$$

то

$$\mathbb{E}(y_t|\mathcal{F}^y_{t-1}) = 0, \quad \mathbb{E}(y_t^2|\mathcal{F}^y_{t-1}) = v_t^2 = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i y_{t-i}^2.$$

Значит, $y_t \in \mathsf{ARCH}(p)$.

Лекция 11

ARCH(1)-модель с условным гауссовским распределением

Рассмотрим случайный процесс y_t, определяемый уравнением

$$y_t = \sqrt{a_0 + a_1 y_{t-1}^2}\epsilon_t, \quad t \in \mathbb{N},$$

где y_t, ϵ_t определены на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ с фильтрацией $\{\mathcal{F}_t\}_{t\in\mathbb{N}}$,[19] причем y_t и ϵ_t согласованы с фильтрацией, то есть $y_t, \epsilon_t \sim \mathcal{F}_t$.

Предположим, что закон распределения ϵ_t при условии σ-алгебры $\mathcal{F}_{t-1}$

$$\mathrm{Law}(\epsilon_t|\mathcal{F}_{t-1}) = \mathcal{N}(0,1).$$

Кроме того,

$$a_0, a_1 \in \mathbb{R}, \quad a_0, a_1 \geq 0, \quad a_1 < 1.$$

Тогда

$$\mathrm{Law}(y_t|\mathcal{F}_{t-1}) = \mathcal{N}(0, a_0 + a_1 y_{t-1}^2).$$

Условная дисперсия

$$\mathbb{D}\left(y_t|\mathcal{F}_{t-1}\right) = \mathbb{E}\left(y_t^2|\mathcal{F}_{t-1}\right) - \left(\left(y_t|\mathcal{F}_{t-1}\right)\right)^2 =$$

$$= a_0 + a_1 y_{t-1}^2,$$

так как

$$\mathbb{E}\left(y_t|\mathcal{F}_{t-1}\right) = \mathbb{E}\left(\sqrt{a_0 + a_1 y_{t-1}^2}\epsilon_t|\mathcal{F}_{t-1}\right) =$$

$$= \sqrt{a_0 + a_1 y_{t-1}^2}\mathbb{E}\left(\epsilon_t|\mathcal{F}_{t-1}\right) = 0,$$

а

$$\mathbb{E}\left(y_t^2|\mathcal{F}_{t-1}\right) = a_0 + a_1 y_{t-1}^2.$$

Таким образом,

$$y_t \in \mathsf{ARCH}(1).$$

Теорема 11.1. *1. Выполнено*

$$\mathbb{E}\left(y_t^{2r}|\mathcal{F}_{t-1}\right) = (a_0 + a_1 y_{t-1}^2)^r (2r-1)!!,$$

где $(2r-1)!! = 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2r-1)$.

2. Если $a_0 > 0$, *то* $\exists\, r_0$ *такое, что*

$$\forall\, r > r_0 \ \mathbb{E} y_t^{2r} \to \infty, \quad t \to \infty.$$

[19] Напомним, что, говоря неформально, фильтрация $\{\mathcal{F}_t\}_{t\in\mathbb{N}}$ – такой поток σ-алгебр, индексированных временным параметром и моделирующих поток событий, доступных в момент времени t и для $t_1 < t_2$

$$\mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}\mathcal{F}.$$

3.

$$\mathbb{E}\left(y_t^2\right) \to \frac{a_0}{1-a_1}, \quad t \to \infty,$$

4.

$$\mathbb{E}y_t^4 = \frac{3a_0^2}{(1-a_1)^2}\frac{1-a_1^2}{1-3a_1^2}, \quad \textit{если } a_1 < \frac{1}{\sqrt{3}}.$$

Доказательство Обозначим

$$\mathbb{E}\left(\epsilon_t^{2k}|\mathcal{F}_{t-1}\right) = m^{(2k)}.$$

Тогда

$$\mathbb{E}\left(y_t^{2k}|\mathcal{F}_{t-1}\right) = \mathbb{E}\left(\underbrace{\left(a_0 + a_1 y_{t-1}^2\right)^k}_{\sim\mathcal{F}_{t-1}} \epsilon_t^{2k}|\mathcal{F}_{t-1}\right) =$$

$$= \left(a_0 + a_1 y_{t-1}^2\right)^k \mathbb{E}\left(\epsilon_t^{2k}|\mathcal{F}_{t-1}\right) = \left(a_0 + a_1 y_{t-1}^2\right)^k m^{(2k)}.$$

После интегрирования по частям имеем

$$m^{(2k)} = \frac{1}{2k+1} m^{(2k+2)}, \quad m^{(0)} = 1.$$

Тогда

$$m^{(2k)} = \prod_{j=1}^{k}(2j-1) = (2k-1)!!$$

Отсюда следует пункт 1 теоремы.

Докажем пункт 2. Фиксируем $r \in N$. Определим векторный процесс

$$\overline{w_t} = (w_t^1, \ldots, w_t^r) = (y_t^{2r}, y_t^{2(r-1)}, \ldots, y_t^2),$$

то есть

$$w_t^i = y_t^{2(r-i+1)}.$$

Имеем

$$\mathbb{E}(W_t^k|\mathcal{F}_{t-1}) = \mathbb{E}(y_t^{2(r-k+1)}|\mathcal{F}_{t-1}.$$

Используя пункт 1 данной теоремы, получаем, что

$$\mathbb{E}(W_t^k|\mathcal{F}_{t-1}) = (a_+ a_1 y_{t-1}^2)^{r-k+1}(2(r-k+1)-1)!! =$$

$$= \left(\sum_{i=0}^{r-k+1} \underbrace{y_{t-1}^{2i}}_{w_{t-1}^{r-i+1} \text{ при } i>0} C_{r-k+1}^i a_0^{r-k+1-i} a_1^i\right)(2(r-k+1)-1)!!.$$

Сделаем замену $i' = r - i + 1$. Получим

$$\mathbb{E}(W_t^k|\mathcal{F}_{t-1}) = a_0^{r-k+1}\left(2(r-k+1)-1\right)!! + \sum_{i'=0}^{r} w_t^{i'} C_{r-k+1}^{r-i'+1} a_0^{i'-k} a_1^{r-i'+1}\left(2(r-l+1)-1\right)!!$$

МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

В итоге получаем

$$\mathbb{E}(W_t^k|\mathcal{F}_{t-1}) = \overline{b} + A\overline{w_{t-1}}, \tag{41}$$

где

$$\overline{b} = (b^1, \ldots, b^r),$$
$$b^k = a_0^{r-k+1}\left(2(r-k+1)-1\right)!!,$$

A – верхнетреугольная матрица с диагональными элементами

$$A_{kk} \underbrace{=}_{i'=k} C_{r-k+1}^{r-k+1} a_0^{k-k} a_1^{r-k+1}\left(2(r-k+1)-1\right)!! =$$

$$= a_1^{r-k+1}\left(2(r-k+1)-1\right)!!, \quad 1 \le k \le r.$$

Возьмем теперь матожидание в (41). Получим

$$\underbrace{\mathbb{E}\overline{w_t}}_{=\overline{\xi_t}} = \overline{b} + A\mathbb{E}\overline{w_{t-1}}.$$

Тогда получим, что

$$\overline{\xi_t} = \overline{b} + A\overline{\xi_{t-1}}.$$

Отсюда многократной подстановкой получаем, что

$$\overline{\xi_t} = \overline{b} + A\overline{b} + A^2\overline{b} + \ldots + A^{t-1}\overline{b} + A^t\overline{\xi_0}.$$

Так как $b_i \ge 0$, $A_{ij} \ge 0$, $\forall\, i, j$, то

$$\left(A\overline{b}\right)^i = \sum_{j-1}^{r} A_{ij} b^i \ge A_{ij} b^i,$$

$$\left(A^n\overline{b}\right)^i \ge \left(A_{ij}\right)^n b^i,$$

следовательно,

$$\left(A^n\overline{b}\right)^1 \ge A_{11} b^1.$$

Имеем, во-первых,

$$b^1 = a_0^r(2r-1)!! > 0,$$

так как $a_0 > 0$, во-вторых,

$$A_{11} = a_1^r(2r-1)!! > 1, \quad r > r_0,$$

где r_0 – некоторое число, и, в-третьих,

$$\xi^1 = \mathbb{E}w_t^1 = \mathbb{E}y_t^{2r}.$$

Отсюда следвает, что

$$\mathbb{E}y_t^{2r} \ge (A_{11})^t b^1 \to \infty, \quad t \to \infty,$$

если $r > r_0$. Пункт 2 теоремы доказан.

Перейдем к доказательству пункта 3.

По условию,

$$\mathbb{E}\left(y_t^2|\mathcal{F}_{t-1}\right) = a_0 + a_1 y_{t-1}^2.$$

Беря от этого выражения матожидание, получим

$$\mathbb{E}y_t^2 = a_0 + a_1\mathbb{E}y_{t-1}^2.$$

Тогда с помощью многократной подстановки получим, что для $\forall\ n$

$$\mathbb{E}y_t^2 = \underbrace{a_0(1 + a_1 + a_1^2 + \ldots + a_1^n)}_{\to\frac{a_0}{1-a_1}, n\to\infty} + \underbrace{a_1^{n+1}\mathbb{E}y_{t-n-1}^2}_{\to 0, n\to\infty},$$

то есть

$$\mathbb{E}y_t^2 \to \frac{a_0}{1-a_1}, \quad t\to\infty.$$

Пункт 3 теоремы доказан. Доказательство пункта 4 предоставляется в качестве упражнения.

Теорема доказана.

Следствие 1. *(кластерность процесса y_t)*

$$\mathbb{E}(y_t^2|\mathcal{F}_{t-1}) - \mathbb{E}y_t^2 \overset{\text{по теореме}}{=} a_1\left(y_{t-1}^2 - \mathbb{E}y_{t-1}^2\right).$$

Доказательство Распишем

$$\mathbb{E}(y_t^2|\mathcal{F}_{t-1}) - \mathbb{E}y_t^2 = a_0 + a_1 y_{t-1}^2 - \frac{a_0}{1-a_1} = a_1 y_{t-1}^2 - \frac{a_0 - a_0 + a_1 a_0}{1-a_1} = a_1\left(y_{t-1}^2 - \mathbb{E}y_{t-1}^2\right).$$

Следствие доказано.

Остановимся подробнее на смысле следствия. Если в момент времени $t-1$

$$y_{t-1}^2 > \mathbb{E}y_{t-1}^2,$$

то, так как $a_1 > 0$,

$$\mathbb{E}(y_t^2|\mathcal{F}_{t-1}) > \mathbb{E}y_t^2.$$

Если же в момент $t-1$

$$y_{t-1}^2 < \mathbb{E}y_{t-1}^2,$$

то, так как $a_1 > 0$,

$$\mathbb{E}(y_t^2|\mathcal{F}_{t-1}) < \mathbb{E}y_t^2.$$

Это приводит к так называемой кластерности, когда процесс стремится сохранить повышенные (или пониженные) значения (рис. **??**). Это поведение характерно для процессов доходности.

Следствие 2. *Коэффицент эксцесса процесса y_t равен*

$$k = 3\frac{1-a_1^2}{1-3a_1^2} > 3, \quad a_1 > 0.$$

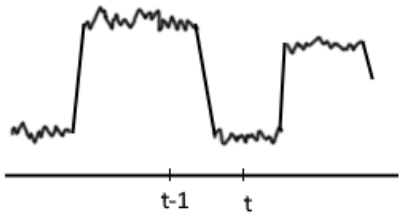

Рис. 11.1. Явление кластерности.

Доказательство Распишем

$$k = \frac{\mathbb{E}y_t^4}{(\mathbb{E}(y_t^2))^2} \overset{\text{по теореме}}{=} \frac{3a_0^2}{(1-a_1)^2}\frac{1-a_1^2}{1-3a_1^2} / \left(\frac{a_0}{1-a_1}\right)^2 =$$

$$= 3\frac{1-a_1^2}{1-3a_1^2} > 3$$

при $a_1 > 0$. Отсюда следует, что процесс y_t имеет тяжелые хвосты.[20]

Следствие доказано.

Процессы $\mathsf{GARCH}(p, q)$

Недостатком при моделировании процессов ARCH является то, что обычно получается большой порядок пространства параметров. Существенно снизить его помогает следующая модель.

Рассмотрим обобщенную модель ARCH – модель GARCH.

Определение 11.1. Процесс $y_t \in \mathsf{GARCH}(p, q)$, если

$$\exists\, c, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \in \mathbb{R}$$

такие, что, во-первых,

$$\mathbb{E}(y_t|\mathcal{F}_{t-1}^y) = 0, \quad \forall t \in \mathbb{Z}, \tag{42}$$

и, во-вторых,

$$\mathbb{D}(y_t|\mathcal{F}_{t-1}^y) = c + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \mathbb{D}(y_{t-j}|\mathcal{F}_{t-j-1}^y). \tag{43}$$

[20] Получается, что в моделях ARCH, с одной стороны, для построение используется нормальное распределение, а с другой, эти модели обладают свойствами реальных процессов доходности, например, тяжелыми хвостами.

Обозначим

$$h_t = \mathbb{D}(y_t|\mathcal{F}^y_{t-1}).$$

Тогда (43) записывается в виде

$$h_t = c + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}.$$

Лекция 12

Теорема о достаточных условиях стационарности

В прошлый раз начали говорить о моделях $\mathsf{GARCH}(p,q)$.

Теорема 12.1. *Пусть процесс $y_t \in \mathsf{GARCH}(p,q)$ и его параметры c, α_i, β_j удовлетворяют условиям*

$$c \geq 0, \quad {}_i \geq 0, \quad 1 \leq i \leq q, \quad \beta_j \geq 0, \quad 1 \leq j \leq p,$$

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1.$$

Тогда y_t стационарен в широком смысле и

$$\mathbb{D}(y_t) = \mathbb{E}y_t = \frac{c}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}.$$

Доказательство Из условия (42) имеем

$$\mathbb{E}(y_t) = \mathbb{E}\left(\underbrace{\mathbb{E}(y_t|\mathcal{F}_{t-1}^y)}_{=0}\right) = 0.$$

Пусть $s \neq t$, $s < t$. Вычислим

$$\mathsf{cov}(y_t, y_s) = \mathbb{E}y_t y_s - \underbrace{\mathbb{E}y_t \mathbb{E}y_s}_{=0} = \mathbb{E}\left(y_s \underbrace{\mathbb{E}(y_t|\mathcal{F}_{t-1}^y)}_{=0}\right) = 0,$$

так как если $s < t$, то $s \leq t-1$, а значит, y_s измеримо относительно $\mathcal{F}_{t-1}^y$. Остается показать, что

$$\mathbb{D}(y_t) = \mathbb{E}y_t^2 = \text{const}.$$

Обозначим через

$$u_t = y_t^2 - h_t.$$

Тогда

$$h_t = y_t^2 - u_t = c + \sum_{i=1}^{q} \alpha_i y_{t-i}^1 + \sum_{j=1}^{p} \beta_j (y_{t-j}^2 - u_{t-j}),$$

откуда получаем, что

$$y_t^2 = c + \sum_{i=1}^{p \vee q} (\alpha_i + \beta_i) y_{t-i}^2 + u_t - \sum_{j=1}^{p} \beta_j u_{t-j},$$

где $\alpha_i = 0$, если $i > q$, и $\beta_j = 0$ при $j > p$, а $p \vee q = \max\{p, q\}$. Теперь возьмем от обеих частей матожидание. Получим

$$\mathbb{E}y_t^2 = c + \sum_{i=1}^{p \vee q} (\alpha_i + \beta_i) \mathbb{E}y_{t-i}^2, \tag{44}$$

так как

$$\mathfrak{u}_{\mathbb{t}} = \mathbb{E}y_t^2 - \underbrace{\mathbb{E}h_t}_{=\mathbb{E}y_t^2} = 0.$$

Уравнение (44) – уравнение типа $\mathsf{AR}(p \vee q)$ с характеристическим уравнением

$$x^{p\vee q} - \sum_{i=1}^{p\vee q}(\alpha_i + \beta_i)x^{p\vee q - i} = 0.$$

Обозначим $n = p \vee q$, $\gamma_i = \alpha_i + \beta_i$. Тогда характеристическое уравнение имеет вид

$$x^n - \sum_{i=1}^{n}\gamma_i x^{n-i} = 0.$$

Обозначим $f_t\mathbb{E}y_t^2$ и перепишем тогда (44)

$$f_t - \sum_{i=1}^{n}\gamma_i f_{t-i} = c.$$

Это уравнение авторегрессии. Тогда

$$\left(1 - \sum_{i=1}^{n} L^i\right) f_t = c,$$

$$\phi(L)f_t = c,$$

$$\phi(L) = 1 - \sum_{i=1}^{n}\gamma_i L^i.$$

Лемма 12.1. *Если $a = const$, то*

$$\phi^{-1}(L)a = \frac{a}{1 - \gamma_1 - \ldots - \gamma_n}.$$

Доказательство Распишем

$$\phi^{-1}(L)a = b \quad \phi(L)b = a$$

$$a = b(1 - \gamma_1 - \ldots - \gamma_n)) \quad b = \phi^{-1}(L)a = \frac{a}{1 - \gamma_1 - \ldots - \gamma_n}.$$

Лемма доказана.

Лемма 12.2.

$$\sum_{i=1}^{n}\gamma_i = \sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$$

тогда и только тогда, когда все корни характеристического уравнения

$$P(x) = x^n - \sum_{i=1}^{n}\gamma_i x^{n-i} = 0$$

по модулю меньше 1.

Доказательство Предположим сначала, что

$$\sum_{i=1}^{n} \gamma_i \geq 1.$$

Тогда

$$P(1) = 1 - \sum_{i=1}^{n} \gamma_i \leq 0.$$

При $x \to \infty$ $P(x) \to \infty$, так как коэффицент при старшем члене x^n равен $1 > 0$.

Следовательно, $\exists\, x_0 \geq 1$ такое, что $P(x_0) = 0$.

Значит, характеристическое уравнение имеет корень по модулю ≥ 1. Тогда если все корни характеристического уравнения по модулю < 1, то

$$\sum_{i=1}^{n} \gamma_i < 1.$$

В обратную сторону. Предположим теперь, что

$$\sum_{i=1}^{n} \gamma_i \geq 1.$$

Предположим обратное, что $\exists\, x_0$, $|x_0| \geq 1$,

$$P(x_0) = 0.$$

Имеем

$$x_0^n = \sum_{i=1}^{n} \gamma_i x_0^{n-i}.$$

Разделив обе части на x_0^n, получим, что

$$1 = \sum_{i=1}^{n} \gamma_i x_0^{-i}.$$

Тогда

$$1 = |\sum_{i=1}^{n} \gamma_j x_0^{-i}| \leq \sum_{i=1}^{n} |\gamma_i x_0^{-i}| = \sum_{i=1}^{n} \gamma_u \left|\frac{1}{|x_0|}\right|^i \leq \sum_{i=1}^{n} \gamma_i < 1.$$

Получили противоречие.

Лемма доказана.

Следовательно, в условиях теоремы

$$f_t = \underbrace{\mathbb{E} y_t^2}_{=\mathbb{D}(y_t)=\text{const}} = \phi^{-1}(L)c = \frac{c}{1 - \sum_{i=1}^{n} y_i} = \frac{c}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j} = \text{const}.$$

Теорема доказана.

Следствие 1. *Пусть процесс* $y_t \in \mathsf{ARCH}(p)$,

$$\mathbb{D}\left(y_t|\mathcal{F}_{t-1}^y\right) = c + \sum_{i=1}^{p} \alpha_i y_{t-i}^2.$$

Тогда, если $c \geq 0$, $\alpha_i \geq 0$, $1 \leq i \leq p$,

$$\sum_{i=1}^{p} \alpha_i < 1,$$

то y_t *стационарен в широком смысле и*

$$\mathbb{D}(y_t) = \frac{c}{1 - \sum \alpha_i}.$$

Эксцесс процесса типа $\mathsf{GARCH}(p, q)$

Найдем значение эксцесса процесса GARCH с условно-гауссовским распределением.

Пусть процесс $X_t \in \mathsf{GARCH}(p, q)$,

$$\mathrm{Law}(X_t|\mathcal{F}_{t-1}) = \mathcal{N}\left(0, \underbrace{c + \sum_{i=1}^{q} \alpha_i X_{t-1}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}}_{=h_t}\right),$$

$$h_t = \mathbb{E}(X_t^2|\mathcal{F}_{t-1}).$$

Лемма 12.3. *Эксцесс процесса* X_t $k(X_t)$ *равен*[21]

$$k(X_t) = \frac{\mathbb{E}X_t^4}{\left(\mathbb{E}X_T^2\right)^2} = 3 + 3\frac{\mathbb{E}\left[\mathbb{E}(X_t^2|\mathcal{F}_{t-1})\right]}{\left(\mathbb{E}X_t^2\right)^2}.$$

Доказательство Запишем

$$\mathbb{E}(X_t^4|\mathcal{F}_{t-1}) = 3\left[\mathbb{E}(X_t^2|\mathcal{F}_{t-1})\right]^2,$$

так как условное распределение $X_t|\mathcal{F}_{t-1}$ – гауссовское.

Возьмем математическое ожидание от обеих частей. Получим

$$\mathbb{E}X_t^4 = 3\mathbb{E}\left[\left(\mathbb{E}(X_t^2|\mathcal{F}_{t-1}\right)^2\right].$$

Кроме того,

$$\left[\mathbb{E}\left(\mathbb{E}(X_t^2|\mathcal{F}_{t-1}\right)\right]^2 = \left(\mathbb{E}X_t^2\right)^2.$$

[21] Здесь первое равенство записано по определению с учетом

$$\mathbb{E}X_t = 0.$$

Тогда

$$k(X_t) = \frac{\mathbb{E}X_t^4}{(\mathbb{E}X_T^2)^2} = 3 + 3\frac{\mathbb{E}\left[\mathbb{E}(X_t^2|\mathcal{F}_{t-1})\right]^2 - \left[\mathbb{E}\left(\mathbb{E}(X_t^2\mathcal{F}_{t-1})\right)\right]^2}{\left[\mathbb{E}X_t^2\right]^2} =$$

$$= 3 + 3\frac{\mathbb{E}\left[\mathbb{E}(X_t^2|\mathcal{F}_{t-1})\right]}{(\mathbb{E}X_t^2)^2}.$$

Лемма доказана.

Следствие 2. *Если*

$$\mathbb{E}\left[\mathbb{E}(X_t^2\mathcal{F}_{t-1})\right] > 0,$$

то

$$k(X_t) > 3,$$

то есть X_t имеет тяжелые хвосты.

Прогнозирование доходности

Рассмотрим процесс с X_t с уравнением

$$X_t = \mu + u_t v_t, \tag{45}$$

относительно которого выполняются следующие предположения:

1. u_t – гауссовский белый шум, то есть

$$Eu_t = 0, \quad \mathbb{D}u_t = 1,$$

u_t и u_s независимы при $s \neq t$ и

$$u_t \sim \mathcal{N}(0, 1).$$

2. В уравнении (45)

$$\mathbb{E}X_t = \mu.$$

Без ограничения общности будем считать[22]

$$\mathbb{E}X_T = 0.$$

3.

$$\mathbb{D}(X_t|v_t) = v_t^2.$$

[22]В практическом смысле, можно считать, что

$$\mu = \mathbb{E}X_t \ll \mathbb{D}(X_t).$$

Лекция 13

Прогнозирование доходности

В прошлый раз закончили на то, что сформулировали предположения для процесса X_t с уравнением

$$X_t = u_t v_t,$$

где u_t – гауссовский белый шум (обновляющий процесс), то есть

$$Eu_t = 0, \quad \mathbb{D}u_t = 1,$$

u_t и u_s независимы при $s \neq t$ и

$$u_t \sim \mathcal{N}(0, 1),$$

а

$$v_t = \mathbb{D}(X_t | \mathcal{F}_{t-1}).$$

Рассмотрим, как процесс v_t (волатильность) может зависеть от $\mathcal{F}_{t-1}$.

3 основных случая:

- Моменты катастроф. Здесь v_t кусочно-постоянна, меняется в моменты $t_1, t_2, \ldots$

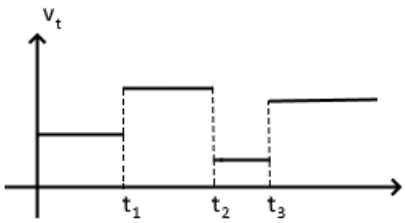

Рис. 13.1. Моменты катастроф.

  – моменты катастроф (рис. 13.1).

- Стохастическая волатильность

$$v_t = f(v_{t-1}, v_{t-2}, \ldots, \eta_t),$$

  где η_t не зависит от $\mathcal{F}_{t-1}$. η_t – *случайный толчок.*

  Например, модель Tauchen Pitts (1983). Эта модель хорошо описывает поведение рынка, но очень сложна для оценивания. Уравнение процесса здесь имеет вид

$$X_t - \mu = \sum_{i=1}^{W_t} \omega_{it},$$

  где ω_{it} – i-ый толчок в день t. Всего в день t происходит случайное количество W_t толчков финансового рынка. Относительно этой модели были приняты следующие предположения:

– ω_{it} и $\omega_j t$ независимы, $\forall\, i \neq j$, $1 \leq i, j \leq W_t$.

– ω_{it} имеет распределение $\mathcal{N}(0, \sigma_\omega^2)$.

В этих предположениях получаем, что

$$v_t^2 = \mathbb{D}(X_t) = W_t \sigma_\omega^2,$$

$$X_t \sim \mathcal{N}(\mu, \sigma_\omega^2 W_t).$$

В этой модели можно учесть объемы продаж

$$Q_t = \sum_{i=1}^{W_t} q_{it},$$

где

$$q_{it} \sim \mathcal{N}(a, b^2),$$

и, соответственно,

$$Q_t \sim \mathcal{N}(aW_t, b^2 W_t).$$

Первоначально в этой модели предполагали, что W_t – белый шум в узком смысле, но, так как тогда X_t – тоже белый шум в узком смысле, от этого предположения отказались.

Тогда предположили, что

$$u_t = \frac{X_t - \mu}{v_t} = \frac{1}{\sqrt{W_t}\sigma_\omega} \underbrace{\sum_{i=1}^{W_t} \omega_{it}}_{\sim \mathcal{N}(0, W_t \sigma_\omega^2)}.$$

Тогда

$$u_t|_{v_t} \sim \mathbb{N}(0, 1),$$

не зависит от v_t, и, значит, u_t не зависит от v_t.

Если

$$v_t = \sqrt{W_t}\sigma_\omega$$

– стационарен в широком смысле и

$$\mathbb{E}(X_t - \mu)^2 = \mathbb{E} v_t^2 u_t^2 = \mathbb{E} V_t^2 \mathbb{E} u_t^2,$$

то

$$\mathbb{E} v_t^2 = \text{const}.$$

Для эксцесса процесса X_t верно, что

$$k_X = \frac{\mathbb{E}(X_t - \mu)^4}{\left(\mathbb{E}(X_t - \mu)^2\right)^2} = \frac{\mathbb{E} v_t^4 u_t^4}{\left(\mathbb{E} u_t^2 v_t^2\right)^2} = \frac{\mathbb{E} u_t^4}{\left(\mathbb{E} u_t^2\right)^2} \frac{\mathbb{E} v_t^4}{\left(\mathbb{E} v_t^2\right)^2} \geq 3$$

Первый из сомножителей равен 3 как эксцесс нормального распределения, а второй из сомножителей ≥ 1.

Отсюда получаем, что у процесса X_t тяжелые хвосты.

Предположим, что

$$\log v_t \sim \mathcal{N}(\alpha, \beta^2),$$
$$u_t \sim (0, 1).$$

Тогда

$$k_x = 3e^{4\beta^2} > 3.$$

Модель не применима практически, так как невозможно оценить v_t на основе значений X_t.

- ARCH-модель, необязательно линейная. Смысл модели в том, что

$$v_t = f(X_{t-1}, X_{t-2}, \ldots).$$

Итак, подведем итог. Для общей модели процесса с уравнением

$$X_t = \mu + u_t v_t$$

выделяют три составляющие волатильности v_t:

- катастрофы;
- стохастическая компонента волатильности (случайные толчки рынка);
- ARCH-модель
$$v_t = f(X_{t-1}, X_{t-2}, \ldots).$$

Первые две составляющие присутствуют в реальных процессах, но оценить их средствами современной финансовой математики невозможно.

МЕХАНИКО-
МАТЕМАТИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in

ЛЕКЦИИ УЧЕНЫХ МГУ