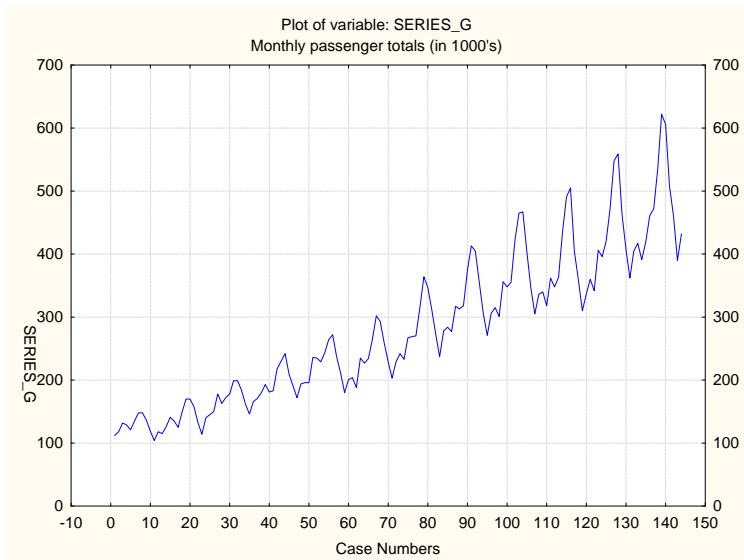


М.М.Осипчук

Аналіз часових рядів



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Прикарпатський національний університет
імені Василя Стефаника
Кафедра математичного і функціонального аналізу

Осипчук М.М.

**АНАЛІЗ
ЧАСОВИХ РЯДІВ**

Івано-Франківськ – 2020

Осипчук М.М. **Аналіз часових рядів**. Навчальний посібник. – Івано-Франківськ: , 2020. – 84 с.
Бібліографія: 21 назв.

Посібник призначено для студентів та аспірантів, що вивчають можливість застосування методів математичної статистики при аналізі процесів, що розвиваються в часі під впливом випадкових чинників. Для засвоєння пропонованого матеріалу необхідні знання вищої математики та основ теорії ймовірностей і математичної статистики.

Рецензент: кандидат фіз.-мат. наук, доцент Івасюк І.Я.

© Осипчук М.М., 2020

Макет підготовлено автором з використанням \LaTeX

Зміст

I	Теоретична частина	5
	Передмова	7
1	Часові ряди	9
	1.1 Адитивна та мультиплікативна моделі часових рядів	9
	1.2 Міри точності прогнозів	11
	1.3 Лаговий оператор	14
	1.4 Стаціонарність часових рядів	15
	1.5 Функція автокореляції	16
	1.6 Функція правдоподібності	17
	1.7 Порядок аналізу часових рядів	20
	1.8 Розбиття часових рядів	22
	1.8.1 Пропорційне розбиття	22
	1.8.2 Поліноміальна інтерполяція	24
2	Методи згладжування часових рядів	25
	2.1 Методи усереднення (moving average)	25
	2.2 Методи експоненціального згладжування	26
	2.3 Інші методи	27
3	Стаціонарні та звідні до них часові ряди	29
	3.1 Білий шум	29
	3.2 $MA(q)$ -процес	30
	3.3 Процес авторегресії	32
	3.4 Перетворення $MA(q)$ -процесів	38
	3.5 $ARMA$ -процес	39
	3.6 Прогнозування на основі $ARMA$ -моделей	42
	3.6.1 Прогнозування на основі залишків	42

3.6.2	Формула Вінера-Колмогорова	44
3.7	Оцінювання невідомих параметрів	50
3.8	ARIMA-процеси	51
3.9	Аналіз часових рядів Бокса-Дженкінса	54
4	Новітні моделі та методи аналізу часових рядів	61
4.1	Огляд моделей та методів	61
4.1.1	Лінійні моделі	61
4.1.2	Моделі часових рядів зі змінною волатильністю.	62
4.1.3	Нелінійні моделі	64
4.1.4	Нейронні мережі	66
4.1.5	Векторні часові ряди	68
4.2	Коротка бібліографія	69
II	Практична частина	73
5	Завдання для практичних занять	75
5.1	Початковий аналіз часового ряду	75
5.2	Згладжування, виділення тренду, прогнозування	76
5.3	Аналіз ARMA-моделей, прогнозування	77
5.4	Аналіз ARIMA-процесів, прогнозування	77
6	Самостійна робота та контроль знань	79
6.1	Задачі	79
6.2	Завдання для контролю знань	82

Частина I

Теоретична частина

Передмова

Протягом останніх десяти – п'ятнадцяти років теорія часових рядів переживає справжню революцію. З'явилась величезна кількість нових прикладних статистичних моделей, ґрунтовне математичне дослідження яких тільки починається. Звичайно, запровадження і вивчення нових моделей було обумовлене насамперед практичними потребами. У зв'язку з суттєво якісними змінами економічних ринків, виникненням нових технологій, інформаційною революцією, відповідно зміною поведінки людини і всього суспільства, стало зрозуміло, що потрібні нові математичні засоби для їх адекватного опису і прогнозування. Старі класичні методи, які базувались на припущеннях стаціонарності вже не відображали швидкозмінних, нелінійних процесів. З іншого боку, після ґрунтовного розвитку теорії стаціонарних, лінійних, гаусових моделей природним наступним кроком стала відмова від таких припущень, узагальнення отриманих результатів. Спочатку це була адаптація розробленої математичної техніки та засобів до моделей, які мали в собі певні елементи схожості до вже вивчених. На сучасному етапі з'являється все більше якісно нових моделей та методів їхнього дослідження. Можна вважати, що зараз ці розділи теорії часових рядів у стадії становлення.

Класичний підхід до побудови моделі часового ряду полягає у розкладі його на декілька компонент, природа кожної з яких принципово відмінна від інших. Далі кожна складова аналізується специфічними для неї методами. Найчастіше на практиці розглядають розклад часового ряду y_t :

$$y_t = tr_t + s_t + r_t$$

Компонента tr_t відповідає складовій, яка змінюється повільно і називається трендом. Це може бути як просто стала, так і многочлен, експонента чи інша подібна функція. Тренд звичайно описує довготривалі тенденції явища. Компонента s_t відповідає складовій, яка змінюється періодично і називається сезонною або циклічною компонентою. Ця компонента часто описує сезонні явища у виробництві протягом року або економічні циклічності. Звичайно складові tr_t та s_t детерміністичні функції. За випадковість у спостереженнях відповідає r_t . Основний напрямок розвитку теорії часових рядів якраз є моделювання і аналіз процесу r_t .

Розділ 1

Часові ряди

В сучасній статистичній теорії існує багато різноманітних методів прогнозування інформації. Значна їх частина відноситься до прогнозування часових рядів. Особливістю прогнозування часових рядів є те, що аналізуються лише дані спостережень без додаткової інформації, без аналізу впливу зовнішніх сил. Звичайно, такий аналіз виглядає досить неповним, але доволі часто прогнози часових рядів є більш точними. Нехай y_1, y_2, \dots, y_T – значення спостережень за деяким процесом протягом T періодів. Ця послідовність є числовими значеннями, кожне з яких має відповідний індекс, який залежить від номера періоду, в який він спостерігався. Така послідовність, записана у порядку зростання індексу, називається часовим рядом. Будемо позначати часовий ряд з T елементами $\{Y_T\}$.

1.1 Адитивна та мультиплікативна моделі часових рядів

Будь-який часовий ряд можна представити як суму детермінованого та випадкового компонентів:

$$y_t = d_t + r_t, \quad t = \overline{1, T}.$$

В свою чергу детермінований компонент складається з трьох частин: трендового, сезонного, циклічного компонентів.

$$d_t = tr_t + s_t + c_t, \quad t = \overline{1, T}.$$

Таким чином, будь-який часовий ряд можна розглядати як суму:

$$y_t = tr_t + s_t + c_t + r_t, \quad t = \overline{1, T}.$$

Такий вигляд часового ряду отримав назву адитивної моделі. Якщо ж замість реальних значень компонентів використовувати їх логарифми, то отримаємо мультиплікативну модель ($t = \overline{1, T}$):

$$\ln y_t = \ln tr_t + \ln s_t + \ln c_t + \ln r_t, \quad \text{або} \quad y_t = tr_t \cdot s_t \cdot c_t \cdot r_t.$$

Детермінований компонент змінюється за певними правилами, які можуть бути визначені за допомогою досліджень і відповідного аналізу часового ряду. Як правило, одним з основних параметрів, від яких залежить детермінований компонент, є час.

Аналіз часового ряду починається з виділення трендового компонента. Його присутність неважко помітити, проаналізувавши графік часового ряду. Дослідники мають змогу описувати такі зміни за допомогою кривих, які можна задати в аналітичному вигляді.

Сезонний компонент показує коливання навколо трендового компонента. Наприклад, його наявність для економічних часових рядів пояснюється сезонним характером виробництва та споживання. Головна ідея виділення сезонних коливань полягає у порівнянні даних за відповідні періоди, а не за минулі періоди.

Циклічний компонент займає проміжне місце між трендом та сезонним компонентом. Тренд – це гладка зміна, яка проявляється на великому проміжку часу. Сезонний компонент – це періодична функція, що залежить від часу, причому його період значно менший за період спостережень. Циклічний компонент розглядається, в основному, як гладка зміна, залежна від часу, але яка не включається ні до тренду, ні до сезонного компонента.

Випадковий компонент є те, що залишилось від часового ряду після виключення тренду, циклічного та сезонного компонентів. Частина таких ефектів може бути віднесена до непередбачених природних катаклізмів (землетруси, пожежі, тощо), частина – до випадкових дій людей. За наявності випадкової компоненти неможливо прогнозувати значення часового ряду без помилки. Але будь-який реальний процес включає випадковий компонент.

Від значення кожного компонента залежить значення часового ряду у кожний період спостережень. Але не завжди представляється можливим характеризувати кожний компонент окремо. Іноді

краще робити прогноз відносно всієї моделі, ніж намагатися виділити кожний компонент окремо.

1.2 Міри точності прогнозів

Розробка прогнозу вимагає оцінки його точності та надійності. Точність і надійність прогнозів – широко поширені в прогностичній літературі терміни, сенс яких, як це уявляється на перший погляд, цілком очевидний. Проте зміст цих термінів часто тлумачиться достатньо суб'єктивно.

Про точність прогнозу прийнято судити по розміру помилки прогнозу – різниці між прогнозним і фактичним значенням досліджуваної змінної. Проте такий підхід до оцінки точності можливий тільки в двох випадках. По-перше, коли період попередження вже закінчився, і дослідник має фактичні значення змінної. При короткостроковому прогнозуванні це цілком реально. По-друге, коли прогноз розробляється, тобто прогнозування здійснюється для деякого моменту часу в минулому, для котрого вже є фактичні дані. При цьому наявна інформація ділиться на дві частини. Одна з них, що охоплює більш ранні дані, служить для оцінювання параметрів прогностичної моделі, а більш пізні дані розглядаються як реалізації відповідних прогностичних оцінок. Отримані ретроспективно помилки прогнозу якоюсь мірою характеризують точність застосованої методики прогнозування і можуть виявитися корисними при зіставленні декількох методів. У той же час, розмір помилки ретроспективного прогнозу не можна розглядати як остаточний доказ придатності, або навпаки, непридатності застосовуваного методу прогнозування. До неї варто ставитися з відомою обережністю і при її застосуванні в якості міри точності необхідно враховувати, що вона отримана при використанні лише частини наявних даних. Проте ця міра точності має більшу наочність і теоретично більш надійна, ніж похибка прогнозу, обчислена для періоду, характеристики котрого вже були використані при оцінюванні параметрів моделі. У останньому випадку похибки, як правило, будуть незначними і мало залежні від теоретичної обґрунтованості, застосованої для прогнозування моделі.

У зв'язку з перевіркою точності прогнозів необхідно зробити ще одне зауваження. Так, якщо для ретроспективного прогнозування застосовується модель, що містить одну або декілька екзогенних

змінних, то точність прогнозу буде значною мірою залежати від того, наскільки точно визначені значення цих змінних на період попередження. При цьому можливі два шляхи: скористатися фактичними значеннями екзогенних змінних (так званий прогноз *ex post*) і очікуваними їхніми значеннями (так званий прогноз *ex ante*). Природно, що точність прогнозу *ex post*, що, як правило, і одержують при перевірці, буде вище, ніж точність прогнозу *ex ante*, тому що в першому випадку буде виключений вплив похибки у значеннях екзогенних змінних.

Перевірка точності одного прогнозу мало що може сказати досліднику. Гарний одиничний прогноз може бути отриманий і по погкій моделі, і навпаки. Звідси випливає, що про якість прогнозів застосовуваних методик і моделей можна судити лише по сукупності зіставлень прогнозів і їхньої реалізації.

Найбільш простою мірою якості прогнозів за умови, що є дані про їхню реалізацію, може стати відносне число випадків, коли фактична реалізація охоплювалася інтервальним прогнозом, до загального числа прогнозів, тобто

$$\eta = \frac{m}{m+n},$$

де m – кількість прогнозів, підтверджених фактичними даними; n – кількість прогнозів, не підтверджених фактичними даними.

Для міри точності прогнозів частіше використовуються такі характеристики: MSE, RMSE, MAD, MAPE. Нехай \hat{y}_t – прогноз значення часового ряду у t -тому періоді, тоді:

$$MSE = \frac{1}{n} \sum_t (y_t - \hat{y}_t)^2$$

– середнє квадратів похибок прогнозу за n кроків.

$$RMSE = \sqrt{MSE}$$

– середньоквадратична похибка прогнозу за n кроків.

$$MAD = \frac{1}{n} \sum_t |y_t - \hat{y}_t|$$

– середня абсолютна похибка за n кроків.

$$RMSPPE = 100 \sqrt{\frac{1}{n} \sum_t \left(\frac{y_t - \hat{y}_t}{y_t} \right)^2}$$

– середньоквадратичне значення відносних до фактичних значень похибок за n кроків (у відсотках).

$$MAPE = \frac{100}{n} \sum_t \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

– середнє значення абсолютних величин відносних похибок за n кроків (у відсотках).

На практиці ці характеристики використовуються досить часто. Перші три критерії виражають похибку у одиницях виміру і тому їх величина залежить від специфіки часового ряду. Останні два критерії вимірюються у відносних одиницях, тому можна говорити про деякий загальний рівень адекватності моделі на основі їх порівняння.

MAPE, RMSE	Точність прогнозу
менше 10%	Висока
10% – 20%	Добра
20% – 40%	Задовільна
40% – 50%	Погана
більше 50%	Незадовільна

Взагалі кажучи, наведені значення не є постійними і змінюються в залежності від мети прогнозування. Дослідник сам встановлює межі показника точності, які його задовольняють.

Вище наведені міри якості прогнозів розглядалися за умови, що дослідник має інформацію про істинні значення часового ряду, що він оцінював у ході розробки прогнозів. Такі міри якості, безсумнівно, являють цінність при вивченні різноманітних методик прогнозування. Проте в практичній роботі проблему точності прогнозу треба вирішувати, як правило, тоді, коли період попередження ще не пройшов, і істинне значення прогнозованої змінної невідомо. У цьому випадку проблема точності може розглядатися в плані зіставлення апріорних якостей, властивостей, властивим альтернативним прогностичним моделям.

Ефективність методу прогнозування залежить від багатьох чинників. На практиці дослідник має досить велику свободу вибору не тільки типу моделі, але і кількості введених до неї параметрів. Вділяють такі критерії:

- кількість зусиль, що витрачаються на побудову моделі і наявність готових машинних програм;
- швидкість, із яким метод уловлює істотні зміни у поведінці ряду, наприклад, раптовий зсув математичного сподівання або збільшення кута нахилу лінії тренду;
- існування серійної кореляції у помилках;
- незмінюваність первинних даних;
- повний обсяг роботи в деяких сферах діяльності - тисячі рядів щомісяця потребують оновлення, невеликі витрати і швидкість мають першорядне значення;
- терміновість прогнозування.

Практика розробки різноманітного роду прогнозів спирається на цілу систему методів, серед яких статистичні методи прогнозування займають важливе місце. Вирішальну роль у статистичному підході до прогнозування грає вибір відповідної моделі, що, будучи наповненої числовими параметрами, стає безпосереднім інструментом прогнозування – предиктором. Володіючи предиктором, можна одержати варіанти прогнозу, що відповідають визначеним гіпотезам і умовам, врахованим при його побудові. Мета статистичної моделі – не замінити судження і досвід спеціаліста, а дати йому в руки інструмент, що дозволяє більш глибоко проникнути в сутність досліджуваних явищ, інструмент, у якому специфічним чином узагальнена і зведена у систему різноманітна статистична інформація. Одержувані на основі предикторів прогнози мають сенс тільки в рамках тих умов, гіпотез і припущень, що були враховані при розробці відповідних статистичних моделей і при їхньому застосуванні для прогнозування. Таким чином, розробка і застосування моделей у прогностичних цілях припускають поглиблений статистичний аналіз досліджуваного об'єкта.

1.3 Лаговий оператор

Особливе значення при роботі з часовими рядами має лаговий оператор. Зазвичай його позначають B (від англійського back-shift) або

L (від німецького *der Lagoperator*). За його допомогою можна отримувати значення часового ряду як функції від його інших значень. Зокрема, $B y_t = y_{t-1}$ і, отже, $B^k y_t = y_{t-k}$, $B^{-k} y_t = y_{t+k}$, $B^0 y_t = y_t$. Для довільного полінома з дійсними коефіцієнтами

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

маємо значення дії поліноміального лагового оператора на значення часового ряду

$$P_n(B)y_t = a_0 y_t + a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_n y_{t-n}.$$

1.4 Стаціонарність часових рядів

Для будь-якого часового ряду можна знайти деякі числові характеристики. Аналогічно методам математичної статистики, для часового ряду, як і для будь-якої вибірки, знаходять основні моменти. Для аналізу часових рядів найважливішими моментами є математичне сподівання, дисперсія, коваріація. Математичним сподіванням часового ряду $\{Y_T\}$ є функція

$$\mu_t = \mathbf{E}y_t = \int_{-\infty}^{+\infty} x dF_t(x),$$

де $F_t(x) = \mathbf{P}(y_t < x)$ – функція розподілу y_t , $t = \overline{1, T}$.

Дисперсія часового ряду $\{Y_T\}$ визначається за формулою:

$$\mathbf{D}(y_t) = \mathbf{E}(y_t - \mu_t)^2.$$

Нарешті, автоковаріація часового ряду $\{Y_T\}$ дорівнює:

$$\text{cov}(y_t, y_{t-j}) = \mathbf{E}(y_t - \mu_t)(y_{t-j} - \mu_{t-j}), \quad j = 1, 2, \dots$$

Часовий ряд є стаціонарним, якщо

- математичне сподівання $\mu_t = \mu < \infty$ для всіх $t = \overline{1, T}$,
- дисперсія $\mathbf{D}(y_t) = \gamma_0 < \infty$ для всіх $t = \overline{1, T}$,
- автоковаріація j -того порядку $\text{cov}(y_t, y_{t-j}) = \gamma_j < \infty$ для всіх $t = \overline{1, T}$, $j = 1, 2, 3, \dots$

Таким чином, стаціонарний часовий ряд повинен мати сталі та скінченні математичні сподівання та дисперсію для кожного періоду часу, а також сталу і скінченну автоковаріацію будь-якого порядку в усі моменти часу.

Оцінками цих моментів часового ряду є статистики:

- математичне сподівання

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T y_t;$$

- дисперсія

$$\hat{\mu}_0 = \frac{1}{T-1} \sum_{t=1}^T (y_t - \hat{\mu})^2;$$

- автоковаріація j -того порядку

$$\hat{\mu}_j = \frac{1}{T-j} \sum_{t=j+1}^T (y_t - \hat{\mu})(y_{t-j} - \hat{\mu}).$$

Вище наведені значення можуть бути доволі різними для різноманітних часових рядів. Було б важко розробити методи для аналізу такої несхожої інформації. Тому був розроблений поділ часових рядів на класи в залежності від властивостей його числових характеристик. Одним з найважливіших поділів часових рядів є поняття стаціонарності. Звичайно, жоден з рядів, що представляє реальну інформацію про процес, який відбувається в часі, не може бути ідеально стаціонарним. Але якщо для деякого часового ряду з деяким наближенням виконуються умови стаціонарності, то для його аналізу можна використати широкий спектр методів аналізу та прогнозування стаціонарних часових рядів. Така методологія детально розглядається у третьому розділі.

1.5 Функція автокореляції

Крім вищенаведених характеристик при аналіз часових рядів застосовується автокореляція та автокореляційна функція. Автокореляція j -того порядку визначається за допомогою рівності:

$$\rho_j = \frac{\text{cov}(y_t, y_{t-j})}{\mathbf{D}(y_t)} = \frac{\gamma_j}{\gamma_0}.$$

Цей коефіцієнт визначає ступінь залежності між спостереженнями, які знаходяться на відстані j періодів. Якщо побудувати залежність ρ_j на графіку, то отримаємо корелограму. Вона представляє деяку криву, що показує як змінюється взаємовплив між спостереженнями в залежності від відстані між ними в часі. Тлумачення значень функції автокореляції та її поведінки не є легкою справою. Розглянемо декілька можливостей.

1. *Випадковий процес.* Якщо спостерігається абсолютно випадковий процес з великою кількістю даних, то $\rho_j = 0$ при всіх ненульових значеннях j .
2. *Короткотермінова залежність.* Стаціонарні процеси часто густо мають короткотермінову залежність між спостереженнями, що проявляється в тому, що декілька перших коефіцієнтів автокореляції є відносно великими в абсолютному значенні по відношенню до всіх інших, які наближаються до нуля. Якщо перші коефіцієнти є додатними, то присутня додатна автокореляція, якщо ж перші коефіцієнти автокореляції постійно змінюють свій знак, то спостерігається від'ємна автокореляція.
3. *Нестаціонарні процеси.* Часові ряди, які мають чіткий трендовий компонент, значення ρ_j не буде наближатися до нуля, навіть при досить великих j . Тому, як правило, підрахунок функції автокореляції застосовують тільки для стаціонарних часових рядів.
4. *Сезонні коливання.* При наявності сезонних коливань корелограма представлятиме, як правило, затухаючі хвилі однакової частоти. Якщо виділити сезонний компонент, то функція автокореляції покаже справжню залежність між спостереженнями у часі.

Для кращого розуміння корелограми рекомендується будувати залежність y_{t-j} від y_t і не забувати, що для будь-якого j вірна рівність $\rho_j = \rho_{-j}$.

1.6 Функція правдоподібності

На сучасному етапі розвитку статистики, метод максимальної правдоподібності став ключовим методом аналізу. Загальне розповсю-

дження цього методу викликано багатьма причинами. По-перше, за його допомогою можна отримувати найкращі оцінки. По-друге, ще у 1976 році було показано, що багато методів оцінювання, такі як двох- та трьохетапний метод найменших квадратів та ін., можуть бути інтерпретовані як наближення до оцінок, отриманих за методом максимальної правдоподібності. Іншими словами, цей метод дозволяє отримувати спроможні та асимптотично ефективні оцінки. По-третє, цей метод є дуже загальним, що дозволяє використовувати його для широкого класу задач. По-четверте, незважаючи на відносну складність методу та необхідність чітко знати форму функції розподілу, з розвитком комп'ютерної техніки, вчені позбулися проблем з підрахунком необхідних оцінок. По-п'яте, за допомогою відповідних функцій правдоподібності відбувається тестування гіпотез. Таким чином, можна сказати, що зараз метод максимальної правдоподібності є універсальним методом отримання оцінок при аналізі часових рядів. Основна ідея застосування методу полягає в тому, що ми припускаємо, що дані мають деякий імовірнісний розподіл і обраховуємо ймовірність шуканої події. Це у загальному випадку залежить від деяких невідомих параметрів. Використовуючи дані, ми можемо максимізувати ймовірність цієї події. Коефіцієнти, при яких досягається максимум ймовірності відповідної події є необхідними оцінками параметрів. Іноді дуже важко знайти ці оцінки у аналітичному вигляді. В такому разі використовуються чисельні методи оптимізації функції правдоподібності.

Максимізація функції правдоподібності Розглянемо побудову функції правдоподібності більш детально. Нехай $\{Y_T\}$ є вибірка, яка має ймовірнісний розподіл $\mathbf{P}(y/A)$, де A – набір невідомих параметрів. Припустимо, що y_t є незалежними, кожне із ймовірнісним розподілом $\mathbf{P}(y_t/A)$, та сумісний розподіл всієї сукупності $\{Y_T\}$ задається рівністю

$$\mathbf{P}(y_1, y_2, \dots, y_T/A) = \prod_{t=1}^T \mathbf{P}(y_t/A).$$

Для відповіді на питання, яке саме значення A максимізує ймовірність спостереження саме такої вибірки $\{Y_T\}$, ми повинні максимі-

зувати функцію правдоподібності:

$$L(A) = \prod_{t=1}^T \mathbf{P}(y_t/A).$$

Для подальшої оптимізації необхідно точно знати розподіл вибірки. В загальному випадку ні функція правдоподібності, ні її логарифм не є лінійними, тому знайти максимум функції правдоподібності в аналітичному вигляді дуже важко. Таким чином потрібно використовувати чисельні методи знаходження максимуму функції. Наведемо загальний алгоритм таких методів.

1. Покласти початкові значення для вектора A .
2. Визначаємо напрям руху для A_i , в якому значення $L(A_i)$ збільшується.
3. Визначаємо довжину кроку та визначаємо нове значення A_{i+1} .
4. Перевіряємо критерій зупинки; якщо необхідно алгоритм продовжити, то покладаємо $i = i + 1$ та повертаємося до кроку 2. Звичайним критерієм зупинки є $|L(A_{i+1}) - L(A_i)| < \delta$, де δ – наперед задане мале число.

Перевірка гіпотез за допомогою функції правдоподібності

Тепер розглянемо критерії перевірки гіпотези H_0 про значення параметрів процесу проти альтернативної H_1 в загальному випадку (без жодних обмежень). Існує три основних класи тестових статистик: тест Вальда, тест за допомогою множників Лагранжа, тест на основі відношень значень функції правдоподібності. Всі ці три критерії мають своєю основою максимізацію функції правдоподібності. Вони є асимптотично еквівалентними. Ключовою різницею між цими трьома підходами є вибір оцінки для розрахунків.

Метод відношення значень функції правдоподібності є найстарішим з усіх цих тестів і був розроблений Нейманом та Пірсоном у 1928 році. Основна суть методу полягає у порівнянні значень функції правдоподібності при умові H_0 та при умові H_1 . Відношення значень функції правдоподібності свідчить на користь однієї чи іншої гіпотези. Статистика

$$LRT = 2(\ln L(A_0) - \ln L(A_1)),$$

де A_0, A_1 відповідно множини параметрів, що відповідають гіпотезам H_0, H_1 , має розподіл Пірсона (χ_m^2) з кількістю ступенів волі рівною кількості обмежень щодо параметрів в гіпотезі H_0 . Порівнявши LRT з квантилем χ_m^2 порядку $1 - \alpha$ розподілу Пірсона, робимо висновок про узгодження гіпотези H_0 із спостереженнями, якщо $LRT < \chi_m^2$, і гіпотеза відхиляється в іншому випадку. При цьому ймовірність помилки обмежується зверху значенням α .

1.7 Порядок аналізу часових рядів

Звичайно, метою прикладного статистичного аналізу часових рядів є побудова математичної моделі ряду, за допомогою якої можна пояснити поведінку ряду і здійснити прогноз на майбутні періоди.

Побудова і вивчення графіка. Аналіз часового ряду починається з побудови і вивчення його графіка. Якщо нестационарність часового ряду очевидна, то першою справою треба виділити нестационарну складову ряду. Процес виділення тренду та інших компонент ряду, що призводять до порушення стаціонарності, може проходити в декілька етапів. На кожному з них розглядається ряд залишків, отриманий у результаті вирахування з вихідного ряду підбраної моделі тренду, або результат різницевих і інших перетворень ряду. Крім графіків, ознаками нестационарності часового ряду можуть служити автокореляційна функція, що прямує не до нуля (за винятком дуже великих значень лагів) і наявність яскраво виражених піків на низьких частотах у періодограмі. За допомогою автокореляційної функції досліджують також внутрішні зв'язки між елементами часових рядів.

У вибіркових дослідженнях найпростіші числові характеристики описової статистики (середнє, медіана, дисперсія, стандартне відхилення, коефіцієнти асиметрії й ексцесу) звичайно дають достатньо інформативне уявлення про вибірку. Графічні методи уявлення й аналізу вибірок при цьому грають лише допоміжну роль, дозволяючи краще зрозуміти локалізацію і концентрацію даних, їхній закон розподілу. Роль графічних методів при аналізі часових рядів цілком інша. Табличне уявлення тимчасового ряду й описових статистиків часті усього не дозволяє зрозуміти характер процесу, у той час як за графіком тимчасового ряду можна зробити досить багато висновків.

Надалі вони можуть бути перевірені й уточнені за допомогою розрахунків. Людське око досить упевнено визначає за графіком часового ряду:

- наявність тренду і його характер;
- наявність сезонних і циклічних компонент;
- ступінь повільності або переривчастості змін послідовних значень ряду після усунення тренду (по цьому показнику можна судити про характер і розмір кореляції між сусідніми елементами ряду).

Так графічний аналіз ряду звичайно задає напрямок його подальшого аналізу.

Вибір моделі для часового ряду. Після того, як вихідний процес максимально наближений до стаціонарного, можна приступити до вибору різноманітних моделей отриманого процесу. Мета цього етапу – опис і урахування надалі аналізу кореляційної структури аналізованого процесу. Модель може вважатися підбраною, якщо залишкова компонента ряду є процесом типу, як правило, "білого шуму"¹. Після підбору залишки аналізуються для перевірки адекватності моделі та побудови надійних інтервалів.

Прогнозування або інтерполяція. Останнім етапом аналізу часового ряду може бути прогнозування його майбутніх (екстраполяція) або відновлення пропущених (інтерполяція) значень і визначення точності цього прогнозу на базі підбраної моделі. Добре підібрати математичну модель вдається не для всякого часового ряду. Нерідко буває і так, що для опису підходять відразу декілька моделей. Нездозначність вибору моделі може спостерігатися як на етапі виділення детермінованого компонента ряду, так і при виборі структури ряду залишків. Тому досить часто розробляють декілька прогнозів, зроблених за допомогою різних моделей.

¹Поняття білого шуму буде розглянуто в пункті 3.1

1.8 Розбиття часових рядів

Різні часові ряди можуть бути утворені за допомогою спостережень над певними процесами через різні часові інтервали. Наприклад, більшість макроекономічної інформації подається щоквартально або щорічно. Тому іноді виникає проблема співставлення таких даних з іншими, що мають, наприклад, місячну структуру. Можлива також і така ситуація: для одного чи декількох років не подана щоквартальна розбивка. Тобто відсутня інформація за підперіоди деяких періодів часового ряду. Таким чином, для того щоб можна було використовувати всю наявну інформацію, необхідно вміти розбивати, наприклад, щорічні дані на кварталні. Припустимо, що ми розглядаємо часовий ряд $\{Y_T\}$ з річною структурою даних. Наша задача – утворити новий часовий ряд $\{X_{4T}\}$, який буде відповідати за значення процесу у кварталах кожного року. Якщо величина, яку представляє даний часовий ряд має властивість адитивності, тобто сума її значень за всі підперіоди, дорівнює її значенню за цілий період, то очевидно, що така розбивка часового ряду повинна проводитися за умови $y_t = x_{4t-3} + x_{4t-2} + x_{4t-1} + x_{4t}$. Якщо ж часовий ряд є результатами простих спостережень над значеннями деякого процесу в рівновіддалені моменти часу (наприклад, рівень води в річці 28 числа кожного місяця), то для ряду, що відображає значення цього процесу в проміжні моменти часу (наприклад, 7, 14, 21 і 28 чисел кожного місяця) повинно, очевидно, виконуватись $y_t = x_{4t}$.

Ми розглянемо основні методи, які застосовуються для утворення часового ряду $\{X_{4T}\}$. В першому випадку таке розбиття ми будемо називати дезаграгуванням часового ряду, а ситуаціях, подібних на другий випадок, – відновленням. Відразу зауважимо, що подібна методика може бути застосована і для інших співвідношень в часових структурах рядів.

1.8.1 Пропорційне розбиття

Дезаграгування. Якщо відома розбивка якогось року по кварталах, то ми можемо використати її для утворення нового ряду з кварталною структурою. Нехай для деякого року τ відомі кварталні значення $d_{\tau 1}, d_{\tau 2}, d_{\tau 3}, d_{\tau 4}$. Тоді новий ряд будується за правилом:

$$x_{4t-4+i} = \alpha_i y_t, \quad i = \overline{1, 4}, \quad \text{де } \alpha_i = \frac{d_{\tau i}}{y_{\tau}}$$

Цей метод утворить новий ряд, який матиме пропорційну структуру. Якщо дослідник має квартальну розбивку по декількох роках, номери яких складають множину $S = \{t_1, t_2, \dots, t_n\}$, то

$$\alpha_i = \sqrt[n]{\prod_{\tau \in S} \frac{d_{\tau i}}{y_{\tau}}}.$$

Якщо ж множина S є порожньою, тобто невідомо розбивки по кварталах жодного року, то використовується елементарне усереднення:

$$\alpha_i = \frac{1}{4}, \quad i = \overline{1, 4}.$$

Але слід враховувати, що таке розбиття виключає сезонні коливання.

Відновлення Нехай знову кожен період (наприклад, рік) ряду $\{Y_T\}$ містить в собі чотири підперіоди (квартали), які будуть періодами в ряді $\{X_4T\}$ і для деякого року τ відомі квартальні значення $d_{\tau 1}, d_{\tau 2}, d_{\tau 3}, d_{\tau 4}$. Очевидно, що $y_{\tau} = d_{\tau 4}$.

Розглянемо різниці ($\tau > 1$)

$$\begin{aligned} \Delta_{\tau 1} &= d_{\tau 1} - y_{\tau-1}, \\ \Delta_{\tau 2} &= d_{\tau 2} - d_{\tau 1}, \\ \Delta_{\tau 3} &= d_{\tau 3} - d_{\tau 2}, \\ \Delta_{\tau 4} &= d_{\tau 4} - d_{\tau 3} = y_{\tau} - d_{\tau 3}. \end{aligned}$$

Обчислимо коефіцієнти пропорційності $\alpha_i = \frac{\Delta_{\tau i}}{y_{\tau} - y_{\tau-1}}, i = \overline{1, 4}$. Тоді для періоду t

$$\begin{aligned} x_{4t-3} &= (y_t - y_{t-1})\alpha_1 + y_{t-1}, \\ x_{4t-2} &= (y_t - y_{t-1})\alpha_2 + x_{4t-3}, \\ x_{4t-1} &= (y_t - y_{t-1})\alpha_3 + x_{4t-2}, \\ x_{4t} &= (y_t - y_{t-1})\alpha_4 + x_{4t-1}. \end{aligned}$$

Оскільки, очевидно, що $\sum_{i=1}^4 \Delta_{\tau i} = 1$, то $x_{4t} = y_t$.

Аналогічно попередньому випадку, коли маємо квартальну розбивку по декількох роках, коефіцієнти пропорційності обчислюємо

за формулою

$$\alpha_i = \sqrt[n]{\prod_{\tau \in S} \frac{\Delta_{\tau i}}{y_{\tau} - y_{\tau-1}}}.$$

1.8.2 Поліноміальна інтерполяція

Деаграгування. Цей метод спочатку утворює новий ряд за правилом

$$z_t = \sum_{j=1}^t y_j.$$

Ряд $\{Z_T\}$ представляє собою акумулятивні суми початкового ряду. Після цього, послідовно через чотири точки ряду $\{Z_T\}$ будується кубічна інтерполяція, тобто знаходяться коефіцієнти функції

$$f(t) = at^3 + bt^2 + ct + d.$$

Наприклад, для розбиття y_3 нам необхідно побудувати інтерполяцію по точках z_1, z_2, z_3, z_4 : $f(1) = z_1, f(2) = z_2, f(3) = z_3, f(4) = z_4$.

Тепер обчислимо значення $f(2.25), f(2.5), f(2.75)$. Тоді

$$\begin{aligned} x_9 &= f(2,25) - f(2), \\ x_{10} &= f(2,5) - f(2,25), \\ x_{11} &= f(2,75) - f(2,5), \\ x_{12} &= f(3) - f(2,75). \end{aligned}$$

Легко бачити, що $x_9 + x_{10} + x_{11} + x_{12} = f(3) - f(2) = y_3$.

Відновлення здійснюється звичайною інтерполяцією.

Розділ 2

Методи згладжування часових рядів

Методи згладжування використовуються для зменшення впливу випадкового компонента (випадкових коливань) у часових рядах. Вони дають можливість отримувати більш "чисті" значення, які складаються лише з детермінованих компонентів. Деякі з методів направлені на виділення лише деяких компонентів, наприклад, тренду.

2.1 Методи усереднення (moving average)

Цей метод є одним з найпростіших, який дозволяє виділити тренд. Для застосування цього методу дослідник повинен мати доволі довгий ряд спостережень. Формально метод описується виразом:

$$\tilde{y}_t = \frac{1}{k} \sum_{j=-k_1}^{k_2} y_{t+j}, \quad k = k_1 + k_2 + 1.$$

Члени нового часового ряду є середніми значеннями певної кількості сусідніх до відповідного члена даного ряду. З формули видно, що нова кількість спостережень становить $T - k$. Єдиною складністю є визначення чисел k_1 та k_2 . Як правило, їх сума дорівнює повному циклу сезонності.

Подвійне усереднення Цей метод двічі використовує усереднення часового ряду. При цьому кількість спостережень зменшується на два повних цикли сезонності, тому для використання методу необхідно мати часовий ряд, який складається щонайменше з 3-х повних циклів сезонності.

2.2 Методи експоненціального згладжування

Звичайне експоненціальне згладжування Цей метод значно переважає усі попередні моделі. Найкраще цей метод зарекомендував себе, коли дані мають дуже гладкий, або навіть горизонтальний тренд. Нова послідовність будується за правилом:

$$S_t = \alpha y_t + (1 - \alpha)S_{t-1}, \quad 0 < \alpha < 1.$$

Початкове значення S_1 , треба вибирати доволі обережно, бо від нього залежить, як саме буде поводити себе згладжена послідовність. Найчастіше вибирають $S_1 = y_1$, або $S_1 = \mu$. Однак з деяких міркувань може бути обране й інше значення S_1 .

Ваговий коефіцієнт α може обиратися кількома шляхами. По-перше, якщо обирається значення близьке до 1, то будуть більш важливими при прогнозуванні останні дані часового ряду, при виборі α близьким до 0, більш впливовими будуть минулі значення. По-друге, можна покласти $\alpha = \frac{2}{T+1}$. По-третє – вибір α , при якому мінімізується один з критеріїв точності прогнозів на n періодів. При цьому розрахунки повинні проводитися лише по перших $T - n$ значеннях часового ряду, а отримані прогнози повинні бути порівняні з реальними даними. При використанні цієї методики відкидається обмеження $0 < \alpha < 1$. Прогноз значень часового ряду дорівнює останньому члену послідовності S_t :

$$\hat{y}_{T+p} = S_T, \quad p = 1, 2, \dots$$

Подвійне експоненціальне згладжування Брауна Цей метод будується аналогічно попередньому, тільки процес згладжування робиться двічі:

$$\begin{aligned} S'_t &= \alpha y_t + (1 - \alpha)S'_{t-1} \\ S''_t &= \alpha S'_t + (1 - \alpha)S''_{t-1}, \quad 0 < \alpha < 1 \end{aligned}$$

Метод використовується, коли дані часового ряду нестационарні. Прогноз будується як останнє значення другої послідовності:

$$\hat{y}_{T+p} = S''_T, \quad p = 1, 2, \dots$$

Потрійне експоненціальне згладжування Брауна Цей метод аналогічний двом попереднім, тільки згладжування проводиться тричі. Це дозволяє прогнозувати нестационарні часові ряди з великими перепадами мінімального та максимального значень. Нові послідовності будуються за правилом:

$$\begin{aligned} S'_t &= \alpha y_t + (1 - \alpha)S'_{t-1} \\ S''_t &= \alpha S'_t + (1 - \alpha)S''_{t-1} \\ S'''_t &= \alpha S''_t + (1 - \alpha)S'''_{t-1}, \quad 0 < \alpha < 1 \end{aligned}$$

Прогноз на наступні періоди має вигляд:

$$\hat{y}_{T+p} = S'''_T, \quad p = 1, 2, \dots$$

Адаптивне згладжування Цей метод дозволяє автоматично змінювати константу згладжування в процесі обрахунку. Нова послідовність будується за правилом

$$S_t = \alpha_t y_t + (1 - \alpha_t)S_{t-1},$$

де α_t змінюється з часом в залежності від похибки прогнозування:

$$\begin{aligned} \alpha_t &= \left| \frac{E_t}{M_t} \right| \\ E_t &= \beta(y_t - \hat{y}_t) + (1 - \beta)E_{t-1}, \\ M_t &= \beta|y_t - \hat{y}_t| + (1 - \beta)M_{t-1} \end{aligned}$$

Параметр β знаходиться у межах $(0; 1)$. Для випадкових похибок коефіцієнт α буде близьким до 0,5.

2.3 Інші методи

Несезонна модель Холта-Вінтера. Ця модель схожа на подвійне експоненціальне згладжування, але дозволяє виділяти трендовий

компонент за допомогою другої послідовності:

$$\begin{aligned} S'_2 &= y_2, & S''_2 &= y_2 - y_1, \\ S'_t &= \alpha y_t + (1 - \alpha)(S'_{t-1} + S''_{t-1}), & 0 < \alpha < 1, \\ S''_t &= \beta(S'_t - S'_{t-1}) + (1 - \beta)S''_{t-1}, & 0 < \beta < 1. \end{aligned}$$

Прогноз на наступні періоди: $\hat{y}_{T+p} = S'_T + pS''_T$, $p = 1, 2, \dots$

Адитивна модель із визначенням сезонних коливань. В цьому методі визначаються трендовий та сезонний компоненти часового ряду. Нехай p – цикл сезонності, тобто $s_{t+p} = s_t$ для будь-якого t . Наша задача – оцінити значення s_t за спостереженнями y_t вважаючи, що величина p відома. За допомогою методу усереднення, оцінимо трендовий компонент tr_t . В результаті отримуємо десеzonолізовану (вільну від сезонних впливів) величину:

$$\hat{tr}_t = \begin{cases} \frac{1}{p}(y_{t-k} + \dots + y_t + \dots + y_{t+k}), & \text{якщо } p = 2k + 1; \\ \frac{1}{2p}(y_{t-k} + \dots + y_t + \dots + y_{t+k-1} + \\ + y_{t-k+1} + \dots + y_t + \dots + y_{t+k}), & \text{якщо } p = 2k, \end{cases}$$

при $t = \overline{k+1, T-k}$ Оцінкою сезонного компонента буде послідовність

$$\hat{s}_t = y_t - \hat{tr}_t, \quad t = \overline{k+1, T-k}.$$

Підррахуємо тепер середні значення \bar{s}_t оцінки сезонного компонента \hat{s}_t , для кожного періоду циклу сезонності. Так, для першого періоду це буде середнє арифметичне значень \hat{s}_t , для всіх $t = k+1 + pl$, для другого періоду – середнє арифметичне значень \hat{s}_t , для всіх $t = k+2 + pl$, де $l = 1, 2, \dots$ і т. д. Для всіх $t \geq p+1$ покладемо $\bar{s}_t = \bar{s}_{t-p}$.

Вважаючи, що сума сезонних складових ряду за один теріод дорівнює нулю, визначимо скоригований сезонний компонент за формулою $s_t^* = \bar{s}_t - \frac{1}{p} \sum_{t=1}^p \bar{s}_t$. Отримані значення s_t^* задовольняють, очевидно, рівність $\sum_{t=1}^p s_t^* = 0$ а також $s_t^* = s_{t-p}^*$ для всіх $t \geq p+1$. Ряд $tr_t = y_t - s_t^*$ використаємо для побудови функції, що задаватиме тренд даного часового ряду.

Розділ 3

Стаціонарні та звідні до них часові ряди

В попередніх розділах ми зовсім не цікавились виглядом та властивостями випадкового складника часових рядів. Тут ми розглянемо моделі часових рядів, основною ознакою яких якраз і буде наявність випадкового складника з різними властивостями.

3.1 Білий шум

Для будь-якого часового ряду можна створити деяку модель, яка б відображала розвиток процесу. Але яка б не була модель, неможливо прогнозувати значення часового ряду без похибки. Звичайно, розроблені методи, які намагаються зменшити цю похибку, але ніколи не можна бути впевненим у абсолютній точності прогнозів. Іншими словами, ніколи не було, нема і не буде моделі, яка б могла робити прогноз абсолютно точно. Таким чином, будь-яка модель повинна мати як доданок похибку, за величиною якої вимірюється точність моделі. Цією похибкою якраз і є випадковий складник. Для стаціонарних часових рядів випадковий складник будується з допомогою такого часового ряду, який прийнято називати білим шумом. Білим шумом будемо називати послідовність $\{\varepsilon_t\}$ ($t \in (-\infty, +\infty)$) незалежних (принаймні некорельованих) однаково розподілених випадкових величин, що мають моменти до другого порядку включно.

Будемо вважати, що $\mathbf{E}\varepsilon_t = 0$ і

$$\gamma_l = \text{cov}(\varepsilon_t, \varepsilon_{t-l}) = \mathbf{E}\varepsilon_t \varepsilon_{t-l} = \begin{cases} 0, & \text{якщо } l \neq 0; \\ \sigma^2, & \text{якщо } l = 0. \end{cases}$$

Процеси білого шуму є цеглинами для побудови більш складних моделей часових рядів. Розглянемо декілька з них.

3.2 МА(q)-процес

Процес $\{y_t\}$ рухомого середнього порядку q ($MA(q)$), будується за допомогою білого шуму ε_t за рівнянням :

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

де $\mu, \theta_1, \dots, \theta_q$ – деякі сталі. З формули видно, що цей часовий ряд містить тільки випадковий складник (тренд сталий). Використовуючи лаговий оператор такий процес можна записати у вигляді

$$y_t = \mu + \varepsilon_t + \theta_1 B \varepsilon_t + \dots + \theta_q B^q \varepsilon_t = \mu + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t.$$

На практиці трапляється дуже рідко, щоб реальний процес розвивався за $MA(q)$ -процесом, але дуже часто за його допомогою можна виразити доволі складні процеси, або отримати їх оцінки чи характеристики.

Розглянемо основні характеристики цього процесу.

1. Математичне сподівання:

$$\begin{aligned} \mathbf{E}y_t &= \mathbf{E}\mu + \mathbf{E}\varepsilon_t + \mathbf{E}(\theta_1 \varepsilon_{t-1}) + \dots + \mathbf{E}(\theta_q \varepsilon_{t-q}) = \\ &= \mu + \mathbf{E}\varepsilon_t + \theta_1 \mathbf{E}\varepsilon_{t-1} + \dots + \theta_q \mathbf{E}\varepsilon_{t-q} = \mu \end{aligned}$$

2. Дисперсія:

$$\begin{aligned} \gamma_0 &= \mathbf{D}y_t = \mathbf{D}\varepsilon_t + \theta_1^2 \mathbf{D}\varepsilon_{t-1} + \dots + \theta_q^2 \mathbf{D}\varepsilon_{t-q} = \\ &= \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2) \end{aligned}$$

3. Коваріація l -го порядку ($l = 1, 2, 3, \dots$):

$$\begin{aligned}
 \gamma_l &= \mathbf{E}(y_t - \mu)(y_{t-l} - \mu) = \\
 &= \mathbf{E}(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}) \\
 &\quad (\varepsilon_{t-l} + \theta_1 \varepsilon_{t-l-1} + \dots + \theta_q \varepsilon_{t-l-q}) = \\
 &= \begin{cases} \theta_l \mathbf{E}\varepsilon_{t-l}^2 + \theta_{l+1} \theta_1 \mathbf{E}\varepsilon_{t-l-1}^2 + \dots + \theta_q \theta_{q-l} \mathbf{E}\varepsilon_{t-q}^2, & l = 1, 2, \dots, q; \\ 0, & l > q. \end{cases} \\
 &= \begin{cases} (\theta_l + \theta_{l+1} \theta_1 + \dots + \theta_q \theta_{q-l}) \sigma^2, & l = 1, 2, \dots, q; \\ 0, & l > q. \end{cases}
 \end{aligned}$$

Звідси випливає, що якщо q є скінченним числом, то $MA(q)$ -процес є стаціонарним.

Розглянемо деякі приклади.

1. $MA(1)$ -процес: $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$.

- Математичне сподівання: $\mathbf{E}y_t = \mu$;
- дисперсія: $\gamma_0 = \sigma^2(1 + \theta_1^2)$;
- коваріація першого порядку: $\gamma_1 = \theta_1 \sigma^2$;
- коваріації інших порядків: $\gamma_2 = \gamma_3 = \dots = 0$;
- коефіцієнт автокореляції першого порядку:

$$\rho_1 = \frac{\theta_1 \sigma^2}{\sigma^2(1 + \theta_1^2)} = \frac{\theta_1}{1 + \theta_1^2};$$

- коефіцієнти автокореляції інших порядків:

$$\rho_2 = \rho_3 = \dots = 0.$$

2. $MA(2)$ -процес: $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$.

- Математичне сподівання: $\mathbf{E}y_t = \mu$;
- дисперсія: $\gamma_0 = \sigma^2(1 + \theta_1^2 + \theta_2^2)$;
- коваріація першого порядку: $\gamma_1 = \theta_1(1 + \theta_2)\sigma^2$;
- коваріація другого порядку: $\gamma_2 = \theta_2 \sigma^2$;
- коваріації інших порядків: $\gamma_3 = \gamma_4 = \dots = 0$.

3. $MA(\infty)$ -процес: $y_t = \mu + \sum_{j=0}^{+\infty} \theta_j \varepsilon_{t-j}$, тут $\theta_0 = 1$.

- Математичне сподівання: $\mathbf{E}y_t = \mu$;
- дисперсія: $\gamma_0 = \sigma^2 \sum_{j=0}^{+\infty} \theta_j^2$;
- коваріація l -го порядку ($l = 1, 2, 3, \dots$): $\gamma_l = \sigma^2 \sum_{j=0}^{+\infty} \theta_{l+j} \theta_j$.

Очевидно, що для стаціонарності $MA(\infty)$ -процесу потрібно, щоб $\sum_{l=0}^{+\infty} \theta_l^2 < +\infty$.

3.3 Процес авторегресії

Розвиток процесів авторегресії пов'язаний з необхідністю аналізу моделей, у яких одна змінна залежить від своїх попередніх значень. Побудова звичайних лінійних регресій не могла дати повне уявлення про напрями розвитку таких часових рядів. Внаслідок цього були запроваджені процеси авторегресії.

Нехай ε_t – білий шум, p – ціле невід'ємне число. Процесом авторегресії порядку p будемо називати процес виду:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

де $\phi_0, \phi_1, \dots, \phi_p$ – деякі сталі. Для побудови ряду за такою моделлю потрібно вказати p послідовних початкових значень y_t .

Використання лагового оператора дає можливість записати

$$\begin{aligned} y_t &= \phi_0 + \phi_1 B y_t + \phi_2 B^2 y_t + \dots + \phi_p B^p y_t + \varepsilon_t = \\ &= \phi_0 + \Phi(B) y_t + \varepsilon_t, \end{aligned}$$

де $\Phi(B) = \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p$. А звідси маємо

$$(1 - \Phi(B)) y_t = \phi_0 + \varepsilon_t.$$

AR(1)-процес. Цей процес можна записати у вигляді

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t.$$

Оскільки $y_{t-1} = \phi_0 + \phi_1 y_{t-2} + \varepsilon_{t-1}$, то

$$y_t = \phi_0(1 + \phi_1) + \phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t.$$

Повторивши попередню процедуру підстановки k разів, отримаємо

$$\begin{aligned} y_t &= \phi_0(1 + \phi_1 + \dots + \phi_1^k) + \phi_1^{k+1} y_{t-k-1} + \phi_1^k \varepsilon_{t-k} + \dots + \phi_1 \varepsilon_{t-1} + \varepsilon_t = \\ &= \phi_0 \sum_{j=0}^k \phi_1^j + \phi_1^{k+1} B^{k+1} y_t + \sum_{j=0}^{k+1} \phi_1^j B^j \varepsilon_t. \end{aligned}$$

Зрозуміло, що випадок $|\phi_1| > 1$ є нереальним, оскільки при цьому на значення часового ряду в даний період найбільший вплив будуть мати події, що відбувалися як завгодно давно. Правильно було б вважати, що чим давніше відбувалася подія, тим менший її вплив на теперішній момент.

Окремої розмови потребує випадок $|\phi_1| = 1$. Зараз ми будемо вважати, що $|\phi_1| < 1$. Це припущення є реалістичним, оскільки найбільша вага при визначенні значень часового ряду надається його останнім елементам.

Вважаючи, що t змінюється від $-\infty$ до $+\infty$, та спрямувавши k до $+\infty$, одержимо

$$\begin{aligned} y_t &= \phi_0 \sum_{j=0}^{+\infty} \phi_1^j + \sum_{j=0}^{+\infty} \phi_1^j B^j \varepsilon_t = \\ &= \frac{\phi_0}{1 - \phi_1} + \sum_{j=0}^{+\infty} \phi_1^j B^j \varepsilon_t. \end{aligned} \quad (3.1)$$

Тобто $AR(1)$ -процес подано у вигляді $MA(\infty)$ -процесу.

Зображення процесу $AR(1)$ у вигляді (3.1) дозволяє легко обчислити його числові характеристики.

- Математичне сподівання: $\mathbf{E}y_t = \frac{\phi_0}{1 - \phi_1}$;
- дисперсія: $\gamma_0 = \sigma^2 \sum_{j=0}^{+\infty} \phi_1^{2j} = \frac{\sigma^2}{1 - \phi_1^2}$;

- коваріація l -того порядку ($l = 1, 2, 3, \dots$):

$$\gamma_l = \sigma^2 \sum_{j=0}^{+\infty} \phi_1^{l+j} \phi_1^j = \sigma^2 \sum_{j=0}^{+\infty} \phi_1^{l+2j} = \frac{\sigma^2 \phi_1^l}{1 - \phi_1^2};$$

- коефіцієнт автокореляції l -того порядку ($l = 1, 2, 3, \dots$):

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \phi_1^l.$$

Отже, для AR(1)-процесу з $|\phi_1| < 1$ можемо зробити такі висновки:

- процес стаціонарний;
- вплив його значень одне на одне зменшується із збільшенням відстані між ними в часі.

Навпаки, припустивши стаціонарність AR(1)-процесу, легко одержимо знайдені вище його числові характеристики.

AR(2)-процес. Після розгляду властивостей процесу авторегресії першого порядку є зрозумілим перехід до більш складного двопорядкового.

Загальним виглядом AR(2)-процесу є $y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$. Перепишемо цю формулу у вигляді $(1 - \phi_1 B - \phi_2 B^2)y_t = \phi_0 + \varepsilon_t$. Ліву частину можна розкласти на множники: $(1 - \lambda_1 B)(1 - \lambda_2 B)y_t = \phi_0 + \varepsilon_t$, де λ_1 і λ_2 – сталі, що задовольняють систему рівнянь

$$\begin{cases} \lambda_1 + \lambda_2 = \phi_1; \\ \lambda_1 \lambda_2 = -\phi_2. \end{cases}$$

Отже, λ_1 і λ_2 є коренями квадратного рівняння $\lambda^2 - \phi_1 \lambda - \phi_2 = 0$. Звідси

$$\lambda_1 = \frac{1}{2} \left(\phi_1 - \sqrt{\phi_1^2 + 4\phi_2} \right), \quad \lambda_2 = \frac{1}{2} \left(\phi_1 + \sqrt{\phi_1^2 + 4\phi_2} \right).$$

Розглянемо випадок, коли $\phi_1^2 + 4\phi_2 > 0$ і, тому λ_1 і λ_2 різні дійсні числа. Таким чином, якщо $|\lambda_1| < 1$ і $|\lambda_2| < 1$, то

$$y_t = \frac{\phi_0 + \varepsilon_t}{(1 - \lambda_1 B)(1 - \lambda_2 B)} = \frac{1}{\lambda_1 - \lambda_2} \left(\frac{\lambda_1}{1 - \lambda_1 B} - \frac{\lambda_2}{1 - \lambda_2 B} \right) (\phi_0 + \varepsilon_t) =$$

$$\begin{aligned}
&= \frac{1}{\lambda_1 - \lambda_2} B^{-1} \sum_{j=1}^{+\infty} \lambda_1^j B^j \phi_0 - \frac{1}{\lambda_1 - \lambda_2} B^{-1} \sum_{j=1}^{+\infty} \lambda_2^j B^j \phi_0 + \\
&+ \frac{1}{\lambda_1 - \lambda_2} B^{-1} \sum_{j=1}^{+\infty} \lambda_1^j B^j \varepsilon_t - \frac{1}{\lambda_1 - \lambda_2} B^{-1} \sum_{j=1}^{+\infty} \lambda_2^j B^j \varepsilon_t = \\
&= \frac{\phi_0}{\lambda_1 - \lambda_2} \sum_{j=1}^{+\infty} \lambda_1^j - \frac{\phi_0}{\lambda_1 - \lambda_2} \sum_{j=1}^{+\infty} \lambda_2^j + \\
&+ \frac{\lambda_1}{\lambda_1 - \lambda_2} \sum_{j=0}^{+\infty} \lambda_1^j \varepsilon_{t-j} - \frac{\lambda_2}{\lambda_1 - \lambda_2} \sum_{j=0}^{+\infty} \lambda_2^j \varepsilon_{t-j} = \\
&= \frac{\phi_0}{1 - \lambda_1 - \lambda_2 + \lambda_1 \lambda_2} + \frac{1}{\lambda_1 - \lambda_2} \sum_{j=0}^{+\infty} (\lambda_1^{j+1} - \lambda_2^{j+1}) \varepsilon_{t-j}
\end{aligned}$$

Ми одержали $MA(\infty)$ -процес. Отже, якщо корені рівняння $1 - \phi_1 z - \phi_2 z^2 = 0$ різні дійсні і за абсолютною величиною більші від одиниці (вони є оберненими до λ_1 і λ_2), то $AR(2)$ -процес є $MA(\infty)$ -процесом. Такий процес буде стаціонарним, якщо

$$\sum_{j=0}^{+\infty} (\lambda_1^{j+1} - \lambda_2^{j+1})^2 < +\infty.$$

Очевидно, це матиме місце, якщо $|\lambda_1| < 1$ і $|\lambda_2| < 1$.

Знайдемо тепер числові характеристики розгляданого процесу за умови стаціонарності.

Математичне сподівання:

$$\begin{aligned}
\mu = \mathbf{E}y_t &= \phi_0 + \phi_1 \mathbf{E}y_{t-1} + \phi_2 \mathbf{E}y_{t-2} + \mathbf{E}\varepsilon_t = \phi_0 + \phi_1 \mu + \phi_2 \mu \Rightarrow \\
\mu &= \frac{\phi_0}{1 - \phi_1 - \phi_2}.
\end{aligned}$$

Для знаходження дисперсії та автоковаріацій перепишемо наш ряд у формі $y_t = \mu(1 - \phi_1 - \phi_2) + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$, звідки $y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \varepsilon_t$. Автоковаріація l -того порядку тоді обчислиться так:

$$\begin{aligned}
\gamma_l &= \mathbf{E}(y_t - \mu)(y_{t-l} - \mu) = \mathbf{E}\phi_1(y_{t-1} - \mu)(y_{t-l} - \mu) + \\
&+ \mathbf{E}\phi_2(y_{t-2} - \mu)(y_{t-l} - \mu) + \mathbf{E}\varepsilon_t(y_{t-l} - \mu) = \\
&= \phi_1 \gamma_{l-1} + \phi_2 \gamma_{l-2}
\end{aligned}$$

Таким чином ми отримали рекурентну формулу для обчислення автоковаріації:

$$\gamma_t = \phi_1 \gamma_{t-1} + \phi_2 \gamma_{t-2}. \quad (3.2)$$

Звідси, оскільки $\gamma_{-1} = \gamma_1$, то $\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1$ і

$$\gamma_1 = \frac{\phi_1}{1 - \phi_2} \gamma_0, \quad \gamma_2 = \frac{\phi_1^2}{1 - \phi_2} \gamma_0 + \phi_2 \gamma_0 = \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1 - \phi_2} \gamma_0$$

Дисперсія знаходиться з формули

$$\begin{aligned} \gamma_0 &= \mathbf{E}(y_t - \mu)^2 = \mathbf{E}(\phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \varepsilon_t)^2 = \\ &= \mathbf{E}\phi_1^2(y_{t-1} - \mu)^2 + \mathbf{E}\phi_2^2(y_{t-2} - \mu)^2 + \mathbf{E}\varepsilon_t^2 + \\ &+ 2\mathbf{E}\phi_1\phi_2(y_{t-1} - \mu)(y_{t-2} - \mu) + 2\mathbf{E}\phi_1(y_{t-1} - \mu)\varepsilon_t + \\ &+ 2\mathbf{E}\phi_2(y_{t-1} - \mu)\varepsilon_t = \\ &= \phi_1^2\gamma_0 + \phi_2^2\gamma_0 + \sigma^2 + 2\phi_1\phi_2\gamma_1 = \phi_1^2\gamma_0 + \phi_2^2\gamma_0 + \sigma^2 + 2\frac{\phi_1^2\phi_2}{1 - \phi_2} \gamma_0. \end{aligned}$$

Отже,

$$\gamma_0 = \frac{1 - \phi_2}{1 - \phi_2 - \phi_1^2 - \phi_2^2 + \phi_2^3 - \phi_1^2\phi_2} \sigma^2.$$

І тому

$$\begin{aligned} \gamma_1 &= \frac{\phi_1}{1 - \phi_2 - \phi_1^2 - \phi_2^2 + \phi_2^3 - \phi_1^2\phi_2} \sigma^2; \\ \gamma_2 &= \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1 - \phi_2 - \phi_1^2 - \phi_2^2 + \phi_2^3 - \phi_1^2\phi_2} \sigma^2. \end{aligned}$$

Автоковаріації вищих порядків знаходяться за наведеною вище рекурентною формулою (3.2).

AR(p)-процес. У цьому пункті ми закінчимо теоретичний аналіз процесів авторегресії, зробимо висновки та узагальнення розглянутих вище моделей. Загальним записом процесу авторегресії порядку $p \in$

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t.$$

Аналогічно попередньому можемо стверджувати, що $AR(p)$ -процес є стаціонарним, коли всі корені рівняння $1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$ різні дійсні і менші за абсолютною величиною від одиниці.

Числові характеристики такого процесу знаходяться аналогічно випадку $AR(2)$. А саме математичне сподівання

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}.$$

Дисперсії і автоковаріації знаходимо із системи рівнянь

$$\begin{aligned} \gamma_0 &= \phi_1\gamma_1 + \phi_2\gamma_2 + \dots + \phi_p\gamma_p + \sigma^2, \\ \gamma_l &= \phi_1\gamma_{l-1} + \phi_2\gamma_{l-2} + \dots + \phi_p\gamma_{l-p}, \quad l = 1, 2, \dots \end{aligned}$$

Тут потрібно враховувати, що $\gamma_{-l} = \gamma_l$.

Аналогічно попередньому для $AR(p)$ -процесу існує єдине представлення у вигляді процесу типу $MA(\infty)$. Дійсно нехай

$$y_t = \mu + \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \dots = \mu + \Psi(B)\varepsilon_t, \quad (3.3)$$

де $\Psi(B) = 1 + \psi_1B + \psi_2B^2 + \dots$

З іншого боку

$$y_t = \phi_0 + \phi_1y_{t-1} + \phi_2y_{t-2} + \dots + \phi_py_{t-p} + \varepsilon_t = \phi_0 + \Phi(B)y_t + \varepsilon_t, \quad (3.4)$$

де $\Phi(B) = \phi_1B + \phi_2B^2 + \dots + \phi_pB^p$.

Тоді

$$\mu = \frac{\phi_0}{1 - \Phi(B)}, \quad \Psi(B) = \frac{1}{1 - \Phi(B)} \Rightarrow \Psi(B)(1 - \Phi(B)) = 1.$$

Тобто $(1 + \psi_1B + \psi_2B^2 + \dots)(1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p) = 1$.

$$\begin{aligned} \psi_1B + \psi_2B^2 + \dots - \phi_1B - \psi_1\phi_1B^2 - \psi_2\phi_1B^3 - \dots - \\ - \phi_2B^2 - \psi_1\phi_2B^3 - \dots - \phi_pB^p - \psi_1\phi_pB^{p+1} - \dots = 0 \end{aligned}$$

Звідси одержуємо систему рівнянь для визначення коефіцієнтів $MA(\infty)$ -процесу (3.3) через коефіцієнти $AR(p)$ -процесу (3.4):

$$\begin{cases} \psi_1 - \phi_1 = 0, \\ \psi_2 - \psi_1\phi_1 - \phi_2 = 0, \\ \psi_3 - \psi_2\phi_1 - \psi_1\phi_2 - \phi_3 = 0, \\ \dots \end{cases}$$

Можна рекурентно записати розв'язок такої системи для $l = 1, 2, \dots$

$$\psi_l = \phi_1\psi_{l-1} + \phi_2\psi_{l-2} + \dots + \phi_p\psi_{l-p},$$

де $\psi_0 = 1$ і $\psi_j = 0$ при $j < 0$.

3.4 Перетворення $MA(q)$ -процесів

У попередніх розділах було показано, що будь-який скінченний $AR(p)$ -процес можна виразити у вигляді нескінченного $MA(\infty)$ -процесу єдиним чином. У цьому розділі буде показано, що і, навпаки, кожен скінченний $MA(q)$ -процес може бути представлений у вигляді нескінченного $AR(\infty)$ -процесу.

Для початку розглянемо $MA(1)$ -процес.

Оскільки $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$, то $\varepsilon_t = y_t - \mu - \theta_1 \varepsilon_{t-1}$, а звідси маємо при $k = 1, 2, \dots$

$$\varepsilon_{t-k} = y_{t-k} - \mu - \theta_1 \varepsilon_{t-k-1}.$$

Тому

$$\begin{aligned} y_t &= \mu + \varepsilon_t + \theta_1(y_{t-1} - \mu - \theta_1 \varepsilon_{t-2}) = \\ &= \mu(1 - \theta_1) + \varepsilon_t + \theta_1 y_{t-1} - \theta_1^2(y_{t-2} - \mu - \theta_1 \varepsilon_{t-3}) = \\ &= \mu(1 - \theta_1 + \theta_1^2) + \varepsilon_t + \theta_1 y_{t-1} - \theta_1^2 y_{t-2} + \theta_1^3 \varepsilon_{t-3} = \dots \end{aligned}$$

Отже,

$$\begin{aligned} y_t &= \mu(1 - \theta_1 + \theta_1^2 - \dots + (-1)^n \theta_1^n + \dots) + \varepsilon_t + \\ &+ \theta_1 y_{t-1} - \theta_1^2 y_{t-2} + \dots + (-1)^{n+1} \theta_1^n y_{t-n} + \dots \end{aligned}$$

або

$$y_t = \mu \sum_{n=0}^{+\infty} (-1)^n \theta_1^n + \sum_{n=1}^{+\infty} (-1)^{n+1} \theta_1^n y_{t-n} + \varepsilon_t.$$

За умови, що $|\theta_1| < 1$, маємо представлення $MA(1)$ -процесу у вигляді процесу типу $AR(\infty)$.

В загальному, подавши $MA(q)$ -процес у вигляді

$$y_t = \mu + \Theta(B)\varepsilon_t,$$

де $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$, а процес $AR(\infty)$ у вигляді

$$y_t = \phi_0 + \Phi(B)y_t + \varepsilon_t,$$

де $\Phi(B) = \phi_1 B + \phi_2 B^2 + \dots$, знайдемо формули для визначення коефіцієнтів ϕ_k ($k = 0, 1, 2, \dots$) через коефіцієнти μ, θ_j ($j = 1, 2, \dots, q$). Легко бачити, що

$$\mu = \frac{\phi_0}{1 - \Phi(B)}, \quad \Theta(B) = \frac{1}{1 - \Phi(B)},$$

звідки $\phi_0 = \mu(1 - \Phi(B))$, $\Theta(B)(1 - \Phi(B)) = 1$ або

$$\begin{aligned}\phi_0 &= \mu(1 - \phi_1 - \phi_2 - \dots), \\ (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)(1 - \phi_1 B - \phi_2 B^2 - \dots) &= 1.\end{aligned}$$

Звідси

$$\begin{aligned}-\phi_1 B - \phi_2 B^2 - \dots &+ \\ +\theta_1 B - \theta_1 \phi_1 B^2 - \theta_1 \phi_2 B^3 - \dots &+ \\ +\theta_2 B^2 - \theta_2 \phi_1 B^3 - \theta_2 \phi_2 B^4 - \dots &+ \\ +\theta_q B^q - \theta_q \phi_1 B^{q+1} - \theta_q \phi_2 B^{q+2} - \dots &= 0\end{aligned}$$

Отже,

$$\begin{aligned}\phi_1 &= \theta_1, \\ \phi_2 &= -\theta_1 \phi_1 + \theta_2, \\ \dots &\dots \dots \dots \\ \phi_q &= -\theta_1 \phi_{q-1} - \theta_2 \phi_{q-2} \dots - \theta_{q-1} \phi_1 + \theta_q, \\ \phi_{q+1} &= -\theta_1 \phi_q - \theta_2 \phi_{q-1} \dots - \theta_q \phi_1, \\ \dots &\dots \dots \dots \\ \phi_n &= -\theta_1 \phi_{n-1} - \theta_2 \phi_{n-2} \dots - \theta_q \phi_{n-q}, \\ \dots &\dots \dots \dots \\ \phi_0 &= \mu(1 - \phi_1 - \phi_2 - \dots).\end{aligned}$$

Очевидно таке представлення можливе, коли ряд $\sum_{n=1}^{+\infty} \phi_n$ - збіжний.

Це має місце у випадку якщо всі корені рівняння

$$1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$$

задовольняють умову $|z_j| > 1$.

На завершення зауважимо, що якщо деякий $AR(\infty)$ -процес відповідає $MA(q)$ -процесу, то він є стаціонарним.

3.5 ARMA-процес

Після вивчення властивостей $AR(p)$ - та $MA(q)$ -процесів є природнім розробити сумісний процес. Таким процесом є $ARMA(p, q)$, який являє собою суму двох розглянутих вище процесів. Особливістю цього

процесу є те, що y_t залежить і від своїх попередніх значень і від збурень минулих періодів. Загальним видом $ARMA(p, q)$ -процесу є:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

або $(1 - \Phi(B))y_t = \phi_0 + \Theta(B)\varepsilon_t$, де $\Phi(B) = \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p$, $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$.

$ARMA(p, q)$ -процес є стаціонарним, коли всі корені z_j рівняння $1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$ задовольняють умову $|z_j| > 1$.

Так само, як процес $AR(p)$ можна подати у вигляді $MA(\infty)$ -процесу, а $MA(q)$ -процес – процесу $AR(\infty)$, так і $ARMA(p, q)$ -процес можна подати як $MA(\infty)$ - або $AR(\infty)$ -процес.

МА(∞)-представлення. Порівнюючи праві частини рівностей

$$y_t = \frac{\phi_0}{1 - \Phi(B)} + \frac{\Theta(B)}{1 - \Phi(B)} \varepsilon_t \text{ та}$$

$$y_t = \mu + \Psi(B)\varepsilon_t,$$

де $\Psi(B) = \psi_0 + \psi_1 B + \psi_2 B^2 + \dots$, одержимо $\Theta(B) = \Psi(B)(1 - \Phi(B))$ або

$$1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q = (\psi_0 + \psi_1 B + \psi_2 B^2 + \dots) \times$$

$$\times (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p).$$

Звідси,

$$1 = \psi_0,$$

$$\theta_1 = \psi_1 - \psi_0 \phi_1 \quad (p \geq 1),$$

або

$$\theta_1 = \psi_1 \quad (p = 0),$$

$$\theta_2 = \psi_2 - \psi_1 \phi_1 - \psi_0 \phi_2 \quad (p \geq 2),$$

або

$$\theta_2 = \psi_2 - \psi_1 \phi_1 \quad (p = 1),$$

або

$$\theta_2 = \psi_2 \quad (p = 0),$$

...

$$\theta_q = \psi_q - \psi_{q-1} \phi_1 - \psi_{q-2} \phi_2 - \dots - \psi_0 \phi_q, \quad (p \geq q),$$

або

$$\theta_q = \psi_q - \psi_{q-1}\phi_1 - \psi_{q-2}\phi_2 - \dots - \psi_{q-p}\phi_p, \quad (p < q),$$

$$0 = \psi_{q+1} - \psi_q\phi_1 - \psi_{q-1}\phi_2 - \dots - \psi_0\phi_{q+1}, \quad (p \geq q + 1),$$

або

$$0 = \psi_{q+1} - \psi_q\phi_1 - \psi_{q-1}\phi_2 - \dots - \psi_{q-p+1}\phi_p, \quad (p < q + 1),$$

... ..

а поклавши $\psi_j = 0$ при $j < 0$, матимемо (з врахуванням $\psi_j = 0$ при $j < 0$)

$$\psi_0 = 1,$$

$$\psi_k = \theta_k + \sum_{j=1}^p \psi_{k-j}\phi_j \quad (k = 1, 2, \dots, q),$$

$$\psi_k = \sum_{j=1}^p \psi_{k-j}\phi_j \quad (k > q).$$

Крім того $\mu = \frac{\phi_0}{1 - \Phi(B)}$ і $(1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu = \phi_0$ або

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}.$$

AR(∞)-представлення. Аналогічно попередньому з рівностей

$$\varepsilon_t = \frac{(1 - \Phi(B))}{\Theta(B)} y_t - \frac{\phi_0}{\Theta(B)} \text{ та}$$

$$\varepsilon_t = (1 - \Gamma(B))y_t - \gamma_0,$$

де $\Gamma(B) = \gamma_1 B + \gamma_2 B^2 + \dots$, маємо $\phi_0 = \Theta(B)\gamma_0$, звідки

$$\gamma_0 = \frac{\phi_0}{1 + \theta_1 + \theta_2 + \dots + \theta_q},$$

і $(1 - \Phi(B)) = \Theta(B)(1 - \Gamma(B))$ або

$$\begin{aligned} 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p &= (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \times \\ &\times (1 - \gamma_1 B - \gamma_2 B^2 - \dots). \end{aligned}$$

Прирівнюючи коефіцієнти при однакових степенях B , одержимо вирази для γ_j .

3.6 Прогнозування на основі ARMA-моделей

Будь-яка модель створюється для аналізу інформації та прогнозування. При цьому точність прогнозів повинна бути як найвища. Розглянемо процес прогнозування на основі ARMA-моделей. Будемо позначати прогнозне значення часового ряду на τ періодів вперед, яке робиться в момент t через $y_{t+\tau|t}$. Для створення прогнозного значення будемо використовувати всю наявну інформацію на період t . Наша мета – мінімізувати один з критеріїв помилок. Зазвичай використовується критерій MSE (mean squared error):

$$MSE = \mathbf{E}(y_{t+\tau} - \hat{y}_{t+\tau|t})^2.$$

3.6.1 Прогнозування на основі залишків

Припустимо, що нам відомі значення збурень за перші t періодів. Тоді, для нескінченного MA(∞)-процесу

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots$$

виконується рівність

$$y_{t+\tau} = \mu + \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t-1+\tau} + \theta_2 \varepsilon_{t-2+\tau} + \dots + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots$$

Прогнозом на період $t+\tau$ в момент t є $\hat{y}_{t+\tau|t} = \mathbf{E}_t y_{t+\tau}$, де \mathbf{E}_t – символ умовного математичного сподівання при умові, що відомі значення часового ряду до моменту t включно. Оскільки ми припускали, що математичне сподівання ε_t є нульовим, то

$$\hat{y}_{t+\tau|t} = \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots$$

Враховуючи, що всі збурення мають однакову дисперсію, підрахуємо середню похибку такого прогнозу:

$$\begin{aligned} y_{t+\tau} - \hat{y}_{t+\tau|t} &= \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t-1+\tau} + \theta_2 \varepsilon_{t-2+\tau} + \dots + \theta_{\tau-1} \varepsilon_{t+1} \\ MSE &= (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_{\tau-1}^2) \sigma^2 \end{aligned}$$

Для MA(q)-процесу:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

прогнозом на період $t + \tau$ у період t є:

$$\hat{y}_{t+\tau|t} = \begin{cases} \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q+\tau}, & \tau = \overline{1, q}, \\ \mu, & \tau > q. \end{cases}$$

Дійсно, в цьому випадку, при $t = 1, 2, \dots, q$

$$y_{t+\tau} = \mu + \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t-1+\tau} + \theta_2 \varepsilon_{t-2+\tau} + \dots + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q+\tau},$$

а при $\tau > q$

$$y_{t+\tau} = \mu + \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t-1+\tau} + \theta_2 \varepsilon_{t-2+\tau} + \dots + \theta_q \varepsilon_{t-q+\tau}.$$

Для похибки тоді маємо вирази

$$MSE = \begin{cases} (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_{\tau-1}^2) \sigma^2, & \tau = \overline{1, q}, \\ (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2, & \tau > q. \end{cases}$$

У найпростішому випадку $MA(1)$ -процесу прогнозом на всі періоди, крім $t + 1$ -го, завжди є середнє значення:

$$\hat{y}_{t+1|t} = \mu + \theta_1 \varepsilon_t, \quad \hat{y}_{t+\tau|t} = \mu \text{ при } \tau > 1.$$

При розгляді AR -процесів, слід пам'ятати, що його можна перетворити до нескінченного MA -процесу. Наприклад, для $AR(1)$ -процесу, що задається рівнянням $y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$ маємо (див. (3.1)):

$$y_t = \mu + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots,$$

де $\mu = \frac{\phi_0}{1 - \phi_1}$. Тому

$$\hat{y}_{t+\tau|t} = \frac{\phi_0}{1 - \phi_1} + \phi_1^\tau \varepsilon_t + \phi_1^{\tau+1} \varepsilon_{t-1} + \dots$$

Похибка становить $MSE = (1 + \phi_1^2 + \phi_1^4 + \dots + \phi_1^{2(\tau-1)}) \sigma^2$. Звідси, між іншим, впливає (згадаємо, що $|\phi_1| < 1$), що

$$MSE \rightarrow \frac{\sigma^2}{1 - \phi_1^2} \text{ при } \tau \rightarrow +\infty.$$

За допомогою вищеописаної методики можна отримувати прогнозні значення на будь-який період. Але будь-який прогноз рідко

може бути ідеально точним. Частіше за все дослідники вказують деякий надійний інтервал, куди повинне попасти реальне значення у відповідний період часу. За допомогою часових рядів також можна будувати інтервальні прогнози з різним рівнем надійності. Основною ідеєю побудови інтервалів є відхилення на однакову визначену величину у напрямі зростання та спадання від точного прогнозного значення. Основним припущенням для реалізації цієї ідеї є нормальність збурень: $\varepsilon_t \sim N(0, \sigma^2)$.

Нехай γ є рівнем надійності, тобто ймовірність попадання прогнозного значення в довірчий інтервал. Враховуючи, що $y_{t+\tau}$ має нормальний розподіл із середнім $\hat{y}_{t+\tau|t}$ та дисперсією MSE , можемо будувати симетричний відносно $\hat{y}_{t+\tau|t}$ надійний інтервал для $y_{t+\tau}$. А саме, з ймовірністю γ

$$y_{t+\tau} \in \left(\hat{y}_{t+\tau|t} - u_{\frac{\gamma+1}{2}} \sqrt{MSE}; \hat{y}_{t+\tau|t} + u_{\frac{\gamma+1}{2}} \sqrt{MSE} \right),$$

де $u_{\frac{\gamma+1}{2}}$ – квантиль порядку $\frac{\gamma+1}{2}$ стандартного нормального розподілу.

Таким чином, знаючи напевно збурення минулих періодів, дослідник може сформулювати прогноз та надійний інтервал на майбутні періоди. Хоча, як правило, ці збурення нам невідомі. Потрібно мати метод оцінки збурень в попередні періоди.

3.6.2 Формула Вінера-Колмогорова

Нехай задано деякий поліном (скінченний чи нескінченний) від лагового оператора B :

$$\Psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots$$

Для деякого $\tau = 1, 2, \dots$ розглянемо

$$\frac{\Psi(B)}{B^\tau} = B^{-\tau} + \psi_1 B^{1-\tau} + \psi_2 B^{2-\tau} + \psi_3 B^{3-\tau} + \dots$$

і позначимо $\left[\frac{\Psi(B)}{B^\tau} \right]_+$ = $\psi_\tau + \psi_{\tau+1} B + \psi_{\tau+2} B^2 + \dots$

Для процесу $MA(q)$:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} = \mu + \Theta(B) \varepsilon_t,$$

де $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ при $^1 \tau \leq q$

$$\hat{y}_{t+\tau|t} = \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q+\tau} = \mu + \left[\frac{\Theta(B)}{B^\tau} \right]_+ \varepsilon_t.$$

а оскільки $\varepsilon_t = \frac{1}{\Theta(B)}(y_t - \mu)$, то

$$\hat{y}_{t+\tau|t} = \mu + \left[\frac{\Theta(B)}{B^\tau} \right]_+ \frac{1}{\Theta(B)} (y_t - \mu). \quad (3.5)$$

Це і є формула Вінера-Колмогорова. Вона залишається вірною і для AR(p)-процесів.

Дійсно, нехай

$$y_t = \phi_0 + \Phi(B)y_t + \varepsilon_t,$$

де $\Phi(B) = \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p$. Подамо його у вигляді MA(∞)-процесу:

$$y_t = \mu + \Theta(B)\varepsilon_t, \quad (3.6)$$

де $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots$. Аналогічно попередньому

$$\hat{y}_{t+\tau|t} = \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots = \mu + \left[\frac{\Theta(B)}{B^\tau} \right]_+ \varepsilon_t.$$

Оскільки з (3.6) $\varepsilon_t = \frac{1}{\Theta(B)}(y_t - \mu)$, то має місце (3.5). Залишилось тільки зауважити, що $\Theta(B) = \frac{1}{1 - \Phi(B)}$.

Прогнозування AR-процесів. Розглянемо застосування цієї формули для випадку AR(1)-процесу: $(1 - \phi_1 B)y_t = \phi_0 + \varepsilon_t$. Тоді

$\Theta(B) = \frac{1}{1 - \phi_1 B}$, $\mu = \frac{\phi_0}{1 - \phi_1}$ (див. с. 33). Оскільки $\Theta(B) = 1 + \phi_1 B + \phi_1^2 B^2 + \phi_1^3 B^3 + \dots$, то

$$\begin{aligned} \left[\frac{\Theta(B)}{B^\tau} \right]_+ &= \phi_1^\tau + \phi_1^{\tau+1} B + \phi_1^{\tau+2} B^2 + \dots = \\ &= \phi_1^\tau (1 + \phi_1 B + \phi_1^2 B^2 + \dots) = \frac{\phi_1^\tau}{1 - \phi_1 B}. \end{aligned}$$

¹Випадок $\tau > q$ мало цікавий, бо $\hat{y}_{t+\tau|t} = \mu$.

Тому за (3.5)

$$\hat{y}_{t+\tau|t} = \frac{\phi_0}{1-\phi_1} + \phi_1^\tau \left(y_t - \frac{\phi_0}{1-\phi_1} \right) = \frac{\phi_0(1-\phi_1^\tau)}{1-\phi_1} + \phi_1^\tau y_t$$

Порівнюючи першу частину цієї рівності з

$$\hat{y}_{t+\tau+1|t} = \frac{\phi_0}{1-\phi_1} + \phi_1^{\tau+1} \left(y_t - \frac{\phi_0}{1-\phi_1} \right)$$

робимо висновок, що

$$\hat{y}_{t+\tau+1|t} = \frac{\phi_0}{1-\phi_1} + \phi_1 \left(\hat{y}_{t+\tau|t} - \frac{\phi_0}{1-\phi_1} \right)$$

Це є один із записів закону ітеративних сподівань. Очевидно, що $\hat{y}_{t|t} = y_t$.

Подібна ситуація має місце і в загальному випадку $AR(p)$ -процесу. Розглянемо часовий ряд

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

і запишемо його у вигляді

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \dots + \phi_p(y_{t-p} - \mu) + \varepsilon_t,$$

де $\mu = \frac{\phi_0}{1-\phi_1-\phi_2-\dots-\phi_p}$. Така рівність повинна зберігатись і для прогнозних значень, тобто

$$\hat{y}_{t+\tau|t} - \mu = \phi_1(\hat{y}_{t+\tau-1|t} - \mu) + \phi_2(\hat{y}_{t+\tau-2|t} - \mu) + \dots + \phi_p(\hat{y}_{t+\tau-p|t} - \mu) + \hat{\varepsilon}_{t+\tau|t}.$$

Оскільки $\hat{\varepsilon}_{t+\tau|t} = 0$, то

$$\hat{y}_{t+\tau|t} - \mu = \phi_1(\hat{y}_{t+\tau-1|t} - \mu) + \phi_2(\hat{y}_{t+\tau-2|t} - \mu) + \dots + \phi_p(\hat{y}_{t+\tau-p|t} - \mu).$$

Враховуючи, що $\hat{y}_{j|t} = y_j$ при $j = 1, 2, \dots, t$, одержимо рекурентні формули для прогнозів

$$\begin{aligned} \hat{y}_{t+1|t} - \mu &= \phi_1(y_t - \mu) + \phi_2(y_{t-1} - \mu) + \dots + \phi_p(y_{t-p+1} - \mu) \\ \hat{y}_{t+2|t} - \mu &= \phi_1(\hat{y}_{t+1|t} - \mu) + \phi_2(y_t - \mu) + \dots + \phi_p(y_{t-p+2} - \mu) \\ \hat{y}_{t+3|t} - \mu &= \phi_1(\hat{y}_{t+2|t} - \mu) + \phi_2(\hat{y}_{t+1|t} - \mu) + \dots + \phi_p(y_{t-p+3} - \mu) \\ &\dots \dots \dots \end{aligned}$$

$$\hat{y}_{t+p+1|t} - \mu = \phi_1(\hat{y}_{t+p|t} - \mu) + \phi_2(\hat{y}_{t+p-1|t} - \mu) + \dots + \phi_p(\hat{y}_{t+1|t} - \mu)$$

.....

де вважаємо, що $\hat{y}_{j|t} = \mu$ при $j \leq 0$ і $t \geq p$.

Прогнозування МА-процесів. Запропоновану методику можна використовувати і для підрахунку прогнозів для МА-процесів. Розглянемо її на прикладі МА(1)-процесу:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} = \mu + (1 + \theta_1 B) \varepsilon_t$$

Застосування формули Вінера-Колмогорова приводить до

$$\hat{y}_{t+\tau|t} = \mu + \left[\frac{1 + \theta_1 B}{B^\tau} \right]_+ \frac{1}{1 + \theta_1 B} (y_t - \mu),$$

що при $\tau \geq 2$ дає $\hat{y}_{t+\tau|t} = \mu$, а

$$\hat{y}_{t+1|t} = \mu + \frac{\theta_1}{1 + \theta_1 B} (y_t - \mu) = \mu + \theta_1 (y_t - \mu) - \theta_1^2 (y_{t-1} - \mu) + \dots$$

Слід зауважити, що для того, щоб досягти досить високого наближення, достатньо підрахувати лише декілька доданків. Якщо значення θ_1 є досить близьким до нуля, важливими є лише декілька перших доданків в одержаній сумі. Але якщо θ_1 є близьким до 1, то потрібно використовувати для розрахунків велику кількість доданків, що значно ускладнює розрахунки.

Тому був розроблений альтернативний метод для прогнозування МА-процесів. Для МА(1)-процесу запишемо $\varepsilon_t = y_t - \theta_1 \varepsilon_{t-1} - \mu$. Припустивши, що $\hat{\varepsilon}_0 = 0$ матимемо

$$\begin{aligned} \hat{\varepsilon}_1 &= y_1 - \theta_1 \hat{\varepsilon}_0 - \mu = y_1 - \mu, \\ \hat{\varepsilon}_2 &= y_2 - \theta_1 \hat{\varepsilon}_1 - \mu = y_2 - \mu - \theta_1 (y_1 - \mu), \\ \hat{\varepsilon}_3 &= y_3 - \theta_1 \hat{\varepsilon}_2 - \mu = y_3 - \mu - \theta_1 (y_2 - \mu) + \theta_1^2 (y_1 - \mu) \text{ і т.д.} \end{aligned}$$

Крім того очевидно, що $y_{t+1} = \mu + \varepsilon_{t+1} + \theta_1 \varepsilon_t$. Тому

$$\begin{aligned} \hat{y}_{t+1|t} &= \mu + \hat{\varepsilon}_{t+1|t} + \theta_1 \hat{\varepsilon}_{t|t} = \mu + \theta_1 \hat{\varepsilon}_{t|t} = \mu + \theta_1 \hat{\varepsilon}_t, \\ \hat{y}_{t+2|t} &= \mu + \hat{\varepsilon}_{t+2|t} + \theta_1 \hat{\varepsilon}_{t+1|t} = \mu, \\ &\dots \dots \dots \\ \hat{y}_{t+\tau|t} &= \mu \end{aligned}$$

У загальному випадку для МА(q)-процесу

$$y_t = \mu + (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t$$

застосування формули Вінера-Колмогорова дає

$$\hat{y}_{t+\tau|t} = \mu + \left[\frac{1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q}{B^\tau} \right]_+ \times \\ \times \frac{1}{1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q} (y_t - \mu).$$

Оскільки

$$\left[\frac{1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q}{B^\tau} \right]_+ = \\ = \begin{cases} \theta_\tau + \theta_{\tau+1} B + \theta_{\tau+2} B^2 + \dots + \theta_q B^{q-\tau}, & \tau = \overline{1, q}, \\ 0, & \tau > q, \end{cases} \\ \frac{1}{1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q} (y_t - \mu) = \varepsilon_t,$$

то

$$\hat{y}_{t+\tau|t} = \begin{cases} \mu + (\theta_\tau + \theta_{\tau+1} B + \theta_{\tau+2} B^2 + \dots + \theta_q B^{q-\tau}) \varepsilon_t, & \tau = \overline{1, q}, \\ \mu, & \tau > q \end{cases}$$

або

$$\hat{y}_{t+\tau|t} = \begin{cases} \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \theta_{\tau+2} \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-(q-\tau)}, & \tau = \overline{1, q}, \\ \mu, & \tau > q. \end{cases}$$

Крім того, в момент часу t середні значення оцінок збурень попередніх періодів збігаються із самими збуреннями та їх оцінками, тобто, $\hat{\varepsilon}_j|t = \varepsilon_j = \hat{\varepsilon}_j$ при $j = 1, 2, \dots, t$. Отже,

$$\hat{y}_{t+\tau|t} = \begin{cases} \mu + \theta_\tau \hat{\varepsilon}_t + \theta_{\tau+1} \hat{\varepsilon}_{t-1} + \dots + \theta_q \hat{\varepsilon}_{t-(q-\tau)}, & \tau = \overline{1, q}, \\ \mu, & \tau > q. \end{cases} = \\ = \begin{cases} \mu + \theta_\tau \hat{\varepsilon}_t + \theta_{\tau+1} \hat{\varepsilon}_{t-1} + \dots + \theta_q \hat{\varepsilon}_{t-(q-\tau)}, & \tau = \overline{1, q}, \\ \mu, & \tau > q. \end{cases}$$

Оцінки збурень у кожен період часу підраховуються за рекурентними формулами:

$$\begin{aligned} \hat{\varepsilon}_1 &= y_1 - \mu, \\ \hat{\varepsilon}_2 &= y_2 - \mu - \theta_1 \hat{\varepsilon}_1, \\ \hat{\varepsilon}_3 &= y_3 - \mu - \theta_1 \hat{\varepsilon}_2 - \theta_2 \hat{\varepsilon}_1, \\ \dots &\dots\dots \\ \hat{\varepsilon}_t &= y_t - \mu - \theta_1 \hat{\varepsilon}_{t-1} - \theta_2 \hat{\varepsilon}_{t-2} - \dots - \theta_q \hat{\varepsilon}_{t-q} \end{aligned}$$

Тут враховано, що $\hat{\varepsilon}_j = 0$ при $j \leq 0$.

Аналогічно у випадку $MA(\infty)$ -процесу одержуємо для всіх $\tau > 0$

$$\hat{y}_{t+\tau|t} = \mu + \theta_\tau \hat{\varepsilon}_t + \theta_{\tau+1} \hat{\varepsilon}_{t-1} + \dots + \theta_{\tau+t-1} \hat{\varepsilon}_1.$$

Прогнозування ARMA-процесів. Розглянемо процес $ARMA(p, q)$

$$(1 - \phi_1 B - \dots - \phi_p B^p)(y_t - \mu) = (1 + \theta_1 B + \dots + \theta_q B^q)\varepsilon_t$$

Тут на відміну від попередніх записів ми зробили заміну

$$\phi_0 = \mu(1 - \phi_1 - \dots - \phi_p).$$

Для періоду $t + \tau$ маємо

$$\begin{aligned} y_{t+\tau} &= \mu + \phi_1(y_{t+\tau-1} - \mu) + \phi_2(y_{t+\tau-2} - \mu) + \dots + \phi_p(y_{t+\tau-p} - \mu) + \\ &+ \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t+\tau-1} + \theta_2 \varepsilon_{t+\tau-2} + \dots + \theta_q \varepsilon_{t+\tau-q}. \end{aligned}$$

Закон ітеративних сподівань дозволяє записати

$$\begin{aligned} \hat{y}_{t+\tau|t} &= \mu + \\ &+ \phi_1(\hat{y}_{t+\tau-1|t} - \mu) + \phi_2(\hat{y}_{t+\tau-2|t} - \mu) + \dots + \phi_p(\hat{y}_{t+\tau-p|t} - \mu) + \\ &+ \hat{\varepsilon}_{t+\tau|t} + \theta_1 \hat{\varepsilon}_{t+\tau-1|t} + \theta_2 \hat{\varepsilon}_{t+\tau-2|t} + \dots + \theta_q \hat{\varepsilon}_{t+\tau-q|t}. \end{aligned}$$

Далі враховуємо, що

$$\hat{y}_{t-j|t} = \begin{cases} y_{t-j}, & \text{при } j = 0, 1, \dots, t-1, \\ 0, & \text{при } j \geq t \end{cases}$$

та при $j \geq 0$

$$\begin{aligned} \hat{\varepsilon}_{t+j+1|t} &= 0, \\ \hat{\varepsilon}_{t-j|t} &= \hat{\varepsilon}_{t-j} = \\ &= y_{t-j} - \mu - \phi_1(y_{t-j-1} - \mu) - \dots - \phi_p(y_{t-j-p} - \mu) - \\ &- \theta_1 \hat{\varepsilon}_{t-j-1} - \theta_2 \hat{\varepsilon}_{t-j-2} - \dots - \theta_q \hat{\varepsilon}_{t-j-q}. \end{aligned}$$

За допомогою перерахованих методів здійснюється прогнозування на основі ARMA-моделей. Дані часових рядів $\{Y_T\}$ є відомими, збурення оцінюються за вищенаведеною методикою і, маючи значення коефіцієнтів відповідних моделей, можна отримати як точкове значення прогнозу, так і відповідний надійний інтервал із заданим рівнем надійності. Єдиною проблемою при аналізі часових рядів залишається оцінка параметрів ϕ_i , $i = 0, 1, \dots, p$ та θ_j , $j = 1, 2, \dots, q$. Вирішенню цієї проблема присвячений наступний пункт.

3.7 Оцінювання невідомих параметрів

Для знаходження невідомих параметрів $ARMA$ -моделей можна використати метод найменших квадратів. Він застосовується при оцінці AR -моделей виду

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t.$$

Суть методу полягає у мінімізації суми квадратів збурень за всі періоди:

$$S = \sum_{t=p+1}^T \varepsilon_t^2 = \sum_{t=p+1}^T (y_t - \phi_0 - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p})^2$$

як функції від $\phi_0, \phi_1, \dots, \phi_p$.

Позначивши

$$y = \begin{pmatrix} y_{p+1} \\ y_{p+2} \\ \dots \\ y_T \end{pmatrix}, \quad X = \begin{pmatrix} 1 & y_p & y_{p-1} & \dots & y_1 \\ 1 & y_{p+1} & y_p & \dots & y_2 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & y_{T-1} & y_{T-2} & \dots & y_{T-p} \end{pmatrix},$$

$$\varepsilon = \begin{pmatrix} \varepsilon_{p+1} \\ \varepsilon_{p+2} \\ \dots \\ \varepsilon_T \end{pmatrix}, \quad \phi = \begin{pmatrix} \phi_0 \\ \phi_1 \\ \dots \\ \phi_p \end{pmatrix},$$

матимемо рівняння множинної регресії $y = X\phi + \varepsilon$, звідки отримуємо оцінки коефіцієнтів $\hat{\phi} = (X'X)^{-1}X'y$.

Описану методику не можна прямо застосовувати для MA -процесів, тому розроблений модифікований алгоритм знаходження оцінок для довільного $MA(q)$ -процесу. Для прикладу розглянемо $MA(1)$ -процес $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$. Нашою метою є мінімізація функції

$$S = \sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T (y_t - \mu - \theta_1 \varepsilon_{t-1})^2.$$

Оскільки нам невідомі попередні збурення, то ми використовуємо їх рекурентні оцінки:

$$\hat{\varepsilon}_0 = 0,$$

$$\begin{aligned}
\hat{\varepsilon}_1 &= y_1 - \mu \\
\hat{\varepsilon}_2 &= y_2 - \mu - \theta_1 \hat{\varepsilon}_1 = y_2 - \mu - \theta_1 (y_1 - \mu) \\
\dots \quad \dots \quad \dots \\
\hat{\varepsilon}_k &= y_k - \mu - \theta_1 (y_{k-1} - \mu) + \theta_1^2 (y_{k-2} - \mu) - \dots + \\
&\quad + (-1)^{k-1} \theta_1^{k-1} (y_1 - \mu)
\end{aligned}$$

Отже, потрібно мінімізувати функцію $S = \sum_{t=1}^T \hat{\varepsilon}_t^2$. Мінімізація такої функції здійснюється за допомогою градієнтних методів, з яких можна окремо виділити нелінійний метод найменших квадратів та метод Гауса.

3.8 ARIMA-процеси

В багатьох прикладних задачах з даних відразу стає зрозуміло, що припущення про фіксований рівень середнього часового ряду хибне. Однією з найбільш поширених моделей часового ряду для опису такого типу нестационарності є модель з одиничними коренем. Така модель виникає, якщо розглянути $ARMA(p, q)$ процес

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

і дозволити многочлену $1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$ мати одиничний корінь. Найпростішим прикладом такого процесу є випадкове блукання $y_t = y_{t-1} + \varepsilon_t$.

Якщо порядок одиничного кореня згаданого многочлена d , то маємо $ARIMA(p, d, q)$ -процес. Моделі такого типу, крім змінного середнього, мають багато інших цікавих властивостей, наприклад – довгу пам'ять (з часом залежність між спостереженнями зменшується повільніше ніж у випадку стаціонарних моделей).

$ARIMA(p, d, q)$ -процеси є загальною моделлю прогнозування нестационарних часових рядів, які можуть бути приведені до стаціонарного виду за допомогою логарифмування або різниць. Запис процесу має три параметри: p – кількість лагових змінних, q – кількість лагових збурень, d – кількість різниць. Під різницею розуміється трансформація часового ряду за правилом

$$\tilde{y}_t = y_t - y_{t-1} = (1 - B)y_t, \quad t = \overline{2, T}.$$

Іноді, значення початкового ряду спочатку логарифмують, тоді

$$\tilde{y}_t = \ln \frac{y_t}{y_{t-1}} = \ln y_t - \ln y_{t-1}, \quad t = \overline{2, T}.$$

В цьому випадку нові значення \tilde{y}_t відповідають рівню зростання процесу.

Процес називається інтегрованим порядку d , коли $(1 - B)^d y_t$, є стаціонарним та перетвореним процесом. Отриманий після цього стаціонарний процес відрізняється від початкового кількістю спостережень.

Отже, після d -кратного застосування до $ARIMA(p, d, q)$ -процесу різницевого оператора $\Delta = 1 - B$ одержуємо $ARMA(p, q)$ -процес.

Припустимо, що часовий ряд $\{Y_T\}$ є таким, що узгоджується з моделлю $ARIMA(p, d, q)$ -процесу. Тобто ряд $(1 - B)^d y_t$, $t = \overline{d + 1, T}$ має структуру $ARMA(p, q)$ -процесу. Отже,

$$(1 - B)^d y_t = \phi_0 + \Phi(B)(1 - B)^d y_t + \Theta(B)\varepsilon_t,$$

де $\Phi(B) = \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p$, $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$. Звідси

$$(1 - B)^d (1 - \Phi(B)) y_t = \phi_0 + \Theta(B)\varepsilon_t$$

і многочлен $(1 - B)^d (1 - \Phi(B))$ має одиничний корінь порядку d . В такому випадку кажуть, що процес має одиничний корінь. Дуже важливим є визначення того, чи є одиничний корінь в процесу. Тоді за допомогою застосування різниць його можна звести до стаціонарного.

Прогнозування $ARIMA$ -процесів. Розглянемо $ARIMA(p, 1, q)$ -процес. Нам відомо, що в цьому випадку ряд $(1 - B)y_t$ можна записати у вигляді $MA(\infty)$ -процесу:

$$z_t = (1 - B)y_t = \mu + \Theta(B)\varepsilon_t, \quad \text{де } \Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots$$

Знайдемо прогнозні значення (див. с. 49)

$$\hat{z}_{t+\tau|t} = \mu + \theta_\tau \hat{\varepsilon}_t + \theta_{\tau+1} \hat{\varepsilon}_{t-1} + \dots + \theta_{\tau+t-1} \hat{\varepsilon}_1.$$

Тут як і раніше вважаємо, що $\hat{\varepsilon}_j = 0$ при $j \leq 0$.

Для отримання прогнозу $\hat{y}_{t+\tau|t}$ необхідно зазначити, що

$$\begin{aligned} y_{t+\tau} &= y_t + (y_{t+1} - y_t) + (y_{t+2} - y_{t+1}) + \dots + (y_{t+\tau} - y_{t+\tau-1}) = \\ &= y_t + z_{t+1} + z_{t+2} + \dots + z_{t+\tau}, \end{aligned}$$

звідки $\hat{y}_{t+\tau|t} = y_t + \hat{z}_{t+1|t} + \hat{z}_{t+2|t} + \dots + \hat{z}_{t+\tau|t}$. Отже,

$$\begin{aligned} \hat{y}_{t+\tau|t} &= y_t + \tau\mu + (\theta_1\hat{\varepsilon}_t + \theta_2\hat{\varepsilon}_{t-1} + \dots + \theta_t\hat{\varepsilon}_1) + \\ &+ (\theta_2\hat{\varepsilon}_t + \theta_3\hat{\varepsilon}_{t-1} + \dots + \theta_{t+1}\hat{\varepsilon}_1) + \dots + (\theta_\tau\hat{\varepsilon}_t + \theta_{\tau+1}\hat{\varepsilon}_{t-1} + \dots + \theta_{\tau+t-1}\hat{\varepsilon}_1) = \\ &= y_t + \tau\mu + (\theta_1 + \theta_2 + \dots + \theta_\tau)\hat{\varepsilon}_t + (\theta_2 + \theta_3 + \dots + \theta_{\tau+1})\hat{\varepsilon}_{t-1} + \dots + \\ &+ (\theta_t + \theta_{t+1} + \dots + \theta_{\tau+t-1})\hat{\varepsilon}_1. \end{aligned}$$

Оцінки $\hat{\varepsilon}_j$ знаходимо за рекурентними формулами (див. с. 48)

$$\begin{aligned} \hat{\varepsilon}_1 &= z_1 - \mu, \\ \hat{\varepsilon}_2 &= z_2 - \mu - \theta_1\hat{\varepsilon}_1, \\ \hat{\varepsilon}_3 &= z_3 - \mu - \theta_1\hat{\varepsilon}_2 - \theta_2\hat{\varepsilon}_1, \\ \dots &\dots\dots \\ \hat{\varepsilon}_t &= z_t - \mu - \theta_1\hat{\varepsilon}_{t-1} - \theta_2\hat{\varepsilon}_{t-2} - \dots - \theta_q\hat{\varepsilon}_{t-q} \end{aligned}$$

Помилка прогнозування становитиме

$$y_{t+\tau} - \hat{y}_{t+\tau|t} = y_t + z_{t+1} + z_{t+2} + \dots + z_{t+\tau} - (y_t + \hat{z}_{t+1|t} + \hat{z}_{t+2|t} + \dots + \hat{z}_{t+\tau|t}).$$

Оскільки

$$\begin{aligned} z_{t+\tau} - \hat{z}_{t+\tau|t} &= \varepsilon_{t+\tau} + \theta_1\varepsilon_{t+\tau-1} + \dots + \theta_{\tau-1}\varepsilon_{t+1} + \theta_\tau(\varepsilon_t - \hat{\varepsilon}_t) + \dots = \\ &= \varepsilon_{t+\tau} + \theta_1\varepsilon_{t+\tau-1} + \dots + \theta_{\tau-1}\varepsilon_{t+1}, \end{aligned}$$

бо $\hat{\varepsilon}_j = \varepsilon_j$ при $j \leq t$, то

$$\begin{aligned} y_{t+\tau} - \hat{y}_{t+\tau|t} &= \varepsilon_{t+1} + \\ &+ \varepsilon_{t+2} + \theta_1\varepsilon_{t+1} + \\ &\dots \dots\dots \\ &+ \varepsilon_{t+\tau} + \theta_1\varepsilon_{t+\tau-1} + \dots + \theta_{\tau-1}\varepsilon_{t+1} = \\ &= \varepsilon_{t+\tau} + (1 + \theta_1)\varepsilon_{t+\tau-1} + \dots + (1 + \theta_1 + \dots + \theta_{\tau-1})\varepsilon_{t+1}. \end{aligned}$$

Отже,

$$MSE = \mathbf{E}(y_{t+\tau} - \hat{y}_{t+\tau|t})^2 = (1 + (1 + \theta_1)^2 + \dots + (1 + \theta_1 + \dots + \theta_{\tau-1})^2)\sigma^2.$$

3.9 Аналіз часових рядів Бокса-Дженкінса

Джордж Бокс та Гвілім Дженкінс розвинули метод аналізу часових рядів, які можна задати у вигляді *ARIMA*-процесів. За допомогою цієї методики можна знаходити всі невідомі параметри, включаючи кількість лагових змінних. Вибір відповідної моделі для часового ряду здійснюється за три етапи.

Ідентифікація моделі. Через порівняння функцій автокореляції та часткових автокореляційних функцій часового ряду з теоретичними функціями різних *ARIMA*(p, d, q)-процесів обираються параметри моделі: p – кількість лагових змінних, q – кількість лагових збурень, d – кількість різниць.

Оцінка моделі. За допомогою відповідної методики, найчастіше при максимізації функції правдоподібності, знаходяться коефіцієнти моделі.

Діагностика моделі. За допомогою різноманітних критеріїв, які перевіряють всі властивості залишків, придатність моделі, встановлюється, чи адекватна модель відповідному процесу.

Ідентифікація моделі. Першим кроком ідентифікації є визначення, чи є поданий часовий ряд стаціонарним, чи необхідно враховувати сезонні коливання. Стаціонарність ряду може бути визначена з графіку часового ряду. Він повинен показувати незначні коливання навколо деякого постійного рівня. Також стаціонарність можна визначити з графіка функції автокореляції. Нестаціонарні процеси часто мають графік з повільно спадаючою кривою. Сезонність або періодичність може бути визначена з графіка функції автокореляції або графіка сезонності. Якщо процес здається нестаціонарним, то застосовують метод різниць. На практиці достатньо одного чи двох застосувань цього методу. Якщо після цього отриманий ряд здається стаціонарним, то величину d можна вважати визначеною. Якщо дуже важко визначити, чи є ряд стаціонарним, то будується корелограма, графік якої не спадає до 0 для нестаціонарних рядів.

Згадаємо, що корелограмою для часових рядів є графік коефіцієнта автокореляції в залежності від його порядку. Для білого шуму при великих T наближений розподіл кожного γ_k є нормальним з нульовим середнім та дисперсією $\frac{1}{T}$. Тому надійним інтервалом з рівнем надійності 0,95, у якому повинні лежати γ_k , щоб процес можна

було віднести до стаціонарних, буде $\left(-\frac{1,96}{\sqrt{T}}; \frac{1,96}{\sqrt{T}}\right)$. На практиці більш важливо дослідити загальну поведінку коефіцієнтів автокореляції, тому потрібно, щоб усі γ_k для $k > q$ попадали у надійний інтервал. Тоді можна говорити, що ряд розвивається як $MA(q)$ -процес. Всі можливі види функції автокореляції наведені у таблиці.

Форма корелограми	Модель часового ряду
Експоненційна, наближається до нуля.	$AR(p)$ (треба використовувати графік часткової автокореляції для визначення порядку p).
Коливання навколо нуля поступово затухають.	$AR(p)$ (треба використовувати графік часткової автокореляції для визначення порядку p).
Один чи більше піків, інші точки близькі до нуля.	$MA(q)$ (порядок q визначається точкою, де графік попадає у надійний інтервал).
Затухання після декількох піків.	$ARMA$ -модель.
Всі значення дорівнюють нулю або близькі до нуля.	Випадкові дані, тобто часовий ряд згенерований випадковими чинниками.
Великі значення повторюються періодично.	Значний вплив сезонних коливань.
Нема затухання до нуля.	Дані нестаціонарні.

Часткова корелограма використовується для визначення кількості лагових змінних у $AR(p)$ -процесі. Очевидно, що з теоретичної точки зору, оцінка коефіцієнта автокореляції k -того порядку при $k > p$ повинна дорівнювати нулю. Таким чином, якщо процес згенерований $AR(p)$ -процесом, то на частковій корелограмі графік повинен спадати до нуля, тобто належати довірчому інтервалу, при $k > p$. Як і в попередньому випадку, використовується інтервал $\left(-\frac{1,96}{\sqrt{T}}; \frac{1,96}{\sqrt{T}}\right)$.

На практиці графіки автокореляції та часткової автокореляції не завжди є доволі ясними, тому іноді є досить важким завданням визначити параметри моделі. Іноді коли на діаграмах не можна побачити наближення до нуля, використовують змішані про-

цеси *ARMA*. Іноді, оцінюють декілька можливих моделей і обирають найкращу за допомогою третього етапу – діагностики моделі. Загальний алгоритм процесу ідентифікації наведений у діаграмі.



Оцінка моделі. Розглянемо застосування методу максимальної правдоподібності, який дозволяє отримувати оцінки коефіцієнтів для будь-якої моделі *ARMA*-процесу.

Запишемо *ARMA*(*p*, *q*)-процес у вигляді

$$y_t - \mu = \sum_{j=1}^p \phi_j (y_{t-j} - \mu) + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}.$$

Розподіл вектора $\{Y_T\}$ залежить від $p + q + 2$ параметрів ϕ_i ($i = \overline{1, p}$), θ_j ($j = \overline{1, q}$), μ , σ^2 . Згадаємо, що μ є математичним сподіванням кожного y_t , σ^2 – дисперсія кожного ε_t .

Припустимо, що елементи білого шуму ε_t мають нормальний розподіл. І тому через лінійність моделі вектор $\{Y_T\}$ також нормально розподілений з середнім μI та коваріаційною матрицею $\sigma^2 V(\phi, \theta)$. Тут $I = (1, 1, \dots, 1)'$, $\phi = (\phi_1, \phi_2, \dots, \phi_p)'$, $\theta = (\theta_1, \theta_2, \dots, \theta_q)'$.

Щільність розподілу такого вектора має вигляд ($x \in \mathbf{R}^T$)

$$f(x, \phi, \theta, \mu, \sigma) = (2\pi)^{-T/2} (\sigma^{2T} \det V(\phi, \theta))^{-1/2} \times \\ \times \exp \left\{ -\frac{(x - \mu I)' (V(\phi, \theta))^{-1} (x - \mu I)}{2\sigma^2} \right\}.$$

Тому натуральний логарифм функції правдоподібності такий:

$$\ln L(Y_T, \phi, \theta, \mu, \sigma) = -\frac{T}{2} \ln 2\pi - T \ln \sigma - \frac{1}{2} \ln \det V(\phi, \theta) - \\ - \frac{(Y_T - \mu I)' (V(\phi, \theta))^{-1} (Y_T - \mu I)}{2\sigma^2}.$$

Оцінки ϕ_i , θ_j , μ , σ^2 отримуються завдяки чисельній максимізації наведеного логарифма функції правдоподібності. При великій кількості спостережень стає важким при розрахунках знаходити обернену матрицю. Для практичних розрахунків, коли, як правило $T < 200$, це не становить для сучасного програмного забезпечення великої проблеми.

Діагностика моделі. Після того, як отримані коефіцієнти відповідної моделі, останнім етапом є її перевірка на адекватність. Це можна зробити декількома шляхами. По-перше, можна визначити залишки і перевірити, чи задовольняють вони властивостям білого шуму. По-друге, можна порівняти модель з іншими, які утворені з даної додаванням чи відніманням одного чи двох лагових змінних чи залишків. По-третє, можна провести спектральний аналіз моделі. У цьому розділі ми послідовно розглянемо всі ці можливості діагностики.

Аналіз залишків. Всі теоретичні *ARMA* моделі включають до свого складу випадковий компонент з властивостями білого шуму. Залишки моделі отримуються завдяки відніманню від реальних значень відомих параметрів моделі. Для *AR(p)*-процесу

$$y_t - \mu = \sum_{j=1}^p \phi_j (y_{t-j} - \mu) + \varepsilon_t$$

послідовність оцінок залишків будується за правилом:

$$\hat{\varepsilon}_t = y_t - \hat{\mu} - \sum_{j=1}^p \hat{\phi}_j (y_{t-j} - \hat{\mu}).$$

Зауважимо, що при визначенні $\hat{\varepsilon}_t$ при $t \leq p$ вважаємо $y_j = \hat{\mu}$ для $j \leq 0$.

Для $MA(q)$ -моделі

$$y_t = \mu + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$

послідовність оцінок залишків будується рекурентно (з врахуванням $\hat{\varepsilon}_j = 0, j \leq 0$):

$$\begin{aligned} \hat{\varepsilon}_1 &= y_1 - \hat{\mu}, \\ \hat{\varepsilon}_2 &= y_2 - \hat{\mu} - \hat{\theta}_1 \hat{\varepsilon}_1, \\ \dots &\dots \dots \\ \hat{\varepsilon}_t &= y_t - \hat{\mu} - \sum_{j=1}^q \hat{\theta}_j \hat{\varepsilon}_{t-j} \text{ при } t > q. \end{aligned}$$

Нарешті, для $ARMA(p, q)$ -процесу

$$y_t - \mu = \sum_{j=1}^p \phi_j (y_{t-j} - \mu) + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j},$$

маємо

$$\begin{aligned} \hat{\varepsilon}_1 &= y_1 - \hat{\mu}, \\ \hat{\varepsilon}_2 &= y_2 - \hat{\mu} - \hat{\phi}_1 (y_1 - \hat{\mu}) - \hat{\theta}_1 \hat{\varepsilon}_1, \\ \dots &\dots \dots \\ \hat{\varepsilon}_t &= y_t - \hat{\mu} - \sum_{j=1}^p \hat{\phi}_j (y_{t-j} - \hat{\mu}) - \sum_{j=1}^q \hat{\theta}_j \hat{\varepsilon}_{t-j} \text{ при } t > \max(p, q). \end{aligned}$$

Отримані залишки повинні бути протестовані на відповідність білому шуму. По-перше, властивістю білого шуму є те, що стандартне відхилення автокореляції є близьким до $\frac{1}{\sqrt{T}}$. Таким чином, можна

порівняти значення автокореляції залишків з цим значенням. Слід пам'ятати, що ці залишки не є білим шумом, а лише згенеровані за допомогою віднімання від реальних значень теоретичних значень, тому є досить природнім, що автокореляція малого порядку буде значно більшою за автокореляцію великого порядку. На питання, чи являються автокореляції перших m порядків задовільними для того, щоб залишки відповідали процесу білого шуму, дає відповідь

така статистика: $Q = T \sum_{j=1}^m \hat{\gamma}_j^2$, яка має χ_{m-p-q}^2 -розподіл. Якщо мо-

дель складається лише з однієї константи, то потрібно відняти ще одну ступінь волі. На жаль, цей тест дає дуже погане наближення, що веде часто до відхилення моделі. Тому цей тест був модифікований. Статистикою критерію є $Q = T(T+2) \sum_{j=1}^m \frac{1}{T-j} \hat{\gamma}_j^2$. Ця величина

повинна мати той самий χ_{m-p-q}^2 -розподіл. Але й цей тест не завжди дає добрі результати.

Порівняння моделі. Нехай ми ідентифікували $ARMA(p, q)$ -процес як задовільну модель для опису часового ряду та оцінили коефіцієнти за допомогою методу максимальної правдоподібності, причому L_0 – максимальне значення логарифму функції правдоподібності. Тепер повторимо цю процедуру для $ARMA(p+1, q)$ - та $ARMA(p, q+1)$ -моделей. Отримаємо значення L_1 та L_2 відповідно. Згідно зі стандартною теорією тестування функції правдоподібності, якщо початкова $ARMA(p, q)$ модель є коректною, то статистики $2(L_1 - L_0)$ та $2(L_2 - L_0)$ розподілені як χ_1^2 -розподіл. Така перевірка є дуже простою. Але, коли дані є сильно корельовані між собою, таке тестування може давати невірні результати.

Числові критерії. Після того, як оцінені коефіцієнти моделі, треба визначити наскільки добре ця модель описує часовий ряд. На відміну від попередніх тестувань, числові критерії лише видають деяке значення, по якому можна судити про адекватність моделі. Загальна характеристика критеріїв з їх властивостями подана у таблиці:

Назва критерію	Формула підрахунку	Бажаний екстремум
Коефіцієнт детермінації	$R^2 = 1 - \frac{\varepsilon'_T \varepsilon_T}{\sum_{t=1}^T (t_t - \bar{y})^2}$	1
Зважений коефіцієнт детермінації	$\bar{R}^2 = 1 - \frac{T-1}{T-p-q} (1 - R^2)$	1
Прогнозний критерій	$FC = \frac{\varepsilon'_T \varepsilon_T}{T-p-q} \left(1 + \frac{p+q}{T} \right)$	min
Інформаційний критерій	$AIC = \ln \left(\frac{\varepsilon'_T \varepsilon_T}{T} \right) + 2 \frac{p+q}{T}$	min
Інформаційний критерій Шварца-Ріссанена	$SIC = \ln \left(\frac{\varepsilon'_T \varepsilon_T}{T} \right) + \frac{p+q}{T} \ln T$	min
Критерій Ханнана-Квіна	$HQ = \ln \left(\frac{\varepsilon'_T \varepsilon_T}{T} \right) + c \frac{p+q}{T} \ln \ln T,$ $c \geq 2$	min

Застосування всіх критеріїв одночасно може призвести до різних рішень. Наприклад, коефіцієнт детермінації може призводити до збільшення параметрів p та q , тому цей критерій застосовувати для прийняття рішень не варто. Зважений коефіцієнт детермінації навпаки призводить до зменшення числа доданків у $ARMA$ -процесах. Інші критерії є більш сильними. Найбільш цікавими є останні два критерії, які є спроможними, тобто при $T \rightarrow +\infty$ повністю ідентифікують істинну модель. Критерій AIC , як правило, призводить до збільшення числа параметрів.

Розділ 4

Новітні моделі та методи аналізу часових рядів

4.1 Огляд моделей та методів

4.1.1 Лінійні моделі

Моделі з довгою пам'яттю. Побудова графіка емпіричної кореляційні функції для багатьох показників, зокрема у фінансах, часто демонструє характер повільного спадання її значень до нуля. Це явище довгої пам'яті, як уже зазначалось, може описуватись *ARIMA*-процесами.

Більш загальна модель, яка зараз починає все ширше застосовуватись, це *ARFIMA*(p, d, q)-процес, який узагальнює *ARIMA*(p, d, q) на випадок нецілих d . Визначимо для $d \in (-0, 5; 0, 5)$:

$$(1 - B)^d = \sum_{k=0}^{+\infty} (-1)^k C_d^k B^k, \text{ де } C_d^k = \frac{d(d-1)\dots(d-k+1)}{k!}.$$

Якщо часовий ряд $(1 - B)^d y_t \in ARMA(p, q)$ -процесом, то кажуть, що $\{y_t\}$ описується *ARFIMA*(p, d, q) моделлю.

4.1.2 Моделі часових рядів зі змінною волатильністю.

У багатьох реальних задачах дослідника цікавить не лише вивчення та передбачення рівня (середнього) часового ряду, але і волатильність (дисперсія) показників. Скажімо для процесу $AR(p)$:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

без одиничних коренів легко отримати оптимальний лінійний прогноз рівня y_t за спостереженнями y_{t-1}, y_{t-2}, \dots :

$$\hat{y}_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p}.$$

Проте в багатьох реальних задачах волатильність такого ряду не є сталою. Наприклад для електронних торгів вона змінюється на протязі доби, у сезонному виробництві – протягом року, інвестори від більш ризикових акцій (з більшою волатильністю) очікують вищі дивіденди як компенсацію. Для опису змінної волатильності одним із найпопулярніших сучасних класів моделей є гетероскедастичність.

Авторегресійна умовна гетероскедастичність. Основна ідея умовної гетероскедастичності полягає у описі білого шуму ε_t знову ж за допомогою процесу авторегресії. Оскільки ми зацікавлені у змінній волатильності, то авторегресійна модель розглядається не для ε_t , а для ε_t^2 .

Кажуть, що ε_t авторегресійний умовно гетероскедастичний процес порядку m ($ARCH(m)$), якщо він задовольняє рівняння

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_m \varepsilon_{t-m}^2 + w_t,$$

де w_t – білий шум, $\alpha_0, \alpha_1, \dots, \alpha_m$ – сталі.

Досить часто зустрічається інше визначення $ARCH(m)$ -процесу:

$$\varepsilon_t = \sqrt{h_t} \nu_t, \quad (4.1)$$

де ν_t – незалежні однаково розподілені випадкові величини з нульовим середнім та одиничною дисперсією, а h_t визначається за формулою

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_m \varepsilon_{t-m}^2. \quad (4.2)$$

У сучасній журнальній літературі можна знайти величезну кількість різноманітних узагальнень $ARCH$ -процесу, які використовують для певних класів прикладних задач. Розглянемо деякі з них.

Узагальнена авторегресійна умовна гетероскедастична модель. Кажуть, що ε_t – узагальнений авторегресійний умовно гетероскедастичний процес ($GARCH(r, m)$), якщо він описується рівняннями :

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

$$h_t = \delta_0 + \delta_1 h_{t-1} + \delta_2 h_{t-2} + \dots + \delta_r h_{t-r} + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_m \varepsilon_{t-m}^2,$$

де $\delta_0, \delta_1, \delta_2, \dots, \delta_r, \alpha_1, \alpha_2, \dots, \alpha_m$ – сталі, а ν_t – такі як в (4.1). Коефіцієнти $\delta_0, \delta_1, \delta_2, \dots, \delta_r, \alpha_1, \alpha_2, \dots, \alpha_m$ вибирають такими, щоб $ARMA$ -процес ε_t^2 був без одиничного кореня. Якщо ж

$$\sum_{i=1}^r \delta_i + \sum_{j=1}^m \alpha_j = 1,$$

то ε_t^2 – процес з одиничним коренем. Відповідну модель називають *інтегрованим GARCH* і позначають $IGARCH$.

ARCH в середньому. Для моделювання таких ситуації в фінансах, коли більш ризикові цінні папери дають більший прибуток у середньому, була запропонована нова регресійна модель $ARCH$ в середньому ($ARCH - M$) яка задається рівняннями (4.1), (4.2) і

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

де x_t, β – вектори, δ – стала. Вплив волатильності ε_t на рівень часового ряду y_t визначає параметр δ .

Моделювання ефекту важеля. Часовий ряд ε_t в (4.1) задається експоненційною $GARCH$ ($EGARCH$) моделлю, якщо h_t описується рівнянням :

$$\ln h_t = \alpha_0 + \sum_{j=1}^{+\infty} \alpha_j (|\nu_{t-j}| - \mathbf{E}|\nu_{t-j}| + c\nu_{t-j})$$

В такій моделі параметр c дає змогу ввести ефект асиметрії. При $c = 0$ рівні за модулем додатні та від'ємні значення ν_{t-j} мають однаковий вплив на волатильність. При $c < 0$ додатні значення ν_{t-j} збільшують волатильність менше ніж від'ємні. Такою моделлю можна описувати відомий в економіці ефект важеля – різке зниження цін акцій веде

до більшого збільшення ризикованості акцій, ніж збільшення їхньої ціни.

Ще одна модель для опису асиметрії базується на заміні других степенів ε_{t-j} на квадратичні функції від ε_{t-j} . В такій *квадратичній ARCH* (*QARCH*) моделі h_t визначається рівнянням:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \varepsilon_{t-1} + \dots + \alpha_m \varepsilon_{t-m}^2 + \beta_m \varepsilon_{t-m}.$$

Параметри α_i та β_i мають обиратись так, щоб значення h_t були невід'ємними.

Інша *GARCH* модель для ефекту важеля задає h_t рівнянням:

$$h_t = \delta_0 + \delta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + c \cdot I(\varepsilon_{t-1}) \varepsilon_{t-1}^2,$$

де $I(x) = \begin{cases} 1, & \text{якщо } x > 0, \\ 0, & \text{інакше} \end{cases}$ – функція Хевісайда.

Нелінійна модель волатильності. Для внесення нелінійності в авторегресію при визначенні волатильності застосовують степеневе перетворення. Таку модель часового ряду називають *нелінійною ARCH* (*NARCH*). В ній

$$h_t = (\alpha_0 + \alpha_1 (\varepsilon_{t-1}^2)^\delta + \dots + \alpha_m (\varepsilon_{t-m}^2)^\delta)^{\frac{1}{\delta}}$$

де $\delta, \alpha_0 > 0$, $\alpha_i \geq 0$ для $i = 1, 2, \dots, m$ – сталі.

4.1.3 Нелінійні моделі

Нелінійна модель, подає значення часового ряду в момент t (скажімо ціну акцій y_t) як деяку нелінійну функцію f від білого шуму (попередніх надходжень інформації)

$$y_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots) \quad (4.3)$$

Специфікація функції f веде до цілого ряду сучасних прикладних моделей. Розглянемо деякі з них.

Білінійна модель. Легко зрозуміти, що лінійна модель виду $y_t = \mu + \sum_{i=1}^{+\infty} \psi_i \varepsilon_{t-i}$ може розглядатись як апроксимація моделі (4.3) при

розкладі функції f в ряд Тейлора з обмеженням лише членами першого порядку. Білінійна модель є природним узагальненням, якщо для покращення апроксимації включити члени другого порядку з розкладу f . В цьому разі матимемо таке зображення нашого ряду:

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{i=1}^m \sum_{j=1}^s \beta_{ij} y_{t-i} \varepsilon_{t-j} + \varepsilon_t,$$

де $\phi_i, \theta_j, \beta_{ij}$ – сталі, p, q, m та s – невід’ємні цілі числа.

Порогова авторегресійна модель. Однією з перших класичних моделей, якими намагались описати змінне середнє була шматково-лінійна модель. У ній припускалось, що весь період спостережень може бути розбитий на певні проміжки часу, спостереження на яких описуються лінійними регресійними моделями. Для покращення лінійної апроксимації до реальних процесів зараз застосовується цілий ряд, так званих, порогових моделей. Узагальнення тут відбулось як у напрямку відмови від простої лінійності, так і в більш гнучкому механізмі переключення від одного режиму до іншого. Часовий ряд y_t породжується *самозбуджуючою TAR (SETAR) моделлю з k режимилами* та пороговими змінними y_{t-d} , якщо

$$y_t = \phi_0^{(j)} + \phi_1^{(j)} y_{t-1} + \dots + \phi_p^{(j)} y_{t-p} + \varepsilon_t^{(j)} \text{ для } c_{j-1} \leq y_{t-d} < c_j$$

де $k, d \in \mathbf{N}$, $j = 1, \dots, k$, c_j – дійсні числа такі, що $-\infty = c_0 < c_1 < \dots < c_{k-1} < c_k = +\infty$, $\varepsilon_t^{(j)}$ – набір білих шумів, які незалежні для різних j . Індекс j використовується для специфікації режиму.

Гладка перехідна модель. До багатьох прикладних даних попередня модель не може бути застосована, оскільки її умовне математичне сподівання не має необхідної гладкості в порогах c_j . Щоб уникнути такого недоліку використовують *гладкі перехідні AR (STAR) моделі*. У цьому разі процес y_t задовольняє рівняння

$$y_t = a_0 + \sum_{i=1}^p \phi_i y_{t-i} + F\left(\frac{y_{t-d} - \Delta}{s}\right) \left(a_1 + \sum_{i=1}^p \psi_i y_{t-i}\right) + \varepsilon_t,$$

де d – параметр затримки, Δ і s – параметри керування моментом та шкалою в моделі переходу, $F(\cdot)$ – деяка гладка функція переходу. На практиці F часто логістична, експоненційна чи функція розподілу.

Модель з марківським переключенням. Неперіодичне переключення між різними станами в багатьох економічних явищах спонукало до виникнення і вивчення нового класу моделей з переключенням. Нехай s_t – двостановий марківський ланцюг зі значеннями в $\{1, 2\}$. Кажуть, що y_t породжений *MSA* моделлю, якщо

$$y_t = \begin{cases} a_1 + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t^{(1)}, & \text{якщо } s_t = 1, \\ a_2 + \sum_{i=1}^p \psi_i y_{t-i} + \varepsilon_t^{(2)}, & \text{якщо } s_t = 2, \end{cases}$$

де $\varepsilon_t^{(1)}$ та $\varepsilon_t^{(2)}$ – незалежні між собою білі шуми.

Моделі з функціональними коефіцієнтами. Останні досягнення в непараметричних методах дали змогу ослабити параметричні обмеження, які часто неявно є в нелінійних моделях. Звичайно, на практиці тепер непараметричні методи використовуються для вибору параметричної нелінійної моделі, яка вже потім детально досліджується. Розглянемо дві такі моделі.

Кажуть, що y_t – процес *авторегресії з функціональними коефіцієнтами (FAR)*, якщо він задовольняє рівняння

$$y_t = f_1(Y_{t-1})y_{t-1} + \dots + f_p(Y_{t-1})y_{t-p} + \varepsilon_t,$$

де $Y_{t-1} = (y_{t-1}, \dots, y_{t-k})$. Стандартні вимоги до функцій f_i – неперервність та двічі диференційованість.

Така модель не є адитивною, що ускладнює дослідження її властивостей і роботу з нею. Поєднанням нелінійності з адитивною структурою є *нелінійна адитивна авторегресія (NAAR)*. Відповідний процес y_t визначається рівнянням

$$y_t = f_0(t) + \sum_{i=1}^p f_i(y_{t-i}) + \varepsilon_t.$$

4.1.4 Нейронні мережі

Одним з найактуальніших методів дослідження в сучасному аналізі часових рядів є *нейронні мережі*. Методи такого дослідження можна

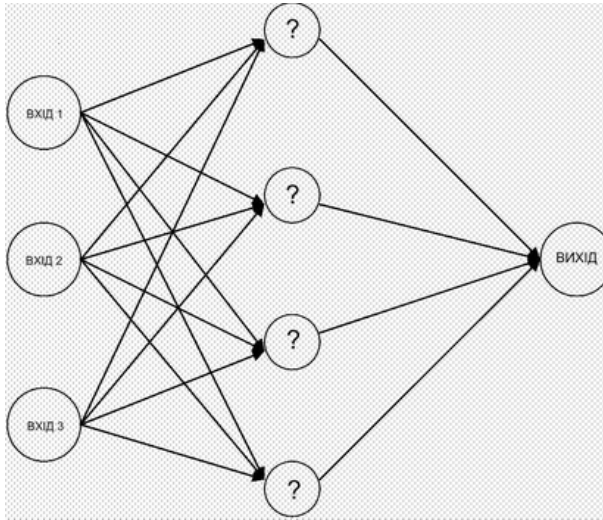


Рисунок 4.1: Приклад нейронної мережі

розглядати як напівпараметричні. Схема найпростішої нейронної мережі зображена на рисунку 4.1.

Вхідні вузли зв'язані з вузлами (нейронами) в прихованому рівні (структуру якого неможливо спостерігати безпосередньо). Кожен нейрон зв'язаний з вихідним вузлом. На рисунку зображено найпростішу нейронну мережу з одним прихованим рівнем і одностороннім напрямком передачі інформації. Звичайно на практиці розглядають мережі з більш складною структурою. Нехай x_i – стан i -того вхідного вузла. Тоді мережа активізується та стан j -того нейрона в прихованому рівні визначається так:

$$h_j = f_j \left(a_j + \sum_{i \rightarrow j} w_{ij} x_i \right),$$

де f_j – деяка функція активації, a_j називають зміщенням, w_{ij} – вагові коефіцієнти, а $\sum_{i \rightarrow j}$ означає суму по всіх вхідних вузлах, які впливають на j -тий нейрон.

Стан вихідного вузла визначається як

$$Y = f_0 \left(a_0 + \sum_{i=0} w_{j0} h_j \right).$$

На практиці часто використовують логістичну функцію

$$f_j(z) = \frac{\exp(z)}{1 + \exp(z)}$$

як функцію активації нейронів, а функцію Хевісайда як функцію активації вихідного вузла.

Застосування нейронної мережі здійснюється у два етапи. Спочатку розглядають тренувальний часовий ряд (y_t, x_t) , $t = 1, \dots, T$, де x_t – вектор входів, y_t – часовий ряд, який становить інтерес. На його базі будується нейронна мережа і оцінюються її параметри. На другому етапі побудована мережа використовується для прогнозування і перевірки точності передбачення.

4.1.5 Векторні часові ряди

Зрозуміло, що в більшості прикладних економічних, фінансових, соціологічних дослідженнях інтерес становить не один, а цілий набір характеристик, показників. Тому природним узагальненням є моделі векторозначних часових рядів. Такі ряди складаються з сукупностей одновимірних часових рядів як компонентів загальної моделі.

Більшість з моделей, які розглядалися раніше можуть бути прямо узагальнені на багатовимірний випадок. Потрібно лише розуміти процеси x_t , y_t , ε_t як векторні, відповідні коефіцієнти як матриці, а функції – як певні функції над векторами чи матрицями. Проте при таких узагальненнях потрібна особлива увага. Оскільки компоненти векторозначного часового ряду часто залежні між собою, то навіть найпростіше узагальнення одновимірної моделі іноді веде до розробки принципово нових методів для її обробки та дослідження. Ми опишемо лише одну модель векторного часового ряду, яка не є простим формальним узагальненням одновимірних моделей, розглянутих раніше.

Коінтегровані часові ряди. Векторний $(n \times 1)$ процес y_t називається *коінтегрованим часовим рядом*, якщо кожна його компонента

зокрема – нестационарний ряд з одиничним коренем $I(1)$ та існує ненульовий $(n \times 1)$ -вектор a такий, що $a'y_t$ – стаціонарний. ($I(1)$ означає, що сам ряд x_t нестационарний, але ряд $\Delta x_t = x_t - x_{t-1}$ є стаціонарним).

Такі ряди виникають в багатьох задачах де кожен з показників нестационарний, але проте існує тісний зв'язок (взаємозв'язання) одного показника через інші. Скажімо, курси кожної з кількох валют зокрема – нестационарні процеси, але різниця їх (домножених на певні сталі коефіцієнти) часто незначно флюктує навколо нуля і може розглядатись як стаціонарна.

4.2 Коротка бібліографія

Класичній теорії стаціонарних часових рядів присвячено величезну кількість книг. Обмежимося посиланням лише на три з них [1, 2, 3] де можна знайти достатньо повну бібліографію. Гарним вступом до теорії часових рядів з одиничним коренем з посиланням на класичні роботи є книга [1]. Її часто розглядають як "біблію" класичної теорії часових рядів. Різноманітні властивості ARIMA процесів обговорюються і в [2]. Першою з монографій про випадкові процеси з довгою пам'яттю є книга [5]. Різноманітні властивості ARCH процесів, їх узагальнення та чисельні посилання на журнальні джерела можна знайти в [6, 7]. Гарним вступом до різних аспектів теорії нелінійних часових рядів можуть слугувати монографії [9, 10]. Огляд методів дослідження білінійних моделей та їх властивостей подано в [11]. Посилання на недавні роботи з теорії порогових моделей можна знайти в [12, 13], а на гладкі перехідні моделі в [14]. Обговорення MSA міститься в [1]. Моделі з функціональними коефіцієнтами досліджувались в [15]. МСМС методи в застосуванні до загальних нелінійних моделей продемонстровано в [16]. Початковий огляд статистичних аспектів нейронних мереж можна знайти в [17, 18]. Опис моделі коінтеграції і чисельні посилання на журнальні публікації про неї містяться в [19, 20]. Моделі, які найбільше поширені в економіці та фінансах з реальним прикладами застосування обговорюються в [8, 21].

Бібліографія

- [1] Hamilton J.D. Time Series Analysis. Princeton, 1994.
- [2] Box G.E.P., Jenkins G.M., Reinsel G.C. Time Series Analysis: Forecasting and Control. 3rd edition, Prentice Hall, 1999.
- [3] Brockwell P.J., Davis R.A. Time Series: Theory and Method. Springer-Verlag, 1991
- [4] Beran J. Statistics for Long-memory Processes. Chapman&Hall, 1994
- [5] Gouriéroux Ch. ARCH Models and Financial Applications. Springer, 1997.
- [6] Engle R.F. ARCH. Oxford Univ press, 1995.
- [7] Шепард Н. Статистические аспекты моделей типа ARCH и стохастической волатильности // Обозрение прикладной и промышленной математики. Том 3., выпуск 6. (1996) Ст. 764-826.
- [8] Time Series Models in Econometrics, Finance and Other Fields. Ed. By D.R.Cox, D.V. Ninbley and O.E.Barndorff-Nielsen. Chapman&Hall, 1996.
- [9] Tong H. Nonlinear Time Series: A Dynamical System Approach. Oxford Univ. Press, 1990.
- [10] Priestley M.B. Non-linear and Nonstationary Time Series Analysis. Academic Press, 1988.
- [11] Terdic G. Bilinear Stochastic Models and Related Problems of Nonlinear Time Series Analysis. Springer, 1999.

- [12] Hansen B.E. Inference in TAR models // Studies in Nonlinear Dynamics and Econometrics, 1(1997). P.119-131.
- [13] Tsay R.S. Testing and modeling multivariate threshold models // Journal of the Americal Stat. Association. 84 (1998). P. 1188-1202.
- [14] Terasvirta T. Specification, estimation, and evaluation of smooth transition autoregressive models // Journal of American Stat. Association , 89 (1994). P.208-218.
- [15] Cai Z., Fan J., Yao Q. Functional-coefficient regression models for nonlinear time series // Working paper. University of North Carolina (1999).
- [16] Carlin B.P., Polson N.G., Stoffer D.S. A Monte Carlo Approach to Nonnormal and Nonlinear State Space Modeling // Journal of The American Stat. Association, 87 (1992). P.493-500.
- [17] Cheng B., Titterington D.M. Neural network: a review from a statistical perspective // Statistical Science 9 (1994). P.2-54.
- [18] Ripley B.D. Statistical aspects of neural networks// Networks and Chaos-Statistical and Probabilistic Aspects, eds. O.E.Brandorff-Nielson, J.L.Jensen, W.S.Kendall. Chapman&Hall (1993). P.40-123.
- [19] Johansen S. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford Univ. Press, 1995.
- [20] Banerjee A., Dolado J.J., Galbraith J.W., Hendry D.F. Cointegrated, Error Correction, and the Econometric Analysis of Non-stationary Data. Oxford Univ. Press, 1993.
- [21] Tsay R.S. Analysis of Financial Time Series. Wiley, 2002.

Частина II

Практична частина

Розділ 5

Завдання для практичних занять

5.1 Початковий аналіз часового ряду

1. Створити базу даних основних макроекономічних показників України за 1996-2006 роки: ВВП, експорт та імпорт товарів та послуг, грошова маса, інфляція, курс долара США до гривні, обсяг вкладів населення в банках тощо. Для роботи вибрати 5 часових рядів (не можуть повторюватися з іншими студентами). Кожен з рядів має містити не менше 30 спостережень. Для визначення потрібної інформації скористатися сайтами:

- www.me.gov.ua,
- www.bank.gov.ua

2. Імпортувати Вашу базу даних до MS Excel.

3. Провести графічний аналіз рядів даних. Визначити наявність трендового, сезонного компонентів.

4. Для обраних рядів даних визначити основні числові характеристики: вибіркове середнє, вибірккову дисперсію, коваріацію перших 10 порядків.

5. Провести агрегацію Ваших часових рядів (квартальні дані перевести у річні, місячні – у квартальні). З отриманими рядами провести дезагрегацію даних за допомогою методів:

- пропорційного розбиття;
- поліноміальної інтерполяції.

Порівняти отримані оцінки з фактичними.

5.2 Зглажування, виділення тренду, прогнозування

1. Для кожного часового ряду в базі даних застосувати метод аналізу:

- плинне середнє (з кроками 3, 4, 5);
- експоненціальне згладжування;
- подвійне експоненціальне згладжування;
- потрійне експоненціальне згладжування;
- метод Холта-Вінтерса;
- адаптивне згладжування;
- проста лінійна регресія.

Підрахувати прогнози на 2007 рік, скориставшись даними за 1996-2006 роки. Обчислити похибки прогнозування.

2. Для Вашої бази даних розробити і обрахувати адитивну модель з визначенням сезонних коливань. Зробити прогнози, обчислити похибки прогнозування.

3. Для Ваших часових рядів побудувати різні види трендів:

- лінійний тренд;
- квадратичний тренд;
- експоненціальний тренд;
- гіперболічний тренд;
- логарифмічний тренд;
- логістичний тренд;

Зробити на їх основі прогнози на поточний рік, підрахувати помилку прогнозування.

5.3 Аналіз стаціонарних часових рядів (*ARMA*-моделі), прогнозування

1. Провести нормування Ваших часових рядів за правилом:

$$\tilde{y}_t = \frac{y_t - \mathbf{E}y_t}{\sqrt{\mathbf{D}y_t}}.$$

2. Для отриманих у попередній задачі рядів обчислити моделі:

- *MA*(1);
- *MA*(2);
- *AR*(1);
- *AR*(2);
- *ARMA*(1, 1);
- *ARMA*(2, 1);
- *ARMA*(1, 2);
- *ARMA*(2, 2).

3. На основі аналізу корелограми та часткової корелограми визначити найкращі параметри моделей для Ваших часових рядів.

4. Для обчислених моделей побудувати прогнози, обчислити похибки прогнозування.

5. Порівняти моделі на основі точності прогнозів.

5.4 Аналіз *ARIMA*-процесів, прогнозування

1. Для Ваших часових рядів побудувати *ARIMA* моделі таких видів:

- *ARIMA*(1, 1, 1); *ARIMA*(1, 2, 1);
- *ARIMA*(2, 1, 1); *ARIMA*(2, 2, 1);
- *ARIMA*(3, 1, 1); *ARIMA*(3, 2, 1);

- $ARIMA(1, 1, 2)$; $ARIMA(1, 2, 2)$;
- $ARIMA(1, 1, 3)$; $ARIMA(1, 2, 3)$.

2. По всіх обчислених моделях побудувати прогнози, оцінити похибку прогнозування.

Розділ 6

Завдання для самоостійного опрацювання та контролю знань

6.1 Задачі

1. Поданий стаціонарний часовий ряд з 16 спостережень: 1.6, 0.8, 1.2, 0.5, 0.9, 1.1, 1.1, 0.6, 1.5, 0.8, 0.9, 1.2, 0.5, 1.3, 0.8, 1.2. Для цього часового ряду: а) побудувати графік; б) за допомогою графіка визначити наближено коефіцієнт автокореляції першого порядку; в) побудувати графік залежності y_{t+1} від y_t та визначити спочатку наближено, а потім підрахувати точно коефіцієнт автокореляції першого порядку.

2. Нехай поданий часовий ряд $\{Y_T\}$, який складається з місячних спостережень. Сезонний компонент $s_t = s_{t-12}$ є сталим для кожного місяця, тобто рівність виконується при всіх t . Нехай також ε_t – стаціонарний ряд випадкових збурень.

1. Для глобального лінійного тренду та адитивної моделі ми маємо $y_t = a + bt + s_t + \varepsilon_t$. Показати, що застосування оператора Δ^{12} для часового ряду призводить до його стаціонарності.

2. Для глобального лінійного тренду та мультиплікативної моделі ми маємо $y_t = (a + bt)s_t + \varepsilon_t$. Чи буде оператор Δ^{12} зводити часовий ряд до стаціонарного? Якщо ні, то знайти інший різницьовий оператор, який буде зводити $\{Y_T\}$ до стаціонарного ряду.

3. Знайти функцію автокореляції для $MA(2)$ -процесу виду

$$y_t = \varepsilon_t + 0.7\varepsilon_{t-1} - 0.2\varepsilon_{t-2}.$$

4. Показати, що функція автокореляції для $MA(q)$ -процесу виду $y_t = \varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_{t-q}$ подається формулою

$$\rho_k = \begin{cases} 1 - \frac{k}{q+1}, & k = 0, 1, \dots, q, \\ 0, & k > m. \end{cases}$$

5. Показати, що нескінченний $MA(\infty)$ -процес виду

$$y_t = \varepsilon_t + C(\varepsilon_{t-1} + \varepsilon_{t-2} + \dots),$$

де C – константа, є нестационарним. Показати крім того, що якщо до цього ряду застосувати метод різниць, то отриманий ряд стане стаціонарним.

6. Знайти функцію автокореляції для $AR(1)$ -процесу виду

$$y_t - \mu = 0.7(y_{t-1} - \mu) + \varepsilon_t.$$

Зобразити графік ρ_k для $k = 0, \pm 1, \pm 2, \dots, \pm 6$.

7. Нехай $y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$. Показати, що якщо μ – константа, то функція автокореляції не залежить від μ .

8. Стаціонарний процес $\{Y_T\}$ має автоковаріаційну функцію γ_k . Новий стаціонарний процес $\{Z_T\}$ визначений рівністю $z_t = y_t - y_{t-1}$. Виразити його функцію автоковаріації через γ_k та знайти її явно, якщо $\gamma_k = \lambda^{|k|}$.

9. Для кожної з нижченаведених моделей виразити їх за допомогою лагових поліномів, визначити, чи є вони стаціонарними та/або перетворювальними.

а) $y_t = 0.3y_{t-1} + \varepsilon_t$;

б) $y_t = \varepsilon_t - 1.3\varepsilon_{t-1} + 0.4\varepsilon_{t-2}$;

$$в) y_t = 0.5y_{t-1} + \varepsilon_t - 1.3\varepsilon_{t-1} + 0.4\varepsilon_{t-2};$$

$$г) y_t = 0.03 + 0.7y_{t-1} + 0.9y_{t-2} + \varepsilon_t + 0.2\varepsilon_{t-1} - 0.5\varepsilon_{t-2}.$$

Для моделі (а) знайти еквівалентний MA -запис.

10. Показати, що функція автокореляції для $ARMA(1, 1)$ -моделі виду $y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$ подається виразом:

$$\rho_k = \begin{cases} \frac{(1+\phi\theta)(\phi+\theta)}{1+\theta^2+2\phi\theta}, & k = 1 \\ \phi\rho_{k-1}, & k = 2, 3, \dots \end{cases}$$

11. Для моделі $(1 - B)(1 - 0.2B)y_t = (1 - 0.5B)\varepsilon_t$:

а) ідентифікувати модель як $ARIMA(p, d, q)$, тобто знайти параметри p, d, q ;

б) визначити, чи є процес стаціонарним;

в) підрахувати перші три коефіцієнти ϕ та θ .

12. При яких значеннях c $AR(2)$ -процес виду $y_t = y_{t-1} + cy_{t-2} + \varepsilon_t$ є стаціонарним. Показати що $AR(3)$ -процес $y_t = y_{t-1} + cy_{t-2} - cy_{t-3} + \varepsilon_t$ – нестаціонарний для будь-якого значення c .

13. Вивести для $AR(1)$ -процесу, який має середнє μ , МНК-оцінки коефіцієнтів.

14. Припустимо, що $AR(2)$ -процес $\{Y_T\}$ – стаціонарний. Підрахуйте коефіцієнти $\theta_0, \theta_1, \theta_2, \theta_3$ його $MA(\infty)$ -представлення.

15. Як оцінити коефіцієнти стаціонарного $AR(2)$ -процесу?

16. Як прогнозувати $y_{t+\tau}$ за допомогою закону ітеративних сподівань? Як виглядає прогноз $\hat{y}_{t+\tau|t}$, при $\tau \rightarrow +\infty$ стаціонарного $AR(2)$ -процесу?

17. При якій умові $ARMA(1, 1)$ -процес $y_t = \phi y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$, де ε_t – білий шум, – стаціонарний, при якій – перетворюваний? Підрахуйте перші три коефіцієнти у $MA(\infty)$ -представлення.

18. Як виглядає оптимальні прогнози для y_{t+1} та y_{t+2} , якщо відомі y_{t-i} та ε_{t-i} , $i \geq 0$, а $\{Y_T\}$ є $ARMA(2, 1)$ -процесом? Опишіть, як потрібно підраховувати прогнози на практиці, коли ε_t невідомі?

19. Нехай $\{Y_T\}$ – випадкове блукання з напрямом (сталим трендом). Як виглядає прогноз $y_{t+\tau}$? Знайдіть прогнозу помилку та середню квадратичну похибку в цьому випадку.

20. Нехай $\{Y_T\}$ описується трендовою моделлю $y_t = \beta_0 + \beta_1 t + \varepsilon_t$, ε_t – білий шум. Як виглядає прогноз $y_{t+\tau}$? Знайдіть прогнозну помилку та середню квадратичну похибку в цьому випадку.

21. Припустимо, що для часового ряду, який складається з 100 спостережень, підраховані коефіцієнти кореляції: $\rho_1 = 0.31$, $\rho_2 = 0.37$, $\rho_3 = 0.05$, $\rho_4 = 0.06$, $\rho_5 = 0.21$, $\rho_6 = 0.11$, $\rho_7 = 0.08$, $\rho_8 = 0.05$, $\rho_9 = 0.12$, $\rho_{10} = 0.01$. Запропонуйте *ARMA* модель, яка буде задовільною.

6.2 Завдання для контролю знань

1. Адитивна та мультиплікативна моделі часового ряду. Особливості складників таких моделей.
2. Прогнозування значень часового ряду. Міри точності прогнозів.
3. Лаговий оператор. Поліном від лагового оператора.
4. Стаціонарність часових рядів.
5. Функція автокореляції. Особливості поведінки ряду та залежності між його елементами.
6. Функція правдоподібності. Метод оцінки параметрів часового ряду.
7. Перевірка гіпотез про часовий ряд з допомогою функції правдоподібності.
8. Загальний порядок аналізу часових рядів.
9. Дезаграгування та відновлення значень часового ряду.
10. Пропорційне розбиття часового ряду.
11. Поліноміальна інтерполяція при розбитті часового ряду.
12. Метод усереднення.
13. Звичайне експоненціальне згладжування часового ряду.
14. Подвійне експоненціальне згладжування часового ряду.
15. Потрійне експоненціальне згладжування часового ряду.
16. Адаптивне згладжування часового ряду.
17. Несезонна модель Холта-Вінтера.
18. Адитивна модель із визначенням сезонних коливань.
19. Означення білого шуму.
20. Означення *MA(q)*-процесу. Представлення з допомогою лагового оператора.
21. Числові характеристики (математичне сподівання, дисперсія, коваріація) *MA(q)*-процесу, його стаціонарність.

22. $MA(1)$ -, $MA(2)$ - та $MA(\infty)$ -процеси (числові характеристики). Стаціонарність $MA(\infty)$ -процесу.
23. Процес авторегресії. Представлення з допомогою лагового оператора.
24. $AR(1)$ -процес, його MA -представлення, стаціонарність та числові характеристики.
25. $AR(2)$ -процес, його MA -представлення, стаціонарність та числові характеристики.
26. $AR(p)$ -процес, його MA -представлення, стаціонарність та числові характеристики.
27. AR -представлення $MA(1)$ -процесу.
28. AR -представлення $MA(q)$ -процесу.
29. Означення та представлення з допомогою лагового оператора $ARMA$ -процесу. Його стаціонарність.
30. $MA(\infty)$ -представлення $ARMA$ -процесу.
31. $AR(\infty)$ -представлення $ARMA$ -процесу.
32. Прогнозування значень $MA(\infty)$ -процесу. Середня похибка прогнозу.
33. Прогнозування значень $MA(q)$ -процесу. Середня похибка прогнозу.
34. Надійний інтервал для прогнозу у випадку нормального розподілу збурень.
35. Прогнозування значень $AR(1)$ -процесу з допомогою його $MA(\infty)$ -представлення. Середня похибка прогнозу та її граничне значення.
36. Формула Вінера-Колмогорова для $MA(q)$ - та $AR(p)$ -процесів.
37. Прогнозування значень $AR(1)$ -процесу з допомогою формули Вінера-Колмогорова. Закон ітеративних сподівань.
38. Прогнозування значень $AR(p)$ -процесу з допомогою формули Вінера-Колмогорова.
39. Прогнозування значень $MA(1)$ -процесу з використанням формули Вінера-Колмогорова. Альтернативний метод.
40. Прогнозування значень $MA(q)$ -процесу з використанням формули Вінера-Колмогорова. Альтернативний метод.
41. Прогнозування значень $ARMA$ -процесів.
42. Оцінювання параметрів AR -процесів. Застосування методу найменших квадратів.
43. Оцінювання параметрів MA -процесів (на прикладі $MA(1)$ -процесу).
44. Процеси з одиничним коренем. $ARIMA(p, d, q)$ -процес та зведення його до $ARMA$ -процесу.

45. Прогнозування *ARIMA*-процесів. Похибка прогнозу та її середній квадрат.
46. Загальна схема Бокса-Дженкінса аналізу часових рядів.
47. Ідентифікація моделі часового ряду.
48. Оцінка моделі часового ряду методом максимальної правдоподібності.
49. Діагностика моделі часового ряду (аналіз залишків).
50. Діагностика моделі часового ряду (порівняння з сусідніми за параметрами).