

Анализ временных рядов в задачах и упражнениях

Борzych Д. А., Демешев Б. Б.

17 марта 2018 г.

Предисловие

План:

0. Автокорреляция в ошибках линейной модели
1. Стационарные/нестационарные процессы, ARMA сюда оператор лага, ACF/PACF
2. Экспоненциальное сглаживание и тета метод
3. ГАРЧ
4. Единичный корень (до VAR) (ADF)
5. VAR/VECM/коинтеграция
6. Midas
7. Байесовские VAR
8. Модели состояние-наблюдение/фильтр Калмана/TVP

Глава 1

Автокорреляция ошибок в линейной модели

1.1 Билл Гейтс оценил модель $y_t = \beta_1 + \beta_2 t + \beta_3 y_{t-1} + \varepsilon_t$ с помощью МНК. Значение статистики Дарбина-Уотсона оказалось равно $DW = 0.55$. Какой из этого следует вывод об автокорреляции ошибок первого порядка?

1.2 Рассмотрим модель $y_t = \beta x_t + \varepsilon_t$, где $\varepsilon_1 = u_1$ и $\varepsilon_t = u_t + u_{t-1}$ при $t \geq 2$. Случайные величины u_i независимы с $\mathbb{E}(u_i) = 0$ и $\text{Var}(u_i) = \sigma^2$.

1. Найдите $\text{Var}(\varepsilon_t)$
2. Являются ли ошибки ε_t гетероскедастичными?
3. Найдите $\text{Cov}(\varepsilon_i, \varepsilon_j)$
4. Являются ли ошибки ε_t автокоррелированными?
5. Как выглядит матрица $\text{Var}(\varepsilon)$?
6. Рассмотрим оценку

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$$

Является ли она несмещенной для β ? Является ли она эффективной в классе линейных по y несмещенных оценок?

7. Если приведенная $\hat{\beta}$ не является эффективной, то приведите формулу для эффективной оценки.

1.3 Имеются данные $y = (1, 2, 0, 0, 2, 1)$. Предполагая модель с автокоррелированной ошибкой, $y_t = \mu + \varepsilon_t$, где $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$ с помощью трёх тестов проверьте гипотезы $H_0: \rho = 0$, $H_0: \mu = 0$, $H_0: \sigma^2 = 1$

$$\begin{cases} \rho = 0 \\ \mu = 0 \\ \sigma^2 = 1 \end{cases}$$

1.4 Рассматривается модель $y_t = \mu + \varepsilon_t$, $t = 1, \dots, T$, где $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$, случайные величины $\varepsilon_0, u_1, \dots, u_T$ независимы, причем $\varepsilon_0 \sim N(0, \sigma^2/(1 - \rho^2))$, $u_t \sim N(0, \sigma^2)$. Имеются наблюдения $y' = (1, 2, 0, 0, 1)$.

1. Выпишите функцию правдоподобия

$$L(\mu, \rho, \sigma^2) = f_{Y_1}(y_1) \prod_{t=2}^T f_{Y_t|Y_{t-1}}(y_t|y_{t-1}).$$

2. Найдите оценки неизвестных параметров модели максимизируя условную функцию правдоподобия

$$L(\mu, \rho, \sigma^2 | Y_1 = y_1) = \prod_{t=2}^T f_{Y_t | Y_{t-1}}(y_t | y_{t-1})$$

- 1.5 Остаются ли в условиях автокорреляции МНК- оценки в линейной модели несмещёнными? Состоятельными?

- 1.6 Продавец мороженого оценил динамическую модель объёмов продаж:

$$\ln \hat{Q}_t = 26.7 + 0.2 \ln \hat{Q}_{t-1} - 0.6 \ln P_t$$

Здесь Q_t — число проданных в день t вафельных стаканчиков, а P_t — цена одного стаканчика в рублях. Продавец также рассчитал остатки \hat{e}_t .

1. Чему, согласно полученным оценкам, равна долгосрочная эластичность объёма продаж по цене?
 2. Предположим, что продавец решил проверить наличие автокорреляции первого порядка с помощью теста Бройша-Годфри. Выпишите уравнение регрессии, которое он должен оценить.
- 1.7 Пусть u_t — независимые нормальные случайные величины с математическим ожиданием 0 и дисперсией σ^2 . Известно, что $\varepsilon_1 = u_1, \varepsilon_t = u_1 + u_2 + \dots + u_t$. Рассмотрим модель $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$.
1. Найдите $\text{Var}(\varepsilon_t), \text{Cov}(\varepsilon_t, \varepsilon_s), \text{Var}(\varepsilon)$
 2. Являются ли ошибки ε_t гетероскедастичными?
 3. Являются ли ошибки ε_t автокоррелированными?
 4. Предложите более эффективную оценку вектора коэффициентов регрессии по сравнению МНК-оценкой.
 5. Результаты предыдущего пункта подтвердите симуляциями Монте-Карло на компьютере.
- 1.8 Ошибки в модели $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$ являются автокоррелированными первого порядка, $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$. Шаман-эконометрист Ойуун выполняет два камлания-преобразования. Поясните смысл камланий:

1. Камлание А, при $t \geq 2$, Ойуун преобразует уравнение к виду $y_t - \rho y_{t-1} = \beta_1(1 - \rho) + \beta_2(x_t - \rho x_{t-1}) + \varepsilon_t - \rho \varepsilon_{t-1}$
2. Камлание Б, при $t = 1$, Ойуун преобразует уравнение к виду $\sqrt{1 - \rho^2} y_1 = \sqrt{1 - \rho^2} \beta_1 + \sqrt{1 - \rho^2} \beta_2 x_1 + \sqrt{1 - \rho^2} \varepsilon_1$.

- 1.9 Рассмотрим модель $y_t = \beta_1 + \beta_2 x_{t1} + \dots + \beta_k x_{tk} + \varepsilon_t$, где ε_t подчиняются автокорреляционной схеме первого порядка, т.е.

1. $\varepsilon_t = \rho \varepsilon_{t-1} + u_t, -1 < \rho < 1$
2. $\text{Var}(\varepsilon_t) = \text{const}, \mathbb{E}(\varepsilon_t) = \text{const}$
3. $\text{Var}(u_t) = \sigma^2, \mathbb{E}(u_t) = 0$
4. Величины u_t независимы между собой
5. Величины u_t и ε_s независимы, если $t \geq s$

Найдите:

1. $\mathbb{E}(\varepsilon_t), \text{Var}(\varepsilon_t)$
2. $\text{Cov}(\varepsilon_t, \varepsilon_{t+h})$
3. $\text{Corr}(\varepsilon_t, \varepsilon_{t+h})$

1.10 Рассматривается модель $y_t = \beta_1 + \beta_2 x_{t1} + \dots + \beta_k x_{tk} + \varepsilon_t$. Ошибки ε_t гомоскедастичны, но в них возможно присутствует автокорреляция первого порядка, $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$. При известном числе наблюдений T на уровне значимости 5% сделайте статистический вывод о наличии автокорреляции.

1. $T = 25, k = 2, DW = 0.8$
2. $T = 30, k = 3, DW = 1.6$
3. $T = 50, k = 4, DW = 1.8$
4. $T = 100, k = 5, DW = 1.1$

1.11 По 100 наблюдениям была оценена модель линейной регрессии $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$. Оказалось, что $RSS = 120, \hat{\varepsilon}_1 = -1, \hat{\varepsilon}_{100} = 2, \sum_{t=2}^{100} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} = -50$. Найдите DW и ρ .

1.12 Применяется ли статистика Дарбина-Уотсона для выявления автокорреляции в следующих моделях

1. $y_t = \beta_1 x_t + \varepsilon_t$
2. $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$
3. $y_t = \beta_1 + \beta_2 y_{t-1} + \varepsilon_t$
4. $y_t = \beta_1 + \beta_2 t + \beta_3 y_{t-1} + \varepsilon_t$
5. $y_t = \beta_1 t + \beta_2 x_t + \varepsilon_t$
6. $y_t = \beta_1 + \beta_2 t + \beta_3 x_t + \beta_4 x_{t-1} + \varepsilon_t$

1.13 По 21 наблюдению была оценена модель линейной регрессии $\hat{y}_{(se)} = \frac{1.2}{(0.3)} + \frac{0.9}{(0.18)} \cdot y_{t-1} + \frac{0.1}{(0.01)} \cdot t$, $R^2 = 0.6, DW = 1.21$. Протестируйте гипотезу об отсутствии автокорреляции ошибок на уровне значимости 5%.

1.14 По 24 наблюдениям была оценена модель линейной регрессии $\hat{y}_{(se)} = \frac{0.5}{(0.01)} + \frac{2}{(0.02)} \cdot t$, $R^2 = 0.9, DW = 1.3$. Протестируйте гипотезу об отсутствии автокорреляции ошибок на уровне значимости 5%.

1.15 По 32 наблюдениям была оценена модель линейной регрессии $\hat{y}_{(se)} = \frac{10}{(2.5)} + \frac{2.5}{(0.5)} \cdot t - \frac{0.1}{(0.01)} \cdot t^2$, $R^2 = 0.75, DW = 1.75$. Протестируйте гипотезу об отсутствии автокорреляции ошибок на уровне значимости 5%.

Глава 2

Стационарные процессы, ARMA

2.1 Запишите процесс $y_t = 4 + 0.4y_{t-1} + 0.3\varepsilon_{t-1} + \varepsilon_t$ с помощью оператора лага.

2.2 Пусть X_t , $t = 0, 1, 2, \dots$ — случайный процесс и $Y_t = (1 + L)^t X_t$. Выразите X_t с помощью Y_t и оператора лага L .

2.3 Пусть F_n — последовательность чисел Фибоначчи. Рассмотрим величину

$$\frac{F_{101} + C_5^1 F_{102} + C_5^2 F_{103} + C_5^3 F_{104} + C_5^4 F_{105} + C_5^5 F_{106}}{F_{111}}$$

1. Запишите величину с помощью оператора лага
2. Упростите величину

2.4 Пусть X_t , $t = \dots - 2, -1, 0, 1, 2, \dots$ — случайный процесс. И $Y_t = X_{-t}$. Какое рассуждение верно?

1. $LY_t = LX_{-t} = X_{-t-1}$;
2. $LY_t = Y_{t-1} = X_{-t+1}$;
3. $X_t LY_t = X_t Y_{t-1}$;
4. $X_t LY_t = X_{t-1} Y_t$;

2.5 Пусть Y_t — стационарный процесс. Верно ли, что стационарны:

1. $Z_t = 2Y_t$
2. $Z_t = Y_t + 1$
3. $Z_t = \Delta Y_t$
4. $Z_t = 2Y_t + 3Y_{t-1}$

2.6 Известно, что временной ряд Y_t порожден стационарным процессом, задаваемым соотношением $Y_t = 1 + 0.5Y_{t-1} + \varepsilon_t$. Имеется 1000 наблюдений.

Вася построил регрессию Y_t на константу и Y_{t-1} . Петя построил регрессию на константу и Y_{t+1} .

Как примерно будут соотноситься между собой их оценки коэффициентов?

2.7 Правильный кубик подбрасывают три раза, обозначим результаты подбрасываний X_1 , X_2 и X_3 .

Также введем обозначения для сумм $L = X_1 + X_2$, $R = X_2 + X_3$ и $S = X_1 + X_2 + X_3$.

1. Интуитивно, без вычислений, определите знак обычных и частных корреляций $\text{Corr}(L, R)$, $\text{Corr}(L, S)$, $\text{pCorr}(L, R; S)$, $\text{pCorr}(L, S; R)$, $\text{Corr}(X_1, R)$, $\text{pCorr}(X_1, R; S)$, $\text{pCorr}(X_1, R; L)$, $\text{pCorr}(L, R; X_2)$, $\text{pCorr}(L, R; X_1)$;

2. Какие из корреляций по модулю равны единице?
3. Найдите все упомянутые обычные и частные корреляции.

2.8 Известно, что ε_t — белый шум. У каких разностных уравнений есть слабо стационарные решения?

1. $y_t = 1 + \varepsilon_t + 0.5\varepsilon_{t-1} + 0.25\varepsilon_{t-2}$
2. $y_t = -2y_{t-1} - 3y_{t-2} + \varepsilon_t + \varepsilon_{t-1}$
3. $y_t = -0.5y_{t-1} + \varepsilon_t$
4. $y_t = 1 - 1.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t - 1.5\varepsilon_{t-1} - 0.5\varepsilon_{t-2}$
5. $y_t = 1 + 0.64y_{t-2} + \varepsilon_t + 0.64\varepsilon_{t-1}$
6. $y_t = 1 + t + \varepsilon_t$
7. $y_t = 1 + y_{t-1} + \varepsilon_t$

2.9 Рассмотрим модель $y_t = \mu + \varepsilon_t$, где ε_t — стационарный AR(1) процесс $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ с $u_t \sim N(0, \sigma^2)$. Найдите условную логарифмическую функцию правдоподобия $l(\mu, \rho, \sigma^2 | y_1)$.

2.10 Известно, что ε_t — белый шум. Классифицируйте в рамках классификации ARIMA процесс $y_t = 1 + \varepsilon_t + 0.5\varepsilon_{t-1} + 0.4\varepsilon_{t-2} + 0.3\varepsilon_{t-3} + 0.2y_{t-1} + 0.1y_{t-2}$.

2.11 На графике представлены данные по уровню озера Гурон в футах в 1875-1972 годах:

```

1 level <- LakeHuron
2 df <- data.frame(level, obs = 1875:1972)
3 n <- nrow(df) # used later for answers
4 v.acf <- acf(level, plot = FALSE)$acf
5 v.pacf <- pacf(level, plot = FALSE)$acf
6 acfs.df <- data.frame(lag = c(1:15, 1:15),
7   acf = c(v.acf[2:16], v.pacf[1:15]),
8   acf.type = rep(c("ACF", "PACF"), each = 15))
9 model <- arima(level, order = c(1, 0, 1))
10 resid <- model$residuals
11 resid.acf <- acf(resid, plot = FALSE)$acf

```

```

1 tikz("../R_plots/huron_ts.tikz", standAlone = FALSE, bareBones = TRUE)
2 ggplot(df, aes(x = obs, y = level)) + geom_line() +
3   labs(x = "Год", y = "Уровень озера (футы)")
4 dev.off()

```

График автокорреляционной и частной автокорреляционной функций:

```

1 ggplot(acfs.df, aes(x = lag, y = acf, fill = acf.type))+
2   geom_histogram(position = "dodge", stat = "identity")+
3   xlab("Лag") + ylab("Корреляция") +
4   guides(fill = guide_legend(title = NULL))+
5   geom_hline(yintercept = 1.96 / sqrt(nrow(df)))+
6   geom_hline(yintercept = -1.96 / sqrt(nrow(df)))

```

1. Судя по графикам, какие модели класса ARMA или ARIMA имеет смысл оценить?
2. По результатам оценки некоей модели ARMA с двумя параметрами, исследователь посчитал оценки автокорреляционной функции для остатков модели. Известно, что для остатков модели первые три выборочные автокорреляции равны соответственно 0.00467, -0.0129 и -0.063 . С помощью подходящей статистики проверьте гипотезу о том, что первые три корреляции ошибок модели равны нулю.

2.12 Процесс x_t — это процесс y_t , наблюдаемый с ошибкой, т.е. $x_t = y_t + \nu_t$. Ошибки ν_t являются белым шумом и не коррелированы с y_t .

1. Является ли процесс x_t MA(1) процессом, если y_t — MA(1) процесс? Если да, то как связаны их автокорреляционные функции?
2. Является ли процесс x_t стационарным AR(1) процессом, если y_t — стационарный AR(1) процесс? Если да, то как связаны их автокорреляционные функции?

2.13 Пусть ε_t — белый шум. Рассмотрим процесс $y_t = 2 + 0.5y_{t-1} + \varepsilon_t$ с различными начальными условиями, указанными ниже.

1. Найдите $\mathbb{E}(y_t)$, $\text{Var}(y_t)$ и определите, является ли процесс стационарным, если:
 - (a) $y_1 = 0$
 - (b) $y_1 = 4$
 - (c) $y_1 = 4 + \varepsilon_1$
 - (d) $y_1 = 4 + \frac{2}{\sqrt{3}}\varepsilon_1$
2. Как точно следует понимать фразу «процесс $y_t = 2 + 0.5y_{t-1} + \varepsilon_t$ является стационарным»?

2.14 Верно ли, что при удалении из стационарного ряда каждого второго наблюдения получается стационарный ряд?

2.15 У эконометрессы Ефросиньи был стационарный ряд. Ей было скучно и она подбрасывала неправильную монетку, выпадающую орлом с вероятностью 0.7. Если выпадал орёл, она оставляла очередной y_t , если решка — то зачёркивала. Получается ли у Ефросиньи стационарный ряд?

2.16 Имеется временной ряд, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{101}$. Величины ε_t нормально распределены, $N(0, \sigma^2)$, и независимы. Построим график этого процесса.

1. Является ли этот процесс белым шумом?
2. Сколько в среднем раз график пересекает ось абсцисс?
3. Оцените вероятность того, что график пересечет ось абсцисс более 60 раз.

2.17 Рассмотрим стационарный AR(1) процесс $y_t = \rho y_{t-1} + \varepsilon_t$, где $\varepsilon_t \sim N(0, 1)$. Имеется ряд y_1, y_2, \dots, y_{101} . Построен график этого процесса. Как от ρ зависит математическое ожидание количества пересечений графика с осью абсцисс?

2.18 Рассмотрим процессы:

А Процесс скользящего среднего:

$$y_t = \varepsilon_t + 2\varepsilon_{t-1} + 3$$

В

$$a_t = \varepsilon_t + \varepsilon_1 + 3$$

C

$$b_t = t\varepsilon_t + 3$$

D

$$c_t = \cos\left(\frac{\pi t}{2}\right)\varepsilon_1 + \sin\left(\frac{\pi t}{2}\right)\varepsilon_2 + 2$$

E Процесс случайного блуждания со смещением:

$$\begin{cases} z_t = \varepsilon_t + z_{t-1} + 3 \\ z_0 = 0 \end{cases}$$

F Процесс с трендом:

$$w_t = 2 + 3t + \varepsilon_t$$

G Еще один процесс:

$$r_t = \begin{cases} 1, & \text{при четных } t \\ -1, & \text{при нечетных } t \end{cases}$$

H Приращение случайного блуждания

$$s_t = \Delta z_t$$

I Приращение процесса с трендом

$$d_t = \Delta w_t$$

Для каждого процесса:

1. Найдите $\mathbb{E}(y_t)$, $\text{Var}(y_t)$
2. Найдите $\gamma_k = \text{Cov}(y_t, y_{t-k})$
3. Найдите $\rho_k = \text{Corr}(y_t, y_{t-k})$. Если ни одна корреляция ρ_k не зависит от времени t , то постройте график зависимости ρ_k от k .
4. Является ли процесс стационарным?
5. Сгенерируйте одну реализацию процесса. Постройте её график и график оценки автокорреляционной функции.

2.19 Эконометресса Антуанетта построила график автоковариационной функции временного ряда и распечатала его:

здесь график

Потом она с ужасом обнаружила, что до презентации исследования остается совсем мало времени, а распечатать надо было график автокорреляционной функции. Что надо исправить Антуанетте на графике, чтобы успеть еще сделать причёску и макияж (это очень важно для презентации)?

2.20 Рассмотрите стационарные процессы

- A. AR(1): $y_t = 5 + 0.3y_{t-1} + \varepsilon_t$
- B. AR(2): $y_t = 5 + 0.3y_{t-1} + 0.1y_{t-2} + \varepsilon_t$
- C. MA(1): $y_t = 5 + 0.3\varepsilon_{t-1} + \varepsilon_t$
- D. MA(2): $y_t = 5 + 0.3\varepsilon_{t-1} + 0.9\varepsilon_{t-2} + \varepsilon_t$
- E. ARMA(1, 1): $y_t = 5 + 0.3y_{t-1} + 0.4\varepsilon_{t-1} + \varepsilon_t$

Если возможно, то представьте каждый процесс в виде:

1. $MA(\infty)$.
2. $AR(\infty)$.
3. $y_t = c + \gamma_1 y_{t-1} + u_t$, где u_t некоррелирован с y_{t-1} . Будет ли u_t белым шумом?
4. $y_t = c + \gamma_1 y_{t+1} + u_t$, где u_t некоррелирован с y_{t+1} . Будет ли u_t белым шумом?
5. $y_t = c + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + u_t$, где u_t некоррелирован с y_{t-1} и y_{t-2} . Будет ли u_t белым шумом?
6. $y_t = c + \gamma_1 y_{t+1} + \gamma_2 y_{t+2} + u_t$, где u_t некоррелирован с y_{t+1} и y_{t+2} . Будет ли u_t белым шумом?

2.21 Рассмотрите стационарные процессы

- A. $AR(1)$: $y_t = 5 + 0.3y_{t-1} + \varepsilon_t$
- B. $AR(2)$: $y_t = 5 + 0.3y_{t-1} + 0.1y_{t-2} + \varepsilon_t$
- C. $MA(1)$: $y_t = 5 + 0.3\varepsilon_{t-1} + \varepsilon_t$
- D. $MA(2)$: $y_t = 5 + 0.3\varepsilon_{t-1} + 0.9\varepsilon_{t-2} + \varepsilon_t$
- E. $ARMA(1, 1)$: $y_t = 5 + 0.3y_{t-1} + 0.4\varepsilon_{t-1} + \varepsilon_t$

Для каждого из процессов:

1. Найдите математическое ожидание $\mathbb{E}(y_t)$.
2. Найдите первые три значения автокорреляционной функции ρ_1, ρ_2, ρ_3 .
3. Найдите первые три значения частной автокорреляционной функции $\phi_{11}, \phi_{22}, \phi_{33}$.

2.22 Известна автокорреляционная функция процесса (y_t) : $\rho_1 = 0.7$, $\rho_2 = 0.3$, и $\rho_k = 0$ при $k \geq 3$. Кроме того, $\mathbb{E}(y_t) = 4$. Выпишите возможные уравнения процесса.

2.23 Известна частная автокорреляционная функция процесса (y_t) : $\phi_{11} = 0.7$, $\phi_{22} = 0.3$, и $\phi_{kk} = 0$ при $k \geq 3$. Кроме того, $\mathbb{E}(y_t) = 4$. Выпишите возможные уравнения процесса.

2.24 Если возможно, то найдите процесс с данной автокорреляционной или частной автокорреляционной функцией.

1. $ACF = (0.9, -0.9, 0, 0, 0, \dots)$;
2. $PACF = (0.9, -0.9, 0, 0, 0, \dots)$;
3. $PACF = (0.9, 0, 0, 0, 0, \dots)$;
4. $PACF = (0, 0.9, 0, 0, 0, \dots)$;
5. $ACF = (0.9, 0, 0, 0, 0, \dots)$;
6. $ACF = (0, 0.9, 0, 0, 0, \dots)$;

2.25 Рассмотрим стационарный процесс $y_t = 4 + 0.7y_{t-1} - 0.12y_{t-2} + \varepsilon_t$, где ε_t — белый шум, причём $\text{Cov}(\varepsilon_t, y_{t-k}) = 0$ при $k \geq 1$.

1. Найдите автокорреляционную функцию: ρ_1, ρ_2 и общую формулу для ρ_k .
2. Найдите $\lim_{k \rightarrow \infty} \rho_k$.
3. Найдите частную автокорреляционную функцию: $\phi_{11}, \phi_{22}, \dots$

2.26 Рассмотрим стационарный процесс с уравнением

$$y_t = 10 + 0.69y_{t-1} + \varepsilon_t - 0.71\varepsilon_{t-1}.$$

Выпишите гораздо более простой процесс со свойствами близкими к свойствам данного процесса.

2.27 Рассмотрим уравнение

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

Какие из указанных процессов являются его решением? Стационарным решением?

1. $y_t = 0.5^t$;
2. $y_t = \sum_{i=0}^{\infty} 0.5^i \varepsilon_{t-i}$;
3. $y_t = 0.5^t + \sum_{i=0}^{\infty} 0.5^i \varepsilon_{t-i}$;
4. $y_t = 0.5^t \varepsilon_{100} + \sum_{i=0}^{\infty} 0.5^i \varepsilon_{t-i}$;
5. $y_t = 0.5^t + \sum_{i=0}^t 0.5^i \varepsilon_{t-i}$;
6. $y_t = \sum_{i=0}^t 0.5^i \varepsilon_{t-i}$;

2.28 Рассмотрим стационарный процесс y_t , задаваемый уравнением

$$y_t = 2 + 0.6 \cdot y_{t-1} - 0.08y_{t-2} + \varepsilon_t,$$

где $\varepsilon_t \sim \mathcal{N}(0; 4)$.

1. Найдите $\mathbb{E}_t(y_{t+1})$, $\text{Var}_t(y_{t+1})$
2. Найдите $\mathbb{E}_t(y_{t+2})$, $\text{Var}_t(y_{t+2})$
3. Постройте 95%-ый предиктивный интервал для y_{102} , если $y_{99} = 5$, $y_{100} = 5.1$
4. Найдите $\mathbb{E}(y_t)$, $\text{Var}(y_t)$
5. Найдите $\lim_{h \rightarrow \infty} \mathbb{E}_t(y_{t+h})$, $\lim_{h \rightarrow \infty} \text{Var}_t(y_{t+h})$

2.29 Задан процесс $y_t = 7 + u_t + 0.2u_{t-1}$, где u_t независимы и нормальны $u_t \sim \mathcal{N}(0; 4)$. Известно, что $y_{100} = 7.2$, $u_{100} = 1.3$, $y_{100} + (-0.2)y_{99} + (-0.2)^2 y_{98} + \dots + (-0.2)^{99} y_1 = 5.6$.

Пусть $\mathcal{F}_t = \sigma(y_t, y_{t-1}, \dots, y_1, u_t, u_{t-1}, \dots, u_1)$ и $\mathcal{H}_t = \sigma(y_t, y_{t-1}, \dots, y_1)$.

1. Найдите $\mathbb{E}(y_{101} | \mathcal{F}_{100})$, $\text{Var}(y_{101} | \mathcal{F}_{100})$.
2. С помощью $AR(\infty)$ представления примерно найдите $\mathbb{E}(y_{101} | \mathcal{H}_{100})$, $\text{Var}(y_{101} | \mathcal{H}_{100})$. Постройте 95%-ый предиктивный интервал для y_{101} .
3. Найдите $\mathbb{E}(y_{101} | y_{100})$, $\text{Var}(y_{101} | y_{100})$.
4. Найдите $\mathbb{E}(y_{101} | y_{100}, y_{99})$, $\text{Var}(y_{101} | y_{100}, y_{99})$.

2.30 Величины X_t независимы и равновероятно принимают значения 0 и 1. Величины Y_t независимы и нормальны $\mathcal{N}(0; 24)$. Процессы (X_t) и (Y_t) независимы. Для каждого из пунктов ответьте на три вопроса. Верно ли, что величины Z_t одинаково распределены? Верно ли, что они независимы? Верно ли, что процесс (Z_t) — белый шум?

1. $Z_t = X_t(1 - X_{t-1})Y_t$;
2. $Z_t = Y_{t-1}Y_t$;

2.31 Величина Z равновероятно принимает значения 0 и 1. Условное распределение вектора $X = (X_1, X_2)$ при известном Z известно:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} | Z = 0 \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} | Z = 1 \sim \mathcal{N} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}; \begin{pmatrix} 4 & -1 \\ -1 & 9 \end{pmatrix} \right)$$

Найдите

1. Частную корреляцию $\text{pCorr}(X_1, X_2; Z)$;
2. Условную корреляцию $\text{CCorr}(X_1, X_2|Z)$;

Глава 3

GARCH

Положение GARCH-модели среди классических моделей временных рядов

$$Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{j=1}^k \beta_j X_{tj},$$
$$\varepsilon_t = \sigma_t \cdot \xi_t,$$
$$\sigma_t^2 = \omega + \sum_{i=1}^s \delta_i \sigma_{t-i}^2 + \sum_{j=1}^r \gamma_j \varepsilon_{t-j}^2.$$

- при $s = 0, r = 0, k = 0$ ARMAX/GARCH — это классическая ARMA(p, q)-модель,
- при $s = 0, r = 0$ ARMAX/GARCH — это ARMA(p, q)-модель, в которой в качестве объясняющих переменных дополнительно включены экзогенные ряды $\{X_{t1}\}, \dots, \{X_{tk}\}$.

Пример использования GARCH-модели

Пусть P_t — цена акции, фьючерса или значение некоторого индекса цен финансовых инструментов в момент времени t .

- *простой доходностью* называется $\frac{P_t - P_{t-1}}{P_{t-1}}$,
- *логарифмической доходностью* называется $\ln \frac{P_t}{P_{t-1}}$.

Связь между простой и логарифмической доходностью

$$\ln \frac{P_t}{P_{t-1}} = \ln \left(\frac{P_{t-1} + P_t - P_{t-1}}{P_{t-1}} \right) = \ln \left(1 + \frac{P_t - P_{t-1}}{P_{t-1}} \right).$$

Используя формулу Тейлора $\ln(1+x) = x + o(x)$ при $x \rightarrow 0$, можем записать следующее приближенное равенство:

$$\ln \frac{P_t}{P_{t-1}} \approx \frac{P_t - P_{t-1}}{P_{t-1}}$$

при малых значениях простой доходности $\frac{P_t - P_{t-1}}{P_{t-1}}$.

В финансовой математике, как правило, используется логарифмическая доходность. Это связано с тем, что

$$\ln \frac{P_T}{P_0} = \ln \frac{P_1}{P_0} + \ln \frac{P_2}{P_1} + \dots + \ln \frac{P_T}{P_{T-1}},$$

т. е. логарифмическая доходность за период $[0; T]$ есть сумма логарифмических доходностей за периоды $[0; 1], [1; 2], \dots, [T-1; T]$.

- В качестве зависимой переменной Y_t возьмём логарифмическую доходность $\ln \frac{P_t}{P_{t-1}}$ интересующего нас финансового инструмента.
- Простейшая модель для расчёта и прогнозирования волатильности — ARMAX($p = 0, q = 0, k = 0$)/GARCH($s = 1, r = 1$)-модель:

$$\begin{aligned} Y_t &= c + \varepsilon_t, \\ \varepsilon_t &= \sigma_t \cdot \xi_t, \\ \sigma_t^2 &= \omega + \delta \cdot \sigma_{t-1}^2 + \gamma \cdot \varepsilon_{t-1}^2, \end{aligned}$$

- Дальнейшее изложение будем вести на примере данной модели.

Определение 3.1. Пусть $\omega > 0$, $\delta \geq 0$, $\gamma \geq 0$, $\delta + \gamma < 1$ — некоторые параметры, а $\sigma_0, \xi_0, \xi_1, \xi_2, \dots$ — независимые случайные величины такие, что

$$\mathbb{E}\sigma_0^2 = \frac{\omega}{1 - \delta - \gamma}, \quad \mathbb{E}\xi_t = 0, \quad \mathbb{E}\xi_t^2 = 1, \quad t \geq 1.$$

В этом случае говорят, что последовательность случайных величин $\{\varepsilon_t\}_{t=0}^\infty$ образует *GARCH(1,1)-процесс*, если выполнены следующие соотношения:

$$\begin{aligned} \varepsilon_0 &= \sigma_0 \cdot \xi_0, \\ \varepsilon_t &= \sigma_t \cdot \xi_t, \quad \sigma_t^2 = \omega + \delta \cdot \sigma_{t-1}^2 + \gamma \cdot \varepsilon_{t-1}^2, \quad t \geq 1. \end{aligned}$$

Напомним определения слабо стационарного процесса и белого шума.

Определение 3.2. Случайный процесс $\{X_t\}_{t=0}^\infty$ называется *слабо стационарным*, если

1. $\mathbb{E}X_t^2 < \infty$ для всех $t \geq 0$;
2. $\mathbb{E}X_t = \mathbb{E}X_s$ для всех $t, s \geq 0$;
3. $D X_t = D X_s$ для всех $t, s \geq 0$;
4. $\text{cov}(X_{t+h}, X_{s+h}) = \text{cov}(X_t, X_s)$ для всех $t, s \geq 0$ и любого h такого, что $t + h \geq 0$ и $s + h \geq 0$.

Определение 3.3. Слабо стационарный процесс $\{X_t\}_{t=0}^\infty$ называется *белым шумом*, если $\mathbb{E}X_t = 0$ и $\text{cov}(X_t, X_s) = 0$ при $t, s \geq 0, t \neq s$.

Ниже мы покажем, что GARCH(1,1)-процесс $\{\varepsilon_t\}_{t=0}^\infty$ является белым шумом.

Лемма 3.1. Пусть случайные величины X_1, \dots, X_m и Y_1, \dots, Y_n независимы в совокупности. Тогда для любых (борелевских) функций $f: \mathbb{R}^m \rightarrow \mathbb{R}^1$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^1$ случайные величины $U = f(X_1, \dots, X_m)$ и $V = g(Y_1, \dots, Y_n)$ независимы.

Доказательство. См., например, Ширяев А. Н. [5], гл. II, § 6, стр. 256. ■

Лемма 3.2. Пусть независимые случайные величины X и Y имеют конечное математическое ожидание. Тогда

- (i) математическое ожидание случайной величины $X \cdot Y$ конечно;
- (ii) $\mathbb{E}[X \cdot Y] = \mathbb{E}X \cdot \mathbb{E}Y$.

Доказательство. См. Ширяев А. Н. [5], гл. II, § 6, стр. 267, теорема 6. ■

Лемма 3.3. Пусть случайные величины X^2 и Y^2 имеют конечное математическое ожидание. Тогда случайная величина $X \cdot Y$ также имеет конечное математическое ожидание.

Доказательство. В силу свойства математического ожидания $|\mathbb{E}Z| \leq \mathbb{E}|Z|$ и неравенства $|X \cdot Y| \leq \frac{1}{2} \cdot X^2 + \frac{1}{2} \cdot Y^2$ получаем:

$$|\mathbb{E}[X \cdot Y]| \leq \mathbb{E}|X \cdot Y| \leq \frac{1}{2} \cdot \mathbb{E}X^2 + \frac{1}{2} \cdot \mathbb{E}Y^2 < \infty.$$

■

Лемма 3.4. Для любого $t \geq 0$ случайные величины σ_t и ξ_t независимы.

Доказательство. При $t = 0$ независимость случайных величин σ_0 и ξ_0 содержится непосредственно в определении GARCH(1,1)-процесса.

При $t = 1$ независимость σ_1 и ξ_1 следует из того, что случайные величины σ_0, ξ_0, ξ_1 независимы в совокупности, и того, что $\sigma_1 = \sqrt{\omega + \delta \cdot \sigma_0^2 + \gamma \cdot \sigma_0^2 \cdot \xi_0^2}$, т. е. σ_1 является функцией от σ_0, ξ_0 .

Независимость σ_t и ξ_t при $t \geq 2$ обосновывается аналогично тому, как это сделано при $t = 1$. Действительно, σ_t есть функция от $\sigma_0, \xi_0, \xi_1, \dots, \xi_{t-1}$, при этом величины $\sigma_0, \xi_0, \xi_1, \dots, \xi_t$ независимы в совокупности. ■

Утверждение 3.1. Пусть последовательность случайных величин $\{\varepsilon_t\}_{t=0}^\infty$ образует GARCH(1,1)-процесс. Тогда для любого $t \geq 0$

(i) $\mathbb{E}\varepsilon_t^2 < \infty$;

(ii) $\mathbb{E}\varepsilon_t = 0$;

(iii) $\mathbb{E}\varepsilon_t^2 = \frac{\omega}{1-\delta-\gamma}$;

(iv) $\text{cov}(\varepsilon_t, \varepsilon_s) = 0$ при $t \neq s, s \geq 0$.

Доказательство. (i) ($t = 0$) По условию случайные величины σ_0^2 и ξ_0^2 имеют конечное математическое ожидание. При этом независимость σ_0^2 и ξ_0^2 вытекает из независимости σ_0 и ξ_0 . Следовательно, в силу леммы 2 случайная величина $\varepsilon_0^2 = \sigma_0^2 \cdot \xi_0^2$ имеет конечное математическое ожидание.

($t = 1$) Согласно лемме 4, случайные величины σ_1 и ξ_1 независимы. Значит, σ_1^2 и ξ_1^2 также независимы. Кроме того, по условию, математическое ожидание ξ_1^2 конечно, а конечность $\mathbb{E}\sigma_1^2$ вытекает из конечности $\mathbb{E}\sigma_0^2, \mathbb{E}\varepsilon_0^2$ и формулы $\sigma_1^2 = \omega + \delta \cdot \sigma_0^2 + \gamma \cdot \varepsilon_0^2$. Следовательно, $\varepsilon_1^2 = \sigma_1^2 \cdot \xi_1^2$ имеет конечное математическое ожидание.

($t \geq 2$) Доказательство конечности $\mathbb{E}\varepsilon_t^2$ при $t \geq 2$ проводится аналогично случаю $t = 1$.

(ii) Для $t \geq 0$ имеем

$$\mathbb{E}\varepsilon_t = \mathbb{E}[\sigma_t \cdot \xi_t] = \mathbb{E}\sigma_t \cdot \mathbb{E}\xi_t = 0.$$

Здесь мы воспользовались независимостью случайных величин σ_t и ξ_t , а также $\mathbb{E}\xi_t = 0$.

(iii) ($t = 0$) При $t = 0$ имеем

$$\mathbb{E}\varepsilon_0^2 = \mathbb{E}\sigma_0^2 \cdot \mathbb{E}\xi_0^2 = \frac{\omega}{1-\delta-\gamma} \cdot 1 = \frac{\omega}{1-\delta-\gamma}.$$

($t = 1$) Пусть $t = 1$. По лемме 4 и доказанному выше, получаем

$$\begin{aligned} \mathbb{E}\varepsilon_1^2 &= \mathbb{E}\sigma_1^2 \cdot \mathbb{E}\xi_1^2 = \mathbb{E}\sigma_1^2 = \omega + \delta \cdot \mathbb{E}\sigma_0^2 + \gamma \cdot \mathbb{E}\varepsilon_0^2 = \\ &= \omega + \delta \cdot \frac{\omega}{1-\delta-\gamma} + \gamma \cdot \frac{\omega}{1-\delta-\gamma} = \frac{\omega}{1-\delta-\gamma}. \end{aligned}$$

($t \geq 2$) Доказательство утверждения при $t \geq 2$ выполняется аналогично рассмотренному случаю $t = 1$.

(iv) Пусть $0 \leq s < t$. Математическое ожидание ξ_t конечно по определению GARCH(1,1)-процесса. Конечность математического ожидания случайной величины $\sigma_t \cdot \varepsilon_s$ следует из конечности $\mathbb{E}\sigma_t^2$ и $\mathbb{E}\varepsilon_s^2$, а также леммы 3.3. Кроме этого, при $0 \leq s < t$ случайные величины ξ_t и $\sigma_t \cdot \varepsilon_s$ независимы. Поэтому

$$\text{cov}(\varepsilon_t, \varepsilon_s) = \mathbb{E}[\varepsilon_t \cdot \varepsilon_s] = \mathbb{E}[\xi_t \cdot (\sigma_t \cdot \varepsilon_s)] = \mathbb{E}\xi_t \cdot \mathbb{E}[\sigma_t \cdot \varepsilon_s] = 0.$$

■

Замечание 3.1. В ходе доказательства пункта (i) утверждения 3.1 попутно было установлено, что $\mathbb{E}\sigma_t^2 < \infty$ для всех $t \geq 0$.

3.1 Рассмотрим следующий AR(1)-ARCH(1) процесс:

$$Y_t = 1 + 0.5Y_{t-1} + \varepsilon_t, \varepsilon_t = \nu_t \cdot \sigma_t$$

ν_t независимые $N(0; 1)$ величины.

$$\sigma_t^2 = 1 + 0.8\varepsilon_{t-1}^2$$

Также известно, что $Y_{100} = 2, Y_{99} = 1.7$

1. Найдите $E_{100}(\varepsilon_{101}^2), E_{100}(\varepsilon_{102}^2), E_{100}(\varepsilon_{103}^2), E(\varepsilon_t^2)$
2. $\text{Var}(Y_t), \text{Var}(Y_t | \mathcal{F}_{t-1})$
3. Постройте доверительный интервал для Y_{101} :
 - (a) проигнорировав условную гетероскедастичность
 - (b) учтя условную гетероскедастичность

3.2 Рассмотрим GARCH(1,2) процесс $\varepsilon_t = \sigma_t \nu_t, \sigma_t^2 = 0.2 + 0.5\sigma_{t-1}^2 + 0.2\varepsilon_{t-1}^2 + 0.1\varepsilon_{t-2}^2$. Найдите безусловную дисперсию $\text{Var}(y_t)$

3.3 Для GARCH(1,1) процесса $\varepsilon_t = \sigma_t \nu_t, \sigma_t^2 = w + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$ найдите $\mathbb{E}(\mathbb{E}(\varepsilon_t^2 | \mathcal{F}_{t-1}))$

3.4 Рассмотрим GARCH(1,1) процесс $\varepsilon_t = \sigma_t \nu_t, \sigma_t^2 = 0.1 + 0.7\sigma_{t-1}^2 + 0.2\varepsilon_{t-1}^2$. Известно, $\sigma_T = 1, \varepsilon_T = 1$. Найдите $\mathbb{E}(\sigma_{T+2}^2 | \mathcal{F}_T)$.

3.5 Найдите безусловную дисперсию GARCH-процессов

1. $\varepsilon_t = \sigma_t \cdot z_t, \sigma_t^2 = 0.1 + 0.8\sigma_{t-1}^2 + 0.1\varepsilon_{t-1}^2$
2. $\varepsilon_t = \sigma_t \cdot z_t, \sigma_t^2 = 0.4 + 0.7\sigma_{t-1}^2 + 0.1\varepsilon_{t-1}^2$
3. $\varepsilon_t = \sigma_t \cdot z_t, \sigma_t^2 = 0.2 + 0.8\sigma_{t-1}^2 + 0.1\varepsilon_{t-1}^2$

3.6 Являются ли верными следующие утверждения?

1. GARCH-процесс является процессом белого шума, условная дисперсия которого изменяется во времени
2. Модель GARCH(1,1) предназначена для прогнозирования меры изменчивости цены финансового инструмента, а не для прогнозирования самой цены инструмента
3. При помощи GARCH-процесса можно устранять гетероскедастичность
4. Безусловная дисперсия GARCH-процесса изменяется во времени
5. Модель GARCH(1,1) может быть использована для прогнозирования волатильности финансовых инструментов на несколько торговых недель вперёд

3.7 Рассмотрим GARCH-процесс $\varepsilon_t = \sigma_t \cdot z_t, \sigma_t^2 = k + g_1\sigma_{t-1}^2 + a_1\varepsilon_{t-1}^2$. Найдите

1. $\mathbb{E}(z_t), \mathbb{E}(z_t^2), \mathbb{E}(\varepsilon_t), \mathbb{E}(\varepsilon_t^2)$

2. $\text{Var}(z_t), \text{Var}(\varepsilon_t), \text{Var}(\varepsilon_t | \mathcal{F}_{t-1})$
3. $\mathbb{E}(\varepsilon_t | \mathcal{F}_{t-1}), \mathbb{E}(\varepsilon_t^2 | \mathcal{F}_{t-1}), \mathbb{E}(\sigma_t^2 | \mathcal{F}_{t-1})$
4. $\mathbb{E}(z_t z_{t-1}), \mathbb{E}(z_t^2 z_{t-1}^2), \text{Cov}(\varepsilon_t, \varepsilon_{t-1}), \text{Cov}(\varepsilon_t^2, \varepsilon_{t-1}^2)$
5. $\lim_{h \rightarrow \infty} \mathbb{E}(\sigma_{t+h}^2 | \mathcal{F}_t)$

3.8 Используя 500 наблюдений дневных логарифмических доходностей y_t , была оценена GARCH(1,1)-модель: $\hat{y}_t = -0.000708 + \hat{\varepsilon}_t$, $\varepsilon_t = \sigma_t \cdot z_t$, $\sigma_t^2 = 0.000455 + 0.6424\sigma_{t-1}^2 + 0.2509\varepsilon_{t-1}^2$. Также известно, что $\hat{\sigma}_{499}^2 = 0.002568$, $\hat{\varepsilon}_{499}^2 = 0.000014$, $\hat{\varepsilon}_{500}^2 = 0.002178$. Найдите

1. $\hat{\sigma}_{500}^2, \hat{\sigma}_{501}^2, \hat{\sigma}_{502}^2$
2. Волатильность в годовом выражении в процентах, соответствующую наблюдению с номером $t = 500$

3.9 Рассмотрим ARCH(1) процесс

$$\begin{cases} y_t = 2 + \varepsilon_t \\ \varepsilon_t = \sigma_t \cdot \nu_t \\ \sigma_t^2 = 10 + 0.5\varepsilon_t^2 \end{cases}$$

1. Найдите $\text{Var}(y_{101})$, постройте 95%-ый предиктивный интервал для y_{101}
2. Известно, что $y_{100} = 3$, постройте 95%-ый предиктивный интервал для y_{101}
3. Известно, что $y_{100} = 12$, постройте 95%-ый предиктивный интервал для y_{101}

3.10 Может ли у GARCH процесса условная дисперсия ε_t быть больше, чем безусловная? А меньше, чем безусловная?

3.11 Как известно, у GARCH процесса условная дисперсия ε_t может быть как больше, так и меньше безусловной.

1. Имеет ли смысл строить предиктивный интервал для y_t , используя условную дисперсию, если она больше безусловной?
2. При построении предиктивного интервала эконометресса Агнессы использует безусловную дисперсию, если она меньше условной, и условную дисперсию, если она меньше безусловной. Корректно ли поступает Агнесса?

3.12 Рассмотрим процесс AR(1)-GARCH(1,1):

$$\begin{cases} y_t = 2 + 0.6y_{t-1} + \varepsilon_t \\ \varepsilon_t = \sigma_t \cdot \nu_t \\ \sigma_t^2 = 6 + 0.4\sigma_{t-1}^2 + 0.2\varepsilon_t^2 \end{cases}$$

Найдите $\text{Var}(\varepsilon_t | \mathcal{F}_{t-1}), \text{Var}(y_t | \mathcal{F}_{t-1}), \text{Var}(\varepsilon_t), \text{Var}(y_t)$

Глава 4

Единичный корень

4.1 Винни-Пух пытается выявить закономерность в количестве придумываемых им каждый день ворчалок. Винни-Пух решил разобраться, является ли оно стационарным процессом, для этого он оценил регрессию

$$\Delta \hat{y}_t = \underset{(0.5)}{4.5} - \underset{(0.1)}{0.4} y_{t-1} + \underset{(0.5)}{0.7} \Delta y_{t-1}$$

Из-за опилок в голове Винни-Пух забыл, какой тест ему нужно провести, то ли Доктора Ватсона, то ли Дикого Фуллера.

1. Аккуратно сформулируйте основную и альтернативную гипотезы
2. Проведите подходящий тест на уровне значимости 5%
3. Сделайте вывод о стационарности ряда
4. Почему Сова не советовала Винни-Пуху пользоваться широко применяемым в Лесу t -распределением?

Глава 5

Векторная авторегрессия

5.1 Рассмотрим систему уравнений:

$$\begin{cases} x_t = -\frac{1}{6}x_{t-1} + \frac{2}{6}y_{t-1} + \varepsilon_{xt} \\ y_t = -\frac{4}{6}x_{t-1} + \frac{1}{6}y_{t-1} + \varepsilon_{yt} \end{cases}$$

1. Есть ли у данной системы стационарное решение?
2. Если стационарное решение имеется, то найдите $\mathbb{E}(x_t)$ и $\mathbb{E}(y_t)$
3. Нарисуйте в осях (x_t, y_t) типичную траекторию стационарного решения

5.2 Рассмотрим систему уравнений:

$$\begin{cases} x_t = -0.2x_{t-1} + 0.6y_{t-1} + \varepsilon_{xt} \\ y_t = 1.2x_{t-1} + 0.4y_{t-1} + \varepsilon_{yt} \end{cases}$$

1. Есть ли у данной системы коинтегрированное решение?
2. Если коинтегрированное решение имеется, то найдите коинтеграционное соотношение и представьте модель в виде модели коррекции ошибок
3. Нарисуйте в осях (x_t, y_t) типичную траекторию коинтегрированного решения

5.3 Белые шумы ε_t и u_t независимы. Пусть $y_t = 2 - 0.5t + u_t$, $x_t = 1 + 0.5t + \varepsilon_t$.

1. Является ли процесс $z_t = x_t + y_t$ стационарным?
2. Являются ли процессы x_t и y_t коинтегрированными?

5.4 Два процесса (x_t) и (y_t) называются независимыми, если независимы любые случайные величины x_s и y_t .

Докажите каждое утверждение или приведите контр-пример.

1. Сумма двух белых шумов является белым шумом.
2. Сумма двух независимых белых шумов является белым шумом.
3. Сумма двух стационарных процессов стационарна.
4. Сумма двух независимых стационарных процессов стационарна.
5. Сумма двух нестационарных процессов нестационарна.
6. Сумма двух независимых нестационарных процессов нестационарна.

5.5 Какие процессы могут быть коинтегрированы: $x_t \sim I(0)$, $y_t \sim I(1)$, $z_t \sim I(2)$, $w_t \sim I(2)$, $s_t \sim I(1)$?

5.6 Белые шумы (ε_t) и (u_t) независимы.

Классифицируйте каждый процесс¹ как $ARIMA(p, d, q)$, определите порядок интеграции каждого процесса и определите, какие пары процессов коинтегрированы:

1. $a_t = 0.5a_{t-1} + u_t$
2. $b_t = b_{t-1} + u_t, b_0 = 0$
3. $c_t = 0.5b_t + \varepsilon_t$
4. $d_t = 0.3b_t + a_t$
5. $e_t = e_{t-1} + \varepsilon_t$
6. $g_t = g_{t-1} + b_t$
7. $h_t = 0.7h_{t-1} + b_t$

¹Если у уравнения не заданы начальные условия, то подразумевается стационарное решение, если оно, конечно, есть.

Глава 6

Модели состояние-наблюдение

6.1 Представьте процесс AR(1), $y_t = 0.9y_{t-1} - 0.2y_{t-2} + \varepsilon_t$, $\varepsilon \sim \text{WN}(0;1)$ в виде модели состояние-наблюдение.

1. Выбрав в качестве состояний вектор $\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}$
2. Выбрав в качестве состояний вектор $\begin{pmatrix} y_t \\ \hat{y}_{t,1} \end{pmatrix}$

Найдите дисперсии ошибок состояний

6.2 Представьте процесс MA(1), $y_t = \varepsilon_t + 0.5\varepsilon_{t-1}$, $\varepsilon \sim \text{WN}(0;1)$ в виде модели состояние-наблюдение.

1. $\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}$
2. $\begin{pmatrix} \varepsilon_t + 0.5\varepsilon_{t-1} \\ 0.5\varepsilon_t \end{pmatrix}$

6.3 Представьте процесс ARMA(1,1), $y_t = 0.5y_{t-1} + \varepsilon_t + \varepsilon_{t-1}$, $\varepsilon \sim \text{WN}(0;1)$ в виде модели состояние-наблюдение.

Вектор состояний имеет вид x_t, x_{t-1} , где $x_t = \frac{1}{1-0.5L}\varepsilon_t$

6.4 Рекурсивные коэффициенты

1. Оцените модель вида $y_t = a + b_t x_t + \varepsilon_t$, где $b_t = b_{t-1}$.
2. Сравните графики filtered state и smoothed state.
3. Сравните финальное состояние b_T с коэффициентом в обычной модели линейной регрессии, $y_t = a + bx_t + \varepsilon_t$.

Глава 7

Решения и ответы к избранным задачам

1.1. В данном случае статистика DW не применима, так как есть лаг y_{t-1} среди регрессоров.

1.2.

1. $\mathbb{E}(\varepsilon_t) = 0$, $\text{Var}(\varepsilon_1) = \sigma^2$, $\text{Var}(\varepsilon_t) = 2\sigma^2$ при $t \geq 2$. Гетероскедастичная.
2. $\text{Cov}(e_t, e_{t+1}) = \sigma^2$. Автокоррелированная.
3. $\hat{\beta}$ — несмещенная, неэффективная
4. Более эффективной будет $\hat{\beta}_{gls} = (X'V^{-1}X)^{-1}X'V^{-1}y$, где

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Матрица V известна с точностью до константы σ^2 , но в формуле для $\hat{\beta}_{gls}$ неизвестная σ^2 сократится. Другой способ построить эффективную оценку — применить МНК к преобразованным наблюдениям, т.е. $\hat{\beta}_{gls} = \frac{\sum x'_i y'_i}{\sum x'^2_i}$, где $y'_1 = y_1$, $x'_1 = x_1$, $y'_t = y_t - y_{t-1}$, $x'_t = x_t - x_{t-1}$ при $t \geq 2$.

1.3.

Для простоты закроем глаза на малое количество наблюдений и как индейцы пираха будем считать, что пять — это много.

1.4. 1. Поскольку имеют место соотношения $\varepsilon_1 = \rho\varepsilon_0 + u_1$ и $Y_1 = \mu + \varepsilon_1$, то из условия задачи получаем, что $\varepsilon_1 \sim N(0, \sigma^2/(1 - \rho^2))$ и $Y_1 \sim N(\mu, \sigma^2/(1 - \rho^2))$. Поэтому

$$f_{Y_1}(y_1) = \frac{1}{\sqrt{2\pi\sigma^2/(1 - \rho^2)}} \exp\left(-\frac{(y_1 - \mu)^2}{2\sigma^2/(1 - \rho^2)}\right).$$

Далее, найдем $f_{Y_2|Y_1}(y_2|y_1)$. Учитывая, что $Y_2 = \rho Y_1 + (1 - \rho)\mu + u_2$, получаем $Y_2|\{Y_1 = y_1\} \sim N(\rho y_1 + (1 - \rho)\mu, \sigma^2)$. Значит,

$$f_{Y_2|Y_1}(y_2|y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_2 - \rho y_1 - (1 - \rho)\mu)^2}{2\sigma^2}\right).$$

Действуя аналогично, получаем, что для всех $t \geq 2$ справедлива формула

$$f_{Y_t|Y_{t-1}}(y_t|y_{t-1}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - \rho y_{t-1} - (1-\rho)\mu)^2}{2\sigma^2}\right).$$

Таким образом, находим функцию правдоподобия

$$L(\mu, \rho, \sigma^2) = f_{Y_T, \dots, Y_1}(y_T, \dots, y_1) = f_{Y_1}(y_1) \prod_{t=2}^T f_{Y_t|Y_{t-1}}(y_t|y_{t-1}),$$

где $f_{Y_1}(y_1)$ и $f_{Y_t|Y_{t-1}}(y_t|y_{t-1})$ получены выше.

2. Для нахождения неизвестных параметров модели запишем логарифмическую условную функцию правдоподобия:

$$\begin{aligned} l(\mu, \rho, \sigma^2 | Y_1 = y_1) &= \sum_{t=2}^T \log f_{Y_t|Y_{t-1}}(y_t|y_{t-1}) = \\ &= -\frac{T-1}{2} \log(2\pi) - \frac{T-1}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=2}^T (y_t - \rho y_{t-1} - (1-\rho)\mu)^2. \end{aligned}$$

Найдем производные функции $l(\mu, \rho, \sigma^2 | Y_1 = y_1)$ по неизвестным параметрам:

$$\begin{aligned} \frac{\partial l}{\partial \mu} &= -\frac{1}{2\sigma^2} \sum_{t=2}^T 2(y_t - \rho y_{t-1} - (1-\rho)\mu) \cdot (\rho - 1), \\ \frac{\partial l}{\partial \rho} &= -\frac{1}{2\sigma^2} \sum_{t=2}^T 2(y_t - \rho y_{t-1} - (1-\rho)\mu) \cdot (\mu - y_{t-1}), \\ \frac{\partial l}{\partial \sigma^2} &= -\frac{T-1}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=2}^T (y_t - \rho y_{t-1} - (1-\rho)\mu)^2. \end{aligned}$$

Оценки неизвестных параметров модели могут быть получены как решение следующей системы уравнений:

$$\begin{cases} \frac{\partial l}{\partial \mu} = 0, \\ \frac{\partial l}{\partial \rho} = 0, \\ \frac{\partial l}{\partial \sigma^2} = 0. \end{cases}$$

Из первого уравнения системы получаем, что

$$\sum_{t=2}^T y_t - \hat{\rho} \sum_{t=2}^T y_{t-1} = (T-1)(1-\hat{\rho})\hat{\mu},$$

откуда

$$\hat{\mu} = \frac{\sum_{t=2}^T y_t - \hat{\rho} \sum_{t=2}^T y_{t-1}}{(T-1)(1-\hat{\rho})} = \frac{3 - \hat{\rho} \cdot 3}{4 \cdot (1-\hat{\rho})} = \frac{3}{4}.$$

Далее, если второе уравнение системы переписать в виде

$$\sum_{t=2}^T (y_t - \hat{\mu} - \hat{\rho}(y_{t-1} - \hat{\mu}))(y_{t-1} - \hat{\mu}) = 0,$$

то легко видеть, что

$$\hat{\rho} = \frac{\sum_{t=2}^T (y_t - \hat{\mu})(y_{t-1} - \hat{\mu})}{\sum_{t=2}^T (y_{t-1} - \hat{\mu})^2}.$$

Следовательно, $\hat{\rho} = -1/11 = -0.0909$.

Наконец, из третьего уравнения системы

$$\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=2}^T (y_t - \hat{\rho}y_{t-1} - (1-\hat{\rho})\hat{\mu})^2.$$

Значит, $\hat{\sigma}^2 = 165/242 = 0.6818$. Ответы: $\hat{\mu} = 3/4 = 0.75$, $\hat{\rho} = -1/11 = -0.0909$, $\hat{\sigma}^2 = 165/242 = 0.6818$.

1.5. Несмещёнными остаются. Состоятельными не всегда остаются, например, состоятельность исчезает, если все случайные ошибки тождественно равны между собой.

1.6.

1.7.

1.8.

1.9.

1. $\mathbb{E}(\varepsilon_t) = 0$, $\mathbb{V}\text{ar}(\varepsilon_t) = \sigma^2/(1 - \rho^2)$
2. $\mathbb{C}\text{ov}(\varepsilon_t, \varepsilon_{t+h}) = \rho^h \cdot \sigma^2/(1 - \rho^2)$
3. $\mathbb{C}\text{orr}(\varepsilon_t, \varepsilon_{t+h}) = \rho^h$

1.10.

1.11.

1.12.

1.13.

1.14.

1.15.

2.1.

$$(1 - 0.4L)y_t = 4 + (1 + 0.3L)\varepsilon_t$$

2.2. $X_t = (1 - L)^t Y_t$

2.3. $F_n = L(1 + L)F_n$, значит $F_n = L^k(1 + L)^k F_n$ или $F_{n+k} = (1 + L)^k F_n$

Ответ: 1

2.4. а — неверно, б — верно, в — верно, г — нет.

2.5. а, б, в, г — стационарны

2.6. Они будут примерно одинаковы. Оценка наклона определяется автоковариационной функцией.

2.7.

2.8.

1. $y_t = 1 + \varepsilon_t + 0.5\varepsilon_{t-1} + 0.25\varepsilon_{t-2}$ — стационарный
2. $y_t = -2y_{t-1} - 3y_{t-2} + \varepsilon_t + \varepsilon_{t-1}$
3. $y_t = -0.5y_{t-1} + \varepsilon_t$ — стационарный
4. $y_t = 1 - 1.5y_{t-1} - 0.5y_{t-2} + \varepsilon_t - 1.5\varepsilon_{t-1} - 0.5\varepsilon_{t-2}$
5. $y_t = 1 + 0.64y_{t-2} + \varepsilon_t + 0.64\varepsilon_{t-1}$ — стационарный
6. $y_t = 1 + t + \varepsilon_t$ — нестационарный
7. $y_t = 1 + y_{t-1} + \varepsilon_t$ — нестационарный

2.9.

2.10.

ARMA(2,3), ARIMA(2,0,3)

2.11.

1. Процесс $AR(2)$, т.к. две первые частные корреляции значимо отличаются от нуля, а гипотезы о том, что каждая последующая равна нулю не отвергаются.
2. Можно использовать одну из двух статистик

$$\text{Ljung-Box} = n(n+2) \sum_{k=1}^3 \frac{\hat{\rho}_k^2}{n-k} = 0.42886$$

$$\text{Box-Pierce} = n \sum_{k=1}^3 \hat{\rho}_k^2 = 0.4076$$

Критическое значение хи-квадрат распределения с 3-мя степенями свободы для $\alpha = 0.05$ равно $\chi_{3,crit}^2 = 7.81$. Вывод: гипотеза H_0 об отсутствии корреляции ошибок модели не отвергается.

2.12.

2.13. Процесс стационарен только при $y_1 = 4 + \frac{2}{\sqrt{3}}\varepsilon_1$. Фразу нужно понимать как «у стохастического разностного уравнения $y_t = 2 + 0.5y_{t-1} + \varepsilon_t$ есть стационарное решение».

2.14. да, стационарный

2.15. да, получается

2.16. да, это белый шум. Величина N распределена биномиально, $Bin(n = 100, p = 1/2)$, $\mathbb{E}(N) = 50$.

2.17. Среднее количество пересечений равно 50 помножить на вероятность того, что два соседних y_t разного знака. Найдём вдвое меньшую вероятность, $\mathbb{P}(y_1 > 0, y_2 < 0)$.

2.18.

$$\mathbb{E}(b_t) = 3$$

$$\mathbb{V}\text{ar}(b_t) = t^2 \sigma_\varepsilon^2$$

$$\mathbb{C}\text{ov}(b_t, b_{t-k}) = 0, k \geq 1$$

$$\mathbb{C}\text{orr}(b_t, b_{t-k}) = 0, k \geq 1$$

b_t — нестационарный из-за дисперсии

$$\mathbb{E}(c_t) = 2$$

$$\mathbb{V}\text{ar}(c_t) = \sigma_\varepsilon^2$$

$$\mathbb{C}\text{ov}(c_t, c_{t-k}) = \cos(\pi k/2) \sigma_\varepsilon^2, k \geq 1$$

$$\mathbb{C}\text{orr}(c_t, c_{t-k}) = \cos(\pi k/2), k \geq 1$$

c_t — стационарный

2.19. зачеркнуть одну цифру

2.20.

2.21.

2.22.

2.23.

2.24.

1. $ACF = (0.9, -0.9, 0, 0, 0, \dots)$ не бывает, так как определитель корреляционной матрицы 3 на 3 отрицательный;
2. $PACF = (0.9, -0.9, 0, 0, 0, \dots) - AR(2)$;
3. $PACF = (0.9, 0, 0, 0, 0, \dots) - y_t = 0.9y_{t-1} + u_t$;
4. $PACF = (0, 0.9, 0, 0, 0, 0, \dots) - y_t = 0.9y_{t-2} + u_t$;
5. $ACF = (0.9, 0, 0, 0, 0, \dots)$ — не бывает, подозрение падает на $MA(1)$, но решения только с комплексными коэффициентами, геометрически: два угла с косинусом 0.9, то есть примерно по 30 градусов, и они даже в сумме не могут дать перпендикуляр;
6. $ACF = (0, 0.9, 0, 0, 0, 0, \dots)$ — не бывает, если проредить процесс через один, то должна получиться невозможная ACF ;

В целом $PACF$ может быть любая, <http://projecteuclid.org/euclid.aos/1176342881>.

2.25. $\phi_{kk} = 0$ при $k \geq 3$.

2.26. Заметим, что $0.69 \approx 0.71$, сокращаем множитель $1 - 0.7L$, получаем $y_t = 100/3 + \varepsilon_t$.

2.27. Стационарным решением является $y_t = \sum_{i=0}^{\infty} 0.5^i \varepsilon_{t-i}$. Решениями также являются: $y_t = 0.5^t + \sum_{i=0}^{\infty} 0.5^i \varepsilon_{t-i}$, $y_t = 0.5^t \varepsilon_{100} + \sum_{i=0}^{\infty} 0.5^i \varepsilon_{t-i}$, $y_t = 0.5^t + \sum_{i=0}^t 0.5^i \varepsilon_{t-i}$, $y_t = \sum_{i=0}^t 0.5^i \varepsilon_{t-i}$.

2.28.

$$\mathbb{E}_t(y_{t+1}) = 2 + 0.6y_{t-1} - 0.08y_{t-2}, \text{Var}_t(y_{t+1}) = 4$$

$$\mathbb{E}_t(y_{t+2}) = 3.2 + 0.28y_t - 0.048y_{t-1}, \text{Var}_t(y_{t+2}) = 1.36 \cdot 4$$

$$\mathbb{E}_{100}(y_{102}) = 4.388, \text{Var}_{100}(y_{102}) = 5.44.$$

$$\text{Предиктивный интервал } [4.388 - 1.96\sqrt{5.44}; 4.388 + 1.96\sqrt{5.44}]$$

$$\mathbb{E}(y_t) = \frac{2}{0.48} \approx 4.17$$

2.29. Заметим, что $\text{Var}(u_t | \mathcal{F}_t) = 0$. Более того, для обратимого процесса $\text{Var}(u_t | y_t, y_{t-1}, \dots, y_1) \approx \text{Var}(u_t | y_t, y_{t-1}, \dots) = 0$.

$$\mathbb{E}(y_{101} | y_{100}) = 7 + 0 + 0.2\mathbb{E}(u_{100} | y_{100})$$

$$\mathbb{E}(u_{100} | y_{100}) = \beta_1 + \beta_2 y_{100}$$

$$\beta_2 = \frac{\text{Cov}(y_{100}, u_{100})}{\text{Var}(y_{100})} = 4/4.16, \beta_1 = \mathbb{E}(u_{100}) - \beta_2 \mathbb{E}(y_{100}) = -4 \cdot 7/4.16$$

$$\frac{y_t}{1 + 0.2L} = \frac{7}{1 + 0.2L} + u_t$$

Заметим, что $\frac{7}{1+0.2L} = 7/1.2$, так как $L \cdot 7 = 7$ (вчера семь равнялось семи).

По условию $\frac{y_{100}}{1+0.2L} \approx 5.6$. Знак «примерно равно» возникает из-за замены бесконечной суммы на конечную.

2.30.

1. $Z_t = X_t(1 - X_{t-1})Y_t$; Процесс Z_t — белый шум. Величины Z_t зависимы. Например, если $Z_t \neq 0$, то $Z_{t+1} = Z_{t-1} = 0$. Величины Z_t одинаково распределены.
2. $Z_t = Y_{t-1}Y_t$; Процесс Z_t — белый шум. Величины Z_t зависимы. Величины Z_t одинаково распределены.

2.31. Проекции: $\tilde{X}_1 = X_1 + Z$; $\tilde{X}_2 = X_2 + Z$; $\mathbb{E}(X_i|Z) = 1 - Z$; $\text{Cov}(X_i, Z) = -1/4$;
 Величина Z имеет распределение Бернулли, поэтому $\mathbb{E}(Z) = 1/2$ и $\text{Var}(Z) = 1/4$;

$$\text{pCorr}(X_1, X_2; Z) = \frac{-1/2}{12.5} = -\frac{1}{\sqrt{50}}$$

$$\text{Corr}(X_1, X_2|Z) = -Z/6$$

3.1.

3.2.

3.3.

3.4.

3.5. 1, 2, 2

3.6.

3.7.

3.8.

3.9.

3.10. Да, может быть и больше, и меньше.

3.11.

3.12.

$$\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \text{Var}(y_t|\mathcal{F}_{t-1}) = 6 + 0.4\sigma_{t-1}^2 + 0.2\varepsilon_t^2$$

$$\text{Var}(\varepsilon_t) = 6/(1 - 0.4 - 0.2) = 6/0.4 = 15$$

$$\text{Var}(y_t) = 15/(1 - 0.36)$$

4.1.

1. H_0 : ряд содержит единичный корень, $\beta = 0$; H_a : ряд не содержит единичного корня, $\beta < 0$
2. $ADF = -0.4/0.1 = -4$, $ADF_{crit} = -2.89$, H_0 отвергается
3. Ряд стационарен
4. При верной H_0 ряд не стационарен, и t -статистика имеет не t -распределение, а распределение Дики-Фуллера.

5.1.

5.2.

5.3.

z_t стационарный, x_t и y_t не коинтегрированы

5.4.

5.5. y_t и s_t ; z_t и w_t .

5.6.

1. $a_t = 0.5a_{t-1} + u_t$, AR(1)

2. $b_t = b_{t-1} + u_t$, $b_0 = 0$, ARIMA(0, 1, 0)

3. $c_t = 0.5b_t + \varepsilon_t$, ARIMA(0, 1, 1)

4. $d_t = d_{t-1} + a_t$, ARIMA(1, 1, 0)

5. $e_t = e_{t-1} + \varepsilon_t$, ARIMA(0, 1, 0)

6. $g_t = g_{t-1} + b_t$, ARIMA(0, 2, 0)

7. $h_t = 0.7h_{t-1} + b_t$, ARIMA(1, 1, 0)

коинтегрированы: b_t , c_t , d_t , h_t .

6.1.

6.2.

6.3.

6.4.

Литература

- [1] Greene W. H. Econometric Analysis. Prentice Hall, 2012.
- [2] Francq C., Zakoian J.-M. GARCH models: structure, statistical inference, and financial applications. Wiley, 2010.
- [3] Tsay R. S. Analysis of Financial Time Series. Wiley, 2005.
- [4] Ширяев А. Н. Основы стохастической финансовой математики. Т. 1. М.: ФАЗИС, 2004.
- [5] Ширяев А. Н. Вероятность. Т. 1. М.: МЦНМО, 2007.

Предметный указатель

доходность логарифмическая, 17

доходность простая, 17

процесс GARCH, 18

Список обозначений

Оглавление

1	Автокорреляция ошибок в линейной модели	5
2	Стационарные процессы, ARMA	9
3	GARCH	17
4	Единичный корень	23
5	Векторная авторегрессия	25
6	Модели состояние-наблюдение	27
7	Решения и ответы к избранным задачам	29