

Бідюк П.І., Романенко В.Д., Тимошук О.Л.

АНАЛІЗ ЧАСОВИХ РЯДІВ

(навчальний посібник)

*Міністерство Освіти і Науки України
Національний технічний університет України «КПІ»
Інститут Прикладного Системного Аналізу*

Бідюк П.І., Романенко В.Д., Тимощук О.Л.

АНАЛІЗ ЧАСОВИХ РЯДІВ

(навчальний посібник)

Київ 2010

УДК 62-50

Бідюк П.І., Романенко В.Д., Тимошук О.Л.

Аналіз часових рядів (навчальний посібник)

— Київ: Політехніка, 2010. — 317 с.

ISBN 0-0000-0000-0

Книга присвячена дослідженню принципів математичного моделювання та прогнозування часових рядів. Докладно розглянуто теорію стаціонарних і нестаціонарних процесів, методику моделювання процесів довільної природи і структури, представлених часовими рядами, а також методи знаходження повних розв'язків рівнянь авторегресії з ковзним середнім (APKC). Показано застосування розв'язків рівнянь до асимптотичного аналізу поведінки процесів та коротко- і середньострокового прогнозування динаміки їх розвитку. Значна увага присвячена аналізу нестаціонарних процесів з детермінованими та стохастичними трендами, моделюванню коінтегрованих та гетероскедастичних процесів. Окремий розділ присвячено оптимальному оцінюванню станів динамічних систем за допомогою дискретних моделей. Докладно розглянуто сучасні методи моделювання та прогнозування нелінійних процесів на основі методів подібних траєкторій і опорних векторів. Розглянуто можливості застосування ймовірінісного прогнозування на основі байєсових методів.

Виконано аналіз таких популярних методів оцінювання параметрів лінійних та нелінійних моделей як рекурсивний метод найменших квадратів, метод максимальної правдоподібності та Монте Карло для марковських ланцюгів.

Книга рекомендується як навчальний посібник для студентів, аспірантів та викладачів, а також інженерів що спеціалізуються в галузі розв'язку проблем математичного моделювання і прогнозування динаміки технічних, фінансово-економічних процесів та процесів іншої природи, представлених часовими рядами.

Ухвалено до видання Вчену Радою Інституту прикладного системного аналізу Національної академії наук та Міністерства освіти і науки України

Рецензенти:

- Азарков В.М., директор інституту аерокосмічних систем управління, завідувач кафедри систем керування Національного авіаційного університету, Лауреат державної премії в галузі науки і техніки, доктор технічних наук, професор;

- Положаєнко С.А., завідувач кафедри комп'ютеризованих систем управління Одеського національного політехнічного університету, доктор технічних наук, професор.

Бідюк Петро Іванович

Романенко Віктор Демідович

Тимошук Оксана Леонідівна

ISBN 0-0000-0000-0

© Бідюк П.І., Романенко В.Д., Тимошук О.Л.

ЗМІСТ

Стор.

Вступ		7
Розділ 1	Мета моделювання, вимоги до даних і моделей	12
1.1	Мета побудови математичних моделей часових рядів	12
1.2	Поняття структури математичної моделі	14
1.3	Два основних методи побудови математичних моделей	16
1.4	Узагальнений алгоритм побудови моделі	18
1.5	Вимоги до експериментальних даних, оцінок параметрів та моделі ..	20
1.6	Спрощена класифікація математичних моделей	26
1.7	Деякі типи регресійних та різницевих рівнянь для опису ЧР	28
1.8	Запитання і вправи до першого розділу	32
Розділ 2	Аналіз стаціонарності процесів авторегресії з ковзним середнім	34
2.1	Слабка і сильна стаціональність	34
2.2	Необхідні і достатні умови стаціонарності процесу AP(1)	36
2.3	Умови стаціонарності часткового розв'язку рівняння АРКС	39
2.4	Узагальнення умов стаціонарності на процеси АРКС(p, q)	43
2.5	Теоретична автокореляційна функція	46
2.6	Обчислення кореляційних функцій за допомогою рівнянь Юла-Уокера	48
2.7	Часткова автокореляційна функція	53
2.8	Запитання і вправи до другого розділу	58
Розділ 3	Застосування різницевих рівнянь до моделювання часових рядів	60
3.1	Загальні відомості про різницеві рівняння	60
3.2	Ітераційний метод знаходження розв'язків різницевих рівнянь	65
3.3	Методика знаходження однорідних та часткових розв'язків	71
3.4	Приклади знаходження повних розв'язків	96
3.5	Запитання і вправи до третього розділу	105
Розділ 4	Методика побудови математичних моделей часових рядів	107
4.1	Аналіз процесу	108
4.2	Попередня обробка даних	109
4.3	Аналіз наявності нелінійностей	113
4.4	Формування інших елементів структури моделі	115
4.5	Оцінювання коефіцієнтів моделей-кандидатів	120
4.6	Діагностика моделі – вибір кращої з множини оцінених кандидатів	123
4.7	Приклади побудови математичних моделей за допомогою експериментальних даних	129
4.8	Запитання і вправи до четвертого розділу	143

Розділ 5	Методика прогнозування динаміки процесів за допомогою різницевих рівнянь	145
5.1	Для чого необхідні прогнози?	145
5.2	Які складові процесу можна прогнозувати?	147
5.3	Умовні та безумовні статистичні характеристики	149
5.4	Прогнозування без розв'язку рівнянь	152
5.5	Побудова функцій прогнозування на основі розв'язків РР	156
5.6	Прогнозування з мінімальною дисперсією	161
5.7	Функції прогнозування, отримані в даному розділі	173
5.8	Оцінювання якості прогнозів	175
5.9	Довірчий інтервал для прогнозу	182
5.10	Комбінування прогнозів, отриманих різними методами	184
5.11	Запитання і вправи до п'ятого розділу	187
Розділ 6	Оператори зсуву та їх застосування	189
6.1	Властивості операторів зсуву	189
6.2	Приклади використання операторів зсуву	191
6.3	Ще одна форма характеристичного рівняння	196
6.4	Розв'язування різницевого рівняння в прямому напрямі	197
6.5	Запитання і вправи до шостого розділу	199
Розділ 7	Оптимальне оцінювання і прогнозування за допомогою фільтра Калмана	201
7.1	Принцип рекурсивного оцінювання	201
7.2	Дискретний фільтр Калмана для вільних динамічних систем	203
7.3	Дискретний фільтр Калмана для лінійних систем з детермінованими і стохастичними входами	212
7.4	Причини розбіжності та заходи щодо підвищення обчислювальної стійкості оптимального фільтра	216
7.5	Приклади побудови оптимального фільтра	219
7.6	Оцінювання невимірюваних компонент вектора стану за допомогою оптимального фільтра	227
7.7	Функція прогнозування на основі оптимального фільтра	228
7.8	Запитання і вправи до сьомого розділу	234
Розділ 8	Представлення математичних моделей у просторі станів	237
8.1	Вступ	237
8.2	Перехід від лінійних неперервних до лінійних дискретних систем	242
8.3	Дискретні рівняння типу АРКС в просторі станів	244
8.4	Приклади представлення моделей у просторі станів	253
8.5	Запитання і вправи до розділу	260
Розділ 9	Методика прогнозування динаміки процесів на основі математичних моделей часових рядів	262
	3	

	різнатемповою дискретизацією	
9.1	Вступ	262
9.2	Розробка різнатемпової моделі авторегресії і ковзного середнього	268
9.3	Прогнозування одновимірних стаціонарних процесів на основі моделей авторегресії і ковзного середнього з різнатемповою дискретизацією	268
9.4	Прогнозування багатовимірних стаціонарних процесів з різнатемповою дискретизацією	273
9.5	Прогнозування нестаціонарних процесів на основі математичних моделей з різнатемповою дискретизацією	282
9.6	Прогнозування динаміки процесів у просторах змінних для моделей з різнатемповою дискретизацією	283
Література		294
Додаток А	Методи оцінювання параметрів математичних і статистичних моделей	300

В С Т У П

Прогнозування на основі моделей, побудованих за експериментальними (статистичними) даними – один із самих популярних підходів до прогнозування динаміки процесів у соціально-економічних, фінансових, технічних та інших системах, коротко- та середньострокового прогнозування об'ємів виробництва та накопичення продукції на складах, оцінювання альтернативних економічних стратегій, формування бюджетів підприємств та держави, прогнозування та менеджменту ризиків довільної природи та розв'язання інших задач.

Загалом методи прогнозування можна розділити на три широкі класи: (1) прогнозування на основі суджень, тобто, прогнозування, що ґрунтуються на суб'єктивних судженнях (оцінках), інтуїції, поглиблених знаннях конкретної області та іншій інформації, що має відношення до прогнозованого процесу; (2) методи прогнозування на основі використання часового ряду однієї змінної, тобто, на основі авторегресії, авторегресії з ковзним середнім (APKC), APKC плюс тренд і т. ін.; (3) методи прогнозування на основі використання часових рядів декількох змінних (векторні процеси). В останньому випадку ендогенна змінна, що прогнозується, залежить від декількох регресорів або екзогенних змінних у правій частині рівняння. Очевидно, що в загальному випадку процедура прогнозування може поєднувати в собі 2-3 наведених вище методи.

На сьогоднішній день в спеціальній літературі описано досить велика кількість методів прогнозування лінійних та нелінійних процесів на основі використання часових рядів. Найбільш поширеними серед них є метод групового врахування аргументів (МГВА), авторегресія (AP), APKC, авторегресія з інтегрованим ковзним середнім (APIKC), лінійна та нелінійна множинна регресія, квантильна регресія, регресійні дерева, нейромережі, байесівські мережі, нечіткі множини, нечіткі нейромережі та інші. Відносно „універсальними” методами моделювання та прогнозування є МГВА і нечіткі нейромережі. Однак, практика показує, що одного, навіть досить універсального методу, недостатньо для досягнення повноти аналізу процесу. Так, коректний аналіз гетероскедастичних процесів вимагає застосування моделей спеціальної структури для описання умовної дисперсії, що не забезпечує МГВА. Кожний метод має також свої недоліки і

переваги щодо обчислювальних витрат та характеристик точності прогнозу. Так, висока точність прогнозу за допомогою МГВА або нейромережі іноді досягається за рахунок високих обчислювальних витрат. Це особливо стосується застосування отриманої моделі в системі керування реального часу, де модель необхідна і для оцінювання прогнозу і для синтезу керуючого впливу. Суттєвий виграв щодо обчислювальних витрат можна досягти в такому випадку за допомогою набагато простішої моделі АРКС, перевагами якої є простота структури та можливості її оперативної адаптації до характеристик процесу в реальному часі.

Однак, на сьогодні відсутній систематизований підхід до вибору структури математичних моделей та методів прогнозування, а також рекомендацій щодо їх застосування. Під структурою моделі надалі будемо розуміти сукупність таких елементів: (1) вимірність моделі – число рівнянь, що описують процес; (2) порядок рівнянь; (3) нелінійність та її тип (відносно змінних або параметрів); (4) час запізнення по входу та його оцінка; (5) збурення та його тип.

Метою даної роботи є (1) аналіз та систематизація деяких існуючих методів до прогнозування стаціонарних та нестаціонарних нелінійних процесів, представлених часовими рядами і часовими перерізами даних; (2) представлення нових методів прогнозування лінійних та нелінійних нестаціонарних процесів; (3) вироблення рекомендацій щодо комплексного застосування методів прогнозування. Розглядаються спеціальні структури моделей гетероскедастичних та коінтегрованих процесів, які забезпечують отримання високоякісних оцінок прогнозів нестаціонарних процесів із змінною дисперсією та трендами різних порядків.

Структура книги ґрунтується на системному підході до аналізу процесів різної природи, який передбачає врахування в моделях можливих невизначеностей описання процесів, зокрема це невизначеності структурного і статистичного характеру. Структурні невизначеності зумовлені значною нестаціональністю багатьох реальних процесів, яка призводить до швидкої зміни елементів структури (складових) процесу, що вимагає адаптації описання структури моделі до процесу. Невизначеності статистичного характеру зумовлені недостатньою інформацією щодо випадкових збурень. Як правило, такі невизначеності вимагають застосування до їх описання розподілів ймовірностей різних типів, хоча в більшості випадків ми достатньо обґрутовано вважаємо збурення нормальним процесом.

Також відомо, що в більшості випадків моделювання та прогнозування ми будуємо оптимальні за параметрами математичні моделі та оцінюємо моделі і прогнози за множиною деяких статистичних критеріїв. При цьому оптимізація стосується не тільки безпосередньо параметрів (коефіцієнтів) моделі, але й статистичних параметрів якості моделі. Тобто виникає

необхідність багатокритеріальної оптимізації всього процесу побудови моделі та оцінювання прогнозу. Така задача частково розглядається в даній роботі, хоча в явному вигляді на сьогодні розв'язок таких задач є досить проблематичним.

Статистичне моделювання і прогнозування ґрунтуються на двох основних типах експериментальних (статистичних) даних: часові ряди і часові перерізи.

Часовий ряд – це множина рівновіддалених (рідко різновіддалених) в часі спостережень (вимірів), які характеризують поведінку процесу чи об'єкта на вибраному часовому інтервалі.

Іншим різновидом даних є *часовий переріз (panel data)*.

Часовий переріз – це множина спостережень змінних досліджуваної системи (процесу) на вибраний конкретний момент часу. Наприклад, дані, які характеризують стан всіх галузей української економіки на кінець грудня 2007 року.

Часові ряди є основою для аналізу поведінки технічних, технологічних, біологічних, соціально-економічних, фінансових та систем інших видів. Якщо на протязі місяця (тридцяти днів) ми будемо вимірювати о шостій ранку та о шостій вечора *температуру свого тіла, частоту пульсу, артеріальний тиск та поверхневий опір шкіри*, то отримаємо в результаті чотири часових ряди, які характеризують наш поточний фізичний стан. Кожний ряд можна описати своєю математичною моделлю, яка може бути використана для прогнозу конкретної змінної. Можна побудувати одну чотиривимірну модель, яка дозволить прогнозувати всі змінні одночасно. Іншими прикладами часових рядів є значення індексу інфляції, зареєстровані щомісячно на протязі кількох років, або щорічні значення валового внутрішнього продукту (ВВП), зареєстровані на протязі 30 років. Проміжок часу між вимірами називається періодом дискретизації вимірів (часто позначають через T_s , де s – перша буква англійського слова *sampling* – дискретизація).

Часові ряди у фінансово-економічних системах є, як правило, обмеженими за розміром (кілька десятків або сотень значень), а ряди, які характеризують поведінку технічних систем і технологічних процесів, мають практично нескінченну довжину. Наприклад, в сучасних автомобілях встановлюють систему автоматичного керування швидкістю руху, яка підтримує постійну швидкість без втручання водія. Швидкість руху постійно вимірюється тахометром, перетворюється в цифрову форму і поступає на елемент порівняння із заданим значенням швидкості. Відхилення поточного значення швидкості від заданого використовується для зміни кута нахилу заслінки в карбюраторі або регулювання системи вприскування палива, яка регулює подачу горючої суміші в циліндри. Якщо прийняти період дискретизації вимірів, T_s , змінних в системі керування швидкістю $T_s = 20$

мс, то за 1 секунду отримаємо 50 вимірів, за 1 хвилину – 3000 вимірів, а за 1 годину – 180000. Таким чином, в даному випадку ми маємо справу з надзвичайно довгими часовими рядами значень швидкості руху та відхилень від завдання.

Період дискретизації повинен бути строго постійним, оскільки його величина впливає на характеристики часового ряду. Чим більшим є період дискретизації, тим біdnішим, з точки зору частотного наповнення, буде ряд експериментальних даних. В техніці період дискретизації вибирають із умови (теорема Котельникова-Шеннона, яка визначає необхідну умову вибору T_s):

$$T_s \leq \frac{1}{2f_{\max}},$$

де f_{\max} – максимальна частота гармонічної складової сигналу, яка необхідна для подальшого дослідження процесу чи об'єкта. Вона вибирається дослідником з фізичних міркувань та особливостей функціонування процесу. Визначити частотний склад сигналу (представленого, наприклад, часовим рядом) можна за допомогою перетворення Фур'є або іншого зручного ортогонального перетворення даних.

Необхідно зазначити, що при невеликих значеннях періоду дискретизації може з'явитись значна лінійна залежність між вимірами, що призводить до появи лінійно залежних стовпчиків вимірів у відповідній матриці, що використовується в методі оцінювання, наприклад, в методі найменших квадратів (МНК), який реалізується за виразом:

$$\hat{\theta} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y},$$

де $\hat{\theta}$ – вектор оцінок параметрів математичної моделі об'єкта; \mathbf{X} – матриця вимірів незалежних змінних (регресорів або екзогенних змінних); \mathbf{y} – вектор вимірів залежної (ендогенної) змінної. Це приводить до неможливості оцінювання параметрів (коєфіцієнтів) рівнянь на основі рядів даних. Тому вибір значення T_s завжди є компромісним – тобто, не занадто великим щоб не втратити інформативність сигналу (його частотне наповнення) і не занадто малим щоб не з'явилася значна лінійна залежність між вимірами. В технічних системах період дискретизації може складати від кількох мікросекунд до десятків хвилин.

При дослідженні соціально-економічних та фінансових процесів період дискретизації може складати кілька хвилин (ціни на біржі), один день, тиждень, місяць, квартал або рік. При цьому можливості змінювати T_s в сторону зменшення в таких системах дуже обмежені внаслідок особливостей протікання цих процесів та методів збору даних. Незважаючи

на те, що масштаби часу, в яких протікають процеси в технічних та соціально-економічних системах суттєво відрізняються, методи аналізу відповідних часових рядів практично залишаються майже однаковими. Про це свідчать численні приклади із спеціальної літератури.

Економетричні часові ряди містять ряд компонент, математичне описання яких буде докладно розглянуті в даній книзі, зокрема це такі:

- детерміновані і випадкові коливання;
- детерміновані і випадкові тренди;
- сезонні ефекти (складові);
- нелінійності відносно змінних та параметрів;
- значні випадкові імпульсні значення (викиди);
- умовна дисперсія процесу.

Дуже важливо визначитись перед побудовою математичної моделі процесу, які компоненти будуть описані та спрогнозовані за моделлю. Докладно послідовність побудови моделей процесів різної природи розглядається в другому розділі, а методика побудови функцій прогнозування – в четвертому розділі. В подальших розділах розглядаються особливості моделювання нестационарних та нелінійних процесів, а також деякі методи їх оцінювання.

Зокрема, в деяких розділах розглядаються можливості застосування нових методів моделювання та прогнозування, наприклад, методи на основі аналізу подібних траєкторій та опорних векторів. Дані методи дають можливість отримати високоякісні прогнози для лінійних та нелінійних процесів. Актуальною є задача оцінювання ймовірнісних прогнозів, потреба в яких швидко зростає як у суто інженерних застосуваннях, так і у фінансово-економічних системах. Так, Байесові мережі довіри представляють собою ефективний інструмент описання та прогнозування процесів із структурними та параметричними невизначеностями різних типів. Вони можуть бути застосовані також до вирішення проблем передбачення при відповідній постановці задачі.

Автори сподіваються, що книга в цілому буде корисною для студентів, аспірантів, інженерів, викладачів та всім іншим, хто займається практичними задачами математичного моделювання та прогнозування на основі експериментальних і статистичних даних в різних галузях техніки, економіки, фінансів, екології, соціальних дослідженнях та ін.

Розділ 1

МЕТА МОДЕЛЮВАННЯ, ВИМОГИ ДО ДАНИХ І МОДЕЛЕЙ

Розділ присвячено аналізу принципів побудови математичних моделей процесів різної природи, представлених експериментальними або статистичними даними у вигляді часових рядів. Під часовим рядом даних будемо розуміти сукупність вимірів змінної, отриманих з однаковим (рідко неоднаковим) часовим інтервалом, який називається періодом дискретизації вимірів.

Розглядається мета побудови моделі та елементи структури моделі, які необхідно визначити в результаті аналізу процесу або об'єкта, що описується математичною моделлю. Наведено аналіз структурного і функціонального методів побудови моделей; перший представляє собою теоретичний метод на основі аналізу законів та закономірностей функціонування процесу, а другий орієнтований на збір та використання експериментальних даних. Узагальнений алгоритм побудови моделі за експериментальними даними дає уявлення про загальні етапи моделювання, які можна застосовувати до процесів ріної природи. Також сформульовано вимоги до експериментальних даних, оцінок параметрів та моделі в цілому. Дотримання цих вимог дає можливість будувати моделі високого ступеня адекватності та скорочувати витрати часу на побудову та аналіз якості моделі. Розділ завершується переліком деяких типів математичних моделей, що описують стаціонарні і нестаціонарні процеси, а також процеси з нелінійностями відносно змінних та параметрів.

1.1 Мета побудови математичних моделей часових рядів

Можна виділити три цілі побудови математичних моделей процесів та об'єктів:

- *поглиблене вивчення* процесів;
- *прогнозування значень* змінних за допомогою моделі;

- використання моделі для створення (*синтезу*) систем керування.

Поглиблене вивчення процесів за допомогою математичних моделей дозволяє дослідити кількісні зв'язки між входними та вихідними змінними, дослідити яким чином змінюються вихідні змінні при варіації входних в широкому діапазоні та розглянути поведінку процесів на будь-яких часових інтервалах у прийнятному масштабі часу. Математична модель, що будується для цієї мети, може бути дуже складною і трудомісткою, оскільки вона повинна враховувати тонкощі взаємодії кількісних і якісних змінних із можливим врахуванням реального часу, тобто в даному випадку часто використовують імітаційне моделювання. За допомогою математичних моделей можна виявити ефекти і явища, які недоступні безпосереднім спостереженням за допомогою приладів. Крім того, при проектуванні нових систем в різних галузях можна швидко змінювати варіанти реалізації системи завдяки можливості її швидкого дослідження на моделі, виявити вплив початкових умов та обмежень на ключові змінні.

Прогнозування значень змінних виконується, як правило, на основі набагато простіших моделей ніж поглиблене вивчення процесів. Для цієї мети дуже зручними є дискретні моделі у вигляді авторегресійних рівнянь (AP) та авторегресії з ковзним середнім (APКС). Часто використовують також стандартизоване представлення в просторі станів неперервних чи дискретних моделей. Тип моделі залежить також від того, який прогноз необхідний – короткостроковий, середньостроковий чи довгостроковий. Задача прогнозування може бути дуже складною, якщо стохастичний процес нелінійний та нестационарний.

Для розв'язання задачі *синтезу систем керування процесами* використовують різні класи моделей в неперервному та дискретному часі, які представляють, як правило, в уніфікованій формі простору станів. В даному випадку моделі не повинні мати дуже високу точність (адекватність), тому що вони працюють в замкненому контурі з від'ємним зворотним зв'язком. Це сприяє асимптотичному зменшенню похибок керування до мінімально можливих значень навіть при використанні грубих моделей.

Моделювання системи керування на комп'ютері дозволяє за мінімальний час дослідити можливість застосування різних варіантів законів керування, комбінацій керуючих дій та визначити ефективність системи керування при зміні входних керуючих сигналів в широкому діапазоні їх значень. Комп'ютерна система керування може працювати як в автоматичному режимі (при керуванні технічними або технологічними процесами в реальному часі), так і в режимі порадника при керуванні економічними, екологічними, соціальними чи іншими складними

процесами. При використанні комп'ютерної системи як порадника, керуючі дії, згенеровані комп'ютером, порівнюються з експертними оцінками оцінками входів і на основі деякого критерію якості вибираються найбільш прийнятні варіанти керуючих впливів.

1.2 Поняття структури математичної моделі

Введемо поняття структури математичної моделі, яке будемо використовувати надалі. Поняття структури моделі містить у собі наступні елементи (параметри):

Порядок моделі, тобто порядок диференціального, різницевого чи іншого рівняння, що використовується для опису динаміки процеса чи об'єкта. Наприклад, стохастичне різницеве авторегресійне (AP) рівняння другого порядку має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + \varepsilon(k).$$

Тобто, порядок цього різницевого рівняння визначається числом затриманих у часі значень зінної, що використовуються у правій частині рівняння для опису змінної в лівій частині. Стохастичним воно називається тому, що у правій частині присутня випадкова змінна $\varepsilon(k)$, призначення якої ми пояснимо трохи нижче.

2. Вимірність моделі. Вона визначається числом рівнянь, що використовуються для математичного описання об'єкта чи процесу. Процес, котрий описують одним рівнянням, називають одновимірним чи скалярним. Процес, котрий описують двома і більше рівняннями, називають багатовимірним. Хоча більшість процесів у природі є багатовимірними, часто обмежуються обновимірними моделями, виходячи з їх простоти та зручності застосування.

3. Наяvnість нелінійностей i їх характер. Визначити наявність нелінійностей - не завжди проста задача. Так, для механічних і деяких інших систем наявність нелінійностей можна визначити шляхом вивчення законів, закономірностей і особливостей їхнього функціонування. Наприклад, відомо, що для механічних систем характерною є наявність нелінійностей типу «люфт», «тертя», гістерезис.

При побудові регресійних моделей частіше зустрічаються нелінійності відносно змінних і нелінійності відносно параметрів. Прикладом нелінійності відносно змінних може бути поліноміальна регресія виду

$$y(k) = a_0 + a_1 x(k) + a_2 x^2(k) + a_3 x^3(k) + \varepsilon(k).$$

Коефіцієнти цього рівняння можна оцінювати звичайним методом найменших квадратів (МНК) при належній побудові матриці вимірювань (вона буде розглянута нижче). Нелінійність відносно параметрів зумовлена наявністю в моделі добутків коефіцієнтів, наприклад, у виглядді

$$y(k) = a_0 + a_1 a_2 x(k) + a_2 \exp(-bx(k)) + \varepsilon(k).$$

Коефіцієнти (параметри) такої моделі неможливо оцінити за допомогою звичайного МНК, тому для розв'язання цієї задачі використовують нелінійний МНК, метод максимальної правдоподібності чи інші методи нелінійного оцінювання.

4. Час запізнення реакції на виході об'єкта відносно вхідного сигналу. Запізнення по входу (лаг) досить легко враховується як у неперервних, так і в дискретних моделях. Для моделі з дискретними змінними у вигляді різницевого рівняння

$$y(k) = a_0 + a_1 y(k-1) + a_2 x(k-d) + \varepsilon(k)$$

час запізнення d представляє ціле число, що дорівнює числу періодів дискретизації вимірювань, на яке запізнюються вихідний сигнал щодо вхідного. Тривалість періоду дискретизації вимірювань залежить від динаміки конкретного процесу і може змінюватися в межах від декількох мікросекунд у фізико-технічних системах до одного року в макроекономіці.

Розглянемо модель неперервного процесу у вигляді передаточної функції із запізненням:

$$W(p) = \frac{K e^{-p\tau}}{1 + Tp},$$

де K – статичний коефіцієнт передачі об'єкта; p – змінна Лапласа; T – постійний часу; τ – час запізнення по входу. Запізнення у дискретній формі d і запізнення в неперервній формі τ зв'язані між собою наступним чином:

$$\hat{d} = \text{int}(\tau/T_s);$$

тут \hat{d} – оцінка часу запізнення в дискретній формі ($\hat{d} = 0, 1, 2, \dots$).

5. Тип збурень, що діють на процес, і спосіб їх врахування. Під збуреннями розуміють вхідні впливи процесу, які створюють, як правило, негативні умови для його протікання і не використовуються з тих чи інших причин як керуючі. Збурення поділяють на детерміновані і стохастичні, а

враховуються вони в адитивній чи мультиплікативній формі. Вище ми привели різницеві рівняння, у яких збурення $\varepsilon(k)$ входить в адитивній формі. Приклад мультиплікативної форми:

$$h(k) = \psi(k)[\alpha_0 + \alpha_1 h(k-1)],$$

де $\psi(k)$ – мультиплікативне збурювання.

Частіше всього збурення описують розподілами випадкових величин (статистичні моделі), але в окремих випадках його можна виміряти і описати функціонально (математичні моделі збурень). Наприклад, можна виміряти температуру навколошнього середовища, яка впливає на протікання реакції у хімічному реакторі і побудувати відповідну функціональну залежність температури від часу.

Вибір структури моделі, що адекватна процесу, – задача не проста і вирішується, як правило, ітераційно. Спочатку структуру моделі оцінюють наближено на підставі дослідження закономірностей протікання процесу, аналізу кореляційних функцій, візуального аналізу даних. При цьому вибирають декілька найбільш ймовірних структур (кандидатів). Потім обчислюють оцінки параметрів моделей-кандидатів і вибирають кращу з них, використовуючи відповідні статистичні характеристики якості моделей.

Якщо жодна з моделей-кандидатів не може вважатися адекватною для конкретного застосування, то необхідно досліджувати на інформативність експериментальні дані, які можуть бути недостатньо інформативними для оцінювання моделі. У такому випадку може знадобитися повторний чи додатковий збір експериментальних даних і корегування структури моделі.

1.3 Два основних методи побудови математичних моделей

Основними методами побудови математичних моделей є

- структурний;
- функціональний.

Структурний метод передбачає моделювання внутрішнього механізму взаємодії змінних, відображає їх фактичні взаємозв'язки.

Критерієм правильності структурної моделі є одинаковий характер поведінки основних змінних реального процесу і моделі.

Розглянемо, наприклад, зростання інфляції внаслідок випуску

додаткової грошової маси. Оскільки логіка цього процесу досить проста і існують експериментальні (статистичні) дані, які ілюструють зростання інфляції, то можна постулювати, що інфляція описується диференціальним або різницевим рівнянням першого/другого порядку (рис.1.1).

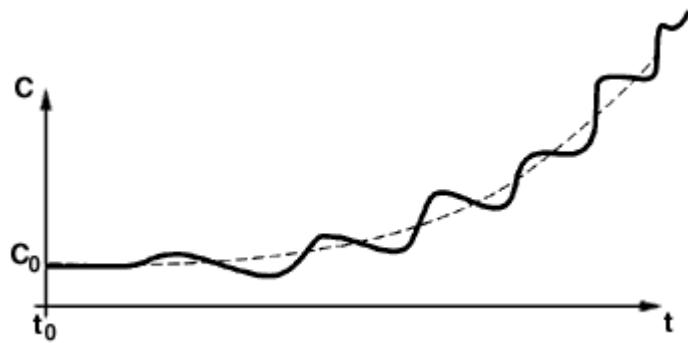


Рис.1.1. Можлива крива зростання інфляції.

Структурним підходом можна скористатись, наприклад, для побудови математичної моделі процесу трансформування власності (модель приватизації наведена в додатку) або макроекономіки в цілому. Для цього необхідно визначити вхідні керуючі змінні, збурення та вихідні змінні, а також визначити якого типу зв'язки існують між ними (рис.1.2).



Рис.1.2. Спрощене зображення макроекономічного процесу.

Серед керуючих змінних макроекономічного процесу можна виділити внутрішні та зовнішні інвестиційні потоки, потоки сировини, робочої сили, нові технології та структурні зміни в промисловості в цілому, а також в окремих галузях. Метою використання керуючих змінних є досягнення заданих рівнів макроекономічних показників – рівень виробництва валового внутрішнього продукту (ВВП), індекси інфляції, індекс людського розвитку, середня заробітна плата і т.і.

Як правило, в такі моделі вводять в явному вигляді збурення – випадкові змінні, які негативно впливають на протікання процесу. Так, у випадку створення моделі макроекономіки, збуреннями можуть бути

- помилкові рішення уряду;
- затримка платежів між підприємствами та державами;

- значні коливання цін на енергоносії;
- хронічна технологічна відсталість;
- швидкі зміни податкового законодавства;
- відлив капіталу за кордон;
- використання недостовірної статистичної інформації.

Очевидно, що врахувати подібну інформацію в моделі надзвичайно складно, а тому випадкові змінні агрегують (об'єднують) і представляють в моделі однією чи двома випадковими змінними, які представляють всі збурення.

На основі знання логіки взаємодії змінних процесу та використання відомих макроекономічних законів (наприклад, врівноваженого розвитку процесів) будується система рівнянь, які описують розвиток окремої галузі або макроекономіки в цілому.

Функціональний підхід використовують для формального описання процесу, не проникаючи глибоко у фактичну структуру цього процесу і взаємодії його змінних.

В даному випадку процес = “*чорний ящик*”

Наприклад, для побудови моделі ціноутворення можна скористатись такими вхідними змінними: I – об’єм імпорту; D – об’єм грошової маси в обороті; P – місячний об’єм виробництва. За вихідну змінну можна взяти індекс споживчих цін C , тобто,

$$C = f(I, D, P).$$

Для побудови цієї моделі необхідно мати чотири часових ряди, які необхідні для обчислення оцінок коефіцієнтів моделі.

Очевидно, що функціональний підхід є простішим від структурного і саме він найчастіше використовується на практиці. Гнучкість даного підходу дає можливість відносно швидко побудувати високоякісні моделі для прогнозування та синтезу систем керування.

1.4 Узагальнений алгоритм побудови моделі

Розглянемо узагальнений алгоритм побудови математичної моделі на основі експериментальних даних у вигляді наступних кроків. Він є узагальненим з точки зору застосування до систем чи процесів практично будь-якого типу.

1. Визначення мети побудови моделі, попереднє вивчення процесу (об'єкта).

На цьому етапі визначається мета побудови моделі, тобто чи буде модель використовуватись для поглибленого вивчення процесу, прогнозування його стану чи керування. Виконується аналіз функціонування процесу на основі літературних джерел та (можливо) експериментальних даних при їх наявності з метою встановлення *числа входів і виходів, логіки взаємодії складових частин процесу, визначення можливих зовнішніх збурень та їх типу*: детерміновані чи випадкові. По можливості необхідно встановити розподіл ймовірностей для випадкових збурень або функціональне описання для детермінованих збурень. Якщо існують моделі подібних процесів, їх також необхідно досконально вивчити та врахувати можливі недоліки.

2. Попередня оцінка структури моделі. На основі вивчення процесу, виконаного на першому етапі, необхідно встановити типи структур моделей-кандидатів. Їх може бути декілька в залежності від того наскільки невизначеною є інформація відносно процесу. Чим більшою є невизначеність, тим більше структур моделей необхідно досліджувати в процесі побудови адекватної моделі.

3. Планування експерименту та підготовка до його виконання. На цьому етапі виконуються наступні дії щодо планування виконання експерименту з метою отримання експериментальних даних:

- визначаються діапазони зміни вхідних та вихідних величин, збурень;
- встановлюється дискретність зміни вхідних величин, період дискретизації вимірювань (якщо змінні неперервні), визначаються типи вимірювальних та реєструючих приладів;
- плануються режими роботи процесу, для яких необхідно зібрати експериментальні дані;
 - якщо дані носять статистичний характер, то визначається періодичність їх збору та занесення в базу даних;
 - задається тип, об'єм і якість продукції, що буде вироблена на протязі експерименту, а також визначаються необхідні об'єми сировини та енергії.

4. Виконання експерименту та формування бази даних. На цьому етапі реалізується розроблений на третьому етапі план експерименту і формуються часові ряди з вимірювань (чи статистичних даних), які будуть використані для обчислювання оцінок параметрів математичних моделей.

5. Обчислення оцінок параметрів (коєфіцієнтів) математичних моделей на основі експериментальних даних. При цьому оцінюють

параметри для всіх моделей-кандидатів, вибраних на другому етапі. Для виконання цієї задачі необхідно:

- вибрати метод оцінювання параметрів моделі в залежності від її структури;
- зробити попередню обробку даних; в залежності від конкретної задачі це може бути масштабування, логарифмування, цифрова фільтрація, видалення недостовірних даних і т.п.;
- обчислити оцінки (векторів) параметрів моделей.

6. *Визначити ступінь адекватності кожної моделі-кандидата* процесу за допомогою статистичних критеріїв. Визначити кращу модель з множини кандидатів.

7. Якщо побудована модель відповідає висунутим вимогам (заточністю прогнозу чи якістю керування), то завершити процедуру; інакше перейти на 8-й крок.

8. *Уточнити структуру моделі*, зібрати, при необхідності, додаткові експериментальні дані і перейти на крок 5.

Хоча планування та виконання експерименту для соціально-економічних та фінансових систем є досить складною задачею, в окремих випадках це цілком можливо, особливо, якщо підприємство має наміри впровадити нові інформаційні технології обробки даних і методи прогнозування розвитку процесів на виробництві. Наприклад, цілком можливо спланувати та провести інвестиційний експеримент, експерименти з новими технологіями, новими типами продукції.

1.5 Вимоги до експериментальних даних, оцінок параметрів та моделі

1.5.1 Вимоги до експериментальних даних

1. Вимога *неперервності* та *синхронності* даних. Експериментальні дані повинні вимірюватись та реєструватись через однакові проміжки часу (період дискретизації вимірювань T_s). Цю вимогу необхідно виконувати для процесів будь-якого типу – технічних, економічних, екологічних і т.д. Порушення цієї вимоги призводить до зміни спектрального складу вимірювального сигналу, що недопустимо, оскільки при цьому змінюється інформативність сигналу. Крім того, вимірювання вхідних та вихідних сигналів необхідно робити синхронно, тобто в одній ті

ж моменти часу. В протилежному випадку вони будуть непридатні для побудови передаточних функцій, оскільки порушуються причинно-наслідкові зв'язки між входами та виходами. Як правило, в системах керування реального часу задача збору вимірювальних даних має найвищий пріоритет.

2. Вибірка даних повинна бути представницькою. Це означає, що вона повинна охоплювати досить довгий період часу, щоб включити в розгляд всі режими роботи, які передбачається описати моделлю. Розрізняють два основних режими роботи процесів: *перехідний та усталений*. В перехідному режимі система керування переводить процес з деякого початкового стану в заданий. Перебування процесу в заданому (номінальному) стані на протязі деякого відносно довгого проміжку часу називають усталеним режимом роботи.

Прикладом широко відомого перехідного процесу є процес нагрівання до кипіння вмісту каструлі на кухонній плиті. На початку цього процесу ми задаємо режим максимальних витрат енергії щоб скоротити тривалість процесу нагрівання. Після досягнення режиму кипіння витрати енергії можна суттєво скоротити і процес переходить в усталений режим “повільного” кипіння. Подачу великої кількості енергії в початковий момент часу можна порівняти з подачею на вхід об'єкта сигналу у вигляді сходинки, а зміну температури вмісту каструлі можна вважати за перехідну характеристику цього процесу. Очевидно, що безліч прикладів такого типу можна знайти в промисловості. На сьогодні в нашому суспільстві спостерігається перехідний процес від соціалістичного ладу, який ґрунтувався на суспільній власності на засоби виробництва, до капіталістичного з приватною власністю.

3. Вибірка вимірювальних даних повинна бути інформативною. Частіше всього інформативність пов'язують з числом похідних, що їх містить вимірювальний сигнал. Чим більше число похідних можна отримати з вимірювань, тим інформативнішим є сигнал. Наприклад, припустимо, що процес описується диференціальним рівнянням другого порядку:

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 = bu(t),$$

де $y(t)$ – вихідний сигнал процесу; $u(t)$ – вхідний сигнал; $\theta^T = [a_0 \ a_1 \ a_2]$ – вектор коефіцієнтів рівняння, які необхідно оцінити за допомогою експериментальних даних. Очевидно, що оцінки коефіцієнтів a_2 і a_1 можна

обчислити тільки в тому випадку, якщо виміри $y(t)$ містять другу і першу похідну по часі.

Іноді інформативність визначають *величиною дисперсії* сигналу, тобто, чим більшою є дисперсія, тим вища інформативність сигналу. Так, константа має нульову дисперсію і, відповідно, мінімальну інформативність.

Вимога інформативності виконується в тому випадку, коли вхідний сигнал задовольняє умові *достатнього збудження* процесу. Основна ідея достатнього рівня збудження полягає в тому, щоб смуга частот вхідного сигналу перекривала амплітудно-частотну характеристику процесу. Тобто, вихідний сигнал $y(t)$ буде інформативним в тому випадку, коли достатньо інформативним буде вхідний сигнал $u(t)$. Умові інформативності (достатнього збудження) задовольняють такі основні вхідні сигнали: *білий шум, псевдовипадковий двійковий сигнал та одиничний імпульс*. Білий шум (гаусів процес) теоретично має нескінчений частотний спектр; достатньо широкі спектри мають і два інші сигнали.

З одного боку, для збудження процесу на його вхід необхідно подавати інформативний сигнал типу білого шуму, а з іншого – такий сигнал може бути недопустимим з точки зору фізики функціонування процесу (подача на вхід подібного сигналу може привести до руйнування процесу чи до створення аварійної ситуації). Тому в таких системах часто використовують як збуджуючий сигнал завдання регулятору, якщо він має форму сходинки, тобто має фронт прямокутного імпульсу. В багатьох випадках можливе використання гармонічних сигналів, які сприймаються "легше" більшістю об'єктів ніж білий шум або одиничний імпульс. Так, при дослідженні механічних систем часто використовують одиничні імпульси, гармонічні збуджуючі сигнали та їх комбінації, а при дослідженні технологічних процесів до вхідного сигналу керування додають 10-15% білого шуму, який забезпечує достатній рівень „збудження” процесу.

1.5.2 Вимоги до оцінок параметрів моделі

Точність оцінок параметрів моделі залежить від якості вимірювальних даних, коректності попередньої обробки даних та від того, наскільки правильно вибрано метод оцінювання. Так, для оцінювання параметрів лінійних та псевдолінійних (нелінійних відносно змінних) моделей можна застосовувати звичайний МНК та його модифікації, а для оцінювання моделей, нелінійних відносно параметрів, необхідно застосовувати нелінійний МНК, метод максимальної правдоподібності та інші методи, розроблені для оцінювання параметрів нелінійних моделей.

Існують такі стандартні вимоги до оцінок параметрів математичних моделей:

1. Оцінки повинні бути *незміщеними*. Це означає, що оцінки параметрів не повинні містити систематичної похибки, яка збільшує або зменшує оцінки параметрів на всіх вибірках даних або на різних відрізках однієї вибірки. Формально незміщеність оцінок параметрів записують так:

$$E[\hat{\theta}] = \theta,$$

де E – символ математичного сподівання; $\hat{\theta}$ – вектор оцінок параметрів; θ – істинне значення вектора параметрів.

2. Оцінки повинні бути *консистентними*, тобто оцінка $\hat{\theta}$ вектора параметрів повинна наблизатись до свого істинного значення θ по мірі збільшення об'єму вибірки даних. Оскільки оцінка $\hat{\theta}$ – це випадкова величина, то наближення до істинного значення можливе тільки в ймовірнісному смыслі. Консистентна оцінка повинна задовольняти наступному співвідношенню:

$$p(|\hat{\theta}_k - \theta| < \varepsilon) \rightarrow 1 \text{ при } k \rightarrow \infty,$$

де $\varepsilon > 0$ – мале число; p – символ ймовірності; $\hat{\theta}_k$ – оцінка вектора параметрів в момент k . Відомо, що довжина перехідного процесу при оцінюванні моделі залежить від кількості параметрів, що оцінюються, та вимірності моделі, а тому об'єми вибірок даних в усіх випадках повинні бути, по можливості, більшими. Проблеми з необхідними об'ємами даних виникають, як правило, при моделюванні економічних та соціальних систем; при побудові моделей технічних систем такі проблеми досить рідкісні.

3. Оцінки повинні бути *ефективними*, а це означає, що із множини допустимих, незміщених та консистентних оцінок необхідно вибрати ті, що є найближчими до оцінюваних параметрів, тобто ті, що мають найменші відхилення від середнього значення.

Іншими словами це вимога мінімальності дисперсії оцінки, яка формально записується так:

$$\operatorname{Var}(\hat{\theta}) \rightarrow \min.$$

Незміщені ефективні оцінки параметрів лінійної моделі можна отримати, наприклад, за допомогою методу найменших квадратів, якщо при оцінювання виконуються наступні умови:

– похибка моделі $e(k) = y(k) - \hat{y}(k)$ є центрованою величиною; де $y(k)$ – значення ряду, отримане експериментально (або статистичні дані); $\hat{y}(k)$ – оцінка змінної, отримана за допомогою побудованої моделі;

– похибка моделі – це некорельований процес, тобто, відсутня автокореляція похибок:

$$\text{cov}[e(k)] = E[e(k)e(k-l)] = \begin{cases} \sigma_e^2, & k = l, \\ 0, & k \neq l. \end{cases}$$

– похибка моделі некорельована із залежною змінною $y(k)$.

Корельованість похибки означає, що вона містить інформацію процес. Тобто необхідно коригувати структуру моделі таким чином, щоб похибка стала некорельованою.

1.5.3 Вимоги до математичної моделі

1. Модель повинна буди *адекватною* процесу чи об'єкту.

Адекватність означає, що модель повинна:

- а) відображати найбільш характерні зв'язки та взаємодію між змінними процесу;
- б) враховувати можливі керуючі дії (сигнали);
- в) враховувати вплив зовнішніх збурень та шуми вимірювань;
- г) враховувати початкові значення змінних та обмеження на них.

Формально адекватність визначають за допомогою ряду статистичних величин. Наприклад, дуже часто використовують середньо-квадратичну похибку моделі (СКП), яка обчислюється за формулою:

$$СКП(x_s, x_m) = \sqrt{\frac{1}{N} \sum_{k=1}^N [x_s(k) - x_m(k)]^2},$$

де $x_s(k)$ – вимір вихідного сигналу об'єкта в момент k ; $x_m(k)$ – оцінка вихідного сигналу об'єкта, отримана за допомогою оціненої моделі. Для лінійних моделей запропоновано кілька статистичних параметрів, що використовуються при оцінюванні адекватності, які будуть розглянуті нижче. Використання одного параметра для визначення ступеня адекватності моделі є некоректним підходом, оскільки *оценки параметрів – це випадкові величини*, а тому збільшення кількості критеріїв адекватності сприяє підвищенню ймовірності вибору адекватної (крашої) моделі.

2. Рівняння моделі повинні мати *розв'язок*, тобто бажано мати *аналітичний* або, якщо це неможливо, то *чисельний* розв'язок. Розв'язок необхідний для асимптотичного аналізу поведінки процесу (аналізу збіжності), а також для обчислення оцінок прогнозів.

Одним із принципів, яких необхідно дотримуватись при побудові моделі є наступний: “*в моделі не повинно бути нічого зайвого крім необхідного*”. Звичайно, що дотримуватись цього принципу досить непросто, і на практиці буває так, що модель дійсно має надзвичайно складну структуру, що також може бути оправдано необхідністю досягнення високого ступеня її адекватності процесу. Це особливо стосується нелінійних процесів. Але при побудові лінійних моделей у вигляді авторегресії чи авторегресії з ковзним середнім достатньо побудувати модель, статистичні характеристики якої співпадають з статистичними характеристиками часового ряду, на основі якого вона оцінюється. Такі спрощені моделі виявляються цілком придатними для прогнозування та керування процесами. Загалом питання складності моделі вирішується в кожному випадку окремо.

3. Модель повинна буди достатньо *універсальною*, щоб її можна було застосувати до описання класу однотипових процесів або до описання функціонування процесу в різних умовах.

Наприклад, для описання *моторної функції* людини (реакція на зовнішні збуджуючі сигнали) застосовують звичайне диференціальне рівняння другого порядку, яке представляють у вигляді функції передачі такого ж порядку:

$$W(s) = \frac{K e^{-\tau s}}{(1 - T_1 s)(1 - T_2 s)},$$

де K – статичний коефіцієнт передачі об’єкта; τ – час запізнення по входу, який в середньому дорівнює для людини 300-350 мс; T_1, T_2 – постійні часу. Така передаточна функція може використовуватись, наприклад, для описання реакції людини на зовнішні відео- або аудіосигнали, що поступають через систему візуального сприйняття чи аудіосистему (поширений приклад – водіння автомобіля або управління іншою машиною). Значення параметрів моделі можуть бути різними для різних людей, але структура моделі залишається незмінною. Таким чином, наведена модель описує широкий клас біологічних систем і цілком відповідає умові універсальності.

При моделюванні технічних систем широко застосовують ланки першого і другого порядку, що відповідають звичайним диференціальним рівнянням таких же порядків. На основі таких простих ланок можна побудувати моделі будь-якої складності. Дуже поширений в техніці та екології клас систем з розподіленими параметрами. Наприклад, процес розповсюдження домішок в атмосфері та водному середовищі, механічні коливання сонячних батарей та антен супутників, крила літака, локомотива з вагонами на залізниці, автомобіля з причепом і багато інших. Динаміку таких систем описують диференціальними рівняннями з частковими похідними.

4. Вимога *робастності* (*robust* – сильний, міцний). Робастність означає, що модель повинна давати прийнятний прогноз вихідної змінної не тільки на тому відрізку часового ряду, на основі якого вона побудована, але і на будь-якому іншому відрізку, що відповідає вибраному режиму роботи. Робастність може розглядатись також як стійкість моделі по відношенню до збурень, похибок та пропусків вимірювань. Вимога робастності є особливо критичною для систем, що працюють в реальному часі, оскільки нестійка модель може стати причиною створення аварійної ситуації.

5. Вимога *адаптивності*. Ця вимога означає, що хоча б частину параметрів моделі (щонайменше один) можна уточнювати по мірі надходження нових вимірювань від об'єкта. Ця вимога є обов'язковою при побудові моделей нестационарних систем, тобто систем, параметри яких є функціями часу. Системи керування, побудовані для нестационарних процесів, називають адаптивними. Такі системи є досить складними з точки зору аналізу збіжності оцінок параметрів та похибок керування, а тому при проектуванні адаптивних систем необхідно особливу увагу приділяти питанням достатнього збудження процесу, вибору методу оцінювання параметрів, збіжності оцінок параметрів, прогнозів та значень керуючих впливів.

1.6 Спрощена класифікація математичних моделей

В спеціальній літературі можна знайти класифікацію математичних моделей за різними критеріями, які приводять до визначення багатьох класів та підкласів моделей. Нижче розглянемо спрощену класифікацію, яка може бути корисною на практиці.

1. Описові (вербалльні) моделі

- Технічне завдання, звіт, логічна схема взаємодії змінних, яка супроводжується слвесним описом.

2. Математичні моделі

Вони будуються, наприклад, на основі рівнянь наступних типів:

- дифференціальні рівняння;
- різницеві рівняння;
- алгебраїчні рівняння.

Математичні моделі розділяють на два широких класи:

- a) *аналітичні* – моделі, які описують вибрані змінні процесу; як правило, такі моделі відтворюють один із аспектів функціонування процесу чи об'єкта, наприклад, динаміку валового внутрішнього продукту, моторну функцію людини чи рух автомобіля, потяга або літака;
- б) *імітаційні* – моделі, які докладно відтворюють поточне функціонування процесу у вибраному масштабі часу.

Імітаційні (IM) моделі нагадують активний фізичний експеримент з використанням фактичних даних, отриманих безпосередньо з процесу.

Переваги імітаційних моделей:

- наглядність результатів (проміжних та остаточних);
- динамічний характер відображення протікання процесу;
- можливість врахування детермінованих і випадкових факторів, а також складних залежностей від них;
- простота корекції моделі (ввести додаткове рівняння, правило і т.і.);
- можливість дослідити процес на множині його реалізацій (тобто, провести статистичний експеримент);
- практично необмежені можливості введення в модель таких елементів, як
 - рівняння будь-якого типу;
 - логічних правил;
 - нечіткої логіки;
 - статистичних розрахунків;
 - оптимізаційних процедур;
 - процедур прийняття рішень;
- об'єднання чітких алгоритмів з нечіткими (це непросто, але корисно при аналізі функціонування процесів з невизначеностями та прийнятті рішень).

Недоліки імітаційних моделей:

- висока вартість побудови та використання;
- вимагають багато часу для розробки.

1.7 Деякі типи регресійних та різницевих рівнянь (РР)

Авторегресія: рівняння авторегресії описує пам'ять процесу, тобто вплив значень попередніх станів на його поточний стан:

$$y(k) = a_0 + a_1 y(k-1) + \dots + a_p y(k-p) = a_0 + \sum_{i=1}^p a_i y(k-i) + \varepsilon(k),$$

де $a_i, i=1,\dots,p$ – коефіцієнти моделі, які оцінюються на основі значень часового ряду; p – порядок авторегресії, який визначається числом затриманих в часі значень ряду, що використовуються в правій частині рівняння для описання динаміки змінної в момент k ; $k = 1, 2, \dots$ дискретний час; $\varepsilon(k)$ – випадкова величина, поява якої зумовлена наступними причинами:

- вплив випадкових збурень на процес, що моделюється;
- похибки рівняння, зумовлені неточно вибраною структурою (можливо, що не враховано деякі регресори, введено непотрібні незалежні змінні або робиться спроба моделювати нелінійний процес за допомогою лінійного рівняння);
- методичні і обчислювальні похибки, які з'являються при обчисленні оцінок коефіцієнтів рівняння.

Парна регресія – вона включає в правій частині незалежну змінну (регресор):

$$y(k) = a_0 + a_1 x(k) + \varepsilon(k),$$

де $x(k)$ – *регресор* (незалежна або *екзогенна* змінна), тобто $x(k)$ має три назви. Залежну змінну $y(k)$ називають ще *ендогенною* змінною.

Множинна регресія – множинна регресія відображає вплив декількох незалежних змінних на залежну:

$$y(k) = a_0 + a_1 x_1(k) + a_2 x_2(k) + \dots + a_p x_p(k) + \varepsilon(k),$$

де $x_1(k), \dots, x_p(k)$ – регресори рівняння. Таке рівняння може включати також авторегресійну частину.

Авторегресія + множинна регресія = змішана регресія:

$$y(k) = a_0 + \sum_{i=1}^l a_i y(k-i) + b_1 x_1(k) + b_2 x_2(k) + \dots + b_p x_p(k) + \varepsilon(k).$$

Авторегресія з ковзним середнім порядку (p, q) (APKC(p, q)):

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j \varepsilon(k-j) + \varepsilon(k).$$

Регресія, нелінійна відносно змінних (псевдолінійна регресія):

$$y(k) = a_0 + a_1 x(k) + a_2 x^2(k) + \dots + a_p x^p(k) + \varepsilon(k).$$

Тобто, в даному випадку це поліноміальна регресія порядку p . Коефіцієнти псевдолінійної регресії оцінюються такими ж методами, що і лінійної, наприклад методом найменших квадратів (МНК) або максимальної правдоподібності (ММП).

Регресія, нелінійна відносно параметрів – Моделі, нелінійні відносно параметрів, містять адитивні члени, які включають в себе добутки параметрів моделей або інші види зв'язку (крім адитивного) між параметрами:

$$y(k) = a_0 + a_1 e^{bx(k)} + \varepsilon(k).$$

Для оцінювання параметрів таких моделей необхідно застосовувати методи нелінійного оцінювання – нелінійний метод найменших квадратів, метод максимальної правдоподібності, Монте Карло та інші.

Моделі гетероскедастичних процесів – тобто процесів, дисперсія яких змінюється в часі. Рівняння для умовної дисперсії (авторегресія першого порядку):

$$h(k) = \beta_0 + \varepsilon^2(k-1) + \varepsilon_1(k),$$

де $h(k)$ – умовна дисперсія процесу в момент k ; $\varepsilon^2(k)$ – квадрат залишків; $\varepsilon_1(k)$ – похибка моделі в момент k . Докладно моделі гетероскедастичних процесів будуть розглянуті нижче.

Авторегресійна умовно гетероскедастична модель порядку p (АРУГ(p)):

$$h(k) = \beta_0 + \sum_{i=1}^p \beta_i \varepsilon^2(k-i) + \varepsilon_1(k).$$

Узагальнена авторегресійна умовно гетероскедастична модель (УАРУГ(p, q)):

$$h(k) = \beta_0 + \sum_{i=1}^p \beta_i \varepsilon^2(k-i) + \sum_{i=1}^q \alpha_i h(k-i) + \varepsilon_1(k),$$

де $\alpha, \beta, \gamma \geq 0$ (щоб уникнути появи від'ємних значень умовних дисперсій).

Експоненціальна модель УАРУГ (умовна дисперсія як асиметрична функція ε , тобто моделювання впливу попередніх значень $\varepsilon(k-i)$ на волатильність):

$$\log(h(k)) = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{|\varepsilon(k-i)|}{h(k-i)} + \sum_{i=1}^p \gamma_i \frac{\varepsilon(k-i)}{h(k-i)} + \sum_{i=1}^q \beta_j \log(h(k-i)).$$

Модель УАРУГ-М (модифікована) – моделювання премії за ризик:

$$y(k) = \beta + \gamma h(k) + \varepsilon(k),$$

$$h(k) = a_0 + a \sum_{i=1}^p \varepsilon^2(k-i) + \sum_{i=1}^q h(k) + \varepsilon_2(k).$$

Модель для прогнозування волатильності за допомогою УАРУГ:

$$\begin{aligned} h(k+1) &= \beta_0 + \beta_1 \varepsilon^2(k) + \gamma h(k), \\ h(k+j) &= \beta_0 + (\beta_1 + \gamma_1) h(k+j-1). \end{aligned}$$

Рівняння для умовної дисперсії та коваріації:

$$\begin{aligned} h(s(k)) &= \alpha_0 + \alpha_1 \varepsilon^2(s(k-1)) + \alpha_2 h(s(k-1)), \\ h(f(k)) &= \beta_0 + \beta_1 \varepsilon^2(f(k-1)) + \beta_2 h(f(k-1)), \\ \text{cov}[s(k), f(k)] &= \gamma_0 + \gamma_1 \varepsilon(s(k-1)) + \gamma_2 \text{cov}[s(k-1), f(k-1)]. \end{aligned}$$

Коефіцієнт хеджування при використанні двофакторної моделі:

$$\hat{b}^*(k) = \frac{\text{cov}[s(k), f(k)]}{\hat{h}^2(f)}.$$

Моделі коінтегрованих процесів, тобто нестационарних процесів, які можна об'єднати в межах однієї моделі, що має стаціонарні статистичні характеристики.

Концепція коінтегрованості змінних передбачає існування довгострокового зв'язку між значеннями змінних. Тобто ми припускаємо існування спільної врівноваженої траєкторії руху цих змінних, від якої вони можуть відхилятися на коротких проміжках часу. Однак економічні механізми в цілому діють таким чином, що рівновага відновлюється і зберігається на довгих часових інтервалах шляхом корегування відповідних відхилень від врівноваженого стану.

У випадку коінтегрованості змінних $x(k)$ і $y(k)$, для них може бути побудована модель корегування похибки, яка поєднує динаміку змінних на коротких проміжках часу з довгостроковим врівноваженим зв'язком та має наступний вигляд:

$$\begin{aligned}\Delta x(k) &= a_{10} + \sum_{i=1}^p b_{1i} \Delta x(k-i) + \sum_{i=1}^p c_{1i} \Delta y(k-i) + \lambda_1 e(k-1) + \varepsilon_1(k), \\ \Delta y(k) &= a_{20} + \sum_{i=1}^p b_{2i} \Delta y(k-i) + \sum_{i=1}^p c_{2i} \Delta x(k-i) + \lambda_2 e_x(k-1) + \varepsilon_2(k).\end{aligned}$$

Коефіцієнти λ_1, λ_2 в наведених рівняннях називають швидкістю пристосування (коригування). Більш докладно побудова та характеристики моделі коінтегрованих процесів буде розглянуто в розділі 13.

Вище ми навели тільки деякі структури (типи) математичних моделей стаціонарних та нестационарних нелінійних процесів, які широко використовуються при описанні динаміки процесів різної природи. На сьогодні існує дуже багато інших різновидів дискретних моделей, які описують різні складові реальних процесів. В подальшому ми будемо намагатися виконати аналіз по можливості більшого числа типів існуючих моделей, але необхідно зазначити, що зробити це в межах однієї книги дуже складно.

Запитання і вправи до першого розділу

1. Назвіть три напрями застосування математичних моделей? Наведіть приклади практичного застосування.
2. Чому математична модель, призначена для застосування в замкненому контурі керування процесом (об'єктом) може мати нижчу ступінь адекватності ніж модель, побудована для прогнозування?
3. Які вимоги висуваються до математичних моделей, призначених для застосування в реальному часі?
4. Сформулюйте необхідну умову для визначення оцінки періоду дискретизації вимірів (теорема Котельникова-Шеннона). Чому ця теорема представляє тільки необхідну умову?
5. Назвіть 5 елементів структури математичної моделі, поясніть їх на прикладі.
6. Поясніть фізичну суть запізнення по входу, наведіть приклади? Чи існують у природі процеси без запізнення і що необхідно для їх існування? Який середній час запізнення (реакції) для людини?
7. Яким чином враховується запізнення по входу в дискретних та неперервних моделях?
8. В чому полягає фізична суть існування авторегресії?
9. Назвіть основні типи нелінійностей процесів та які методи застосовують для їх оцінювання?
10. Сформулюйте узагальнений алгоритм побудови математичної моделі на основі експериментальних даних. Поясніть необхідність оцінювання декількох моделей-кандидатів.
11. Поясніть вимогу неперервності та синхронності до експериментальних даних.
12. Що означають вимоги представництва та інформативності до експериментальних даних?
13. Сформулюйте три основних вимоги до оцінок параметрів математичної моделі. При яких умовах оцінки, отримані за методом найменших квадратів, задовольняють трьом загальним вимогам до оцінок моделей?
14. Яким вимогам повинна задовольняти модель процесу?
15. Що означає *адекватність* моделі процесу? Яким чином можна її визначити?
16. Для чого необхідно знаходити розв'язки рівнянь, що описують динаміку процесу?
17. Що означають *універсальність* та *робастність* моделі?
18. Яким чином досягається адаптивність моделі?
19. Сформулюйте переваги та недоліки імітаційного моделювання?

20. В чому полягає відмінність між парною та множинною регресіями?
21. Чому регресійні рівняння називають стохастичними? Яку ще складову містять, як правило, часові ряди окрім випадкової?
22. Назвіть причини введення випадкової змінної $\varepsilon(k)$ у праву частину рівняння?
23. Якими методами можна коректно оцінювати параметри моделей, нелінійних відносно параметрів?
24. Який процес називають гетероскедастичним? Яка іншу назву мають такі процеси?
25. В чому полягає суть інтегрованості та коінтегрованості випадкових процесів?

Розділ 2

АНАЛІЗ СТАЦІОНАРНОСТІ ПРОЦЕСІВ АВТОРЕГРЕСІЇ З КОВЗНИМ СЕРЕДНІМ

2.1 Слабка і сильна стаціонарність

Процеси з одиничними коренями

Розглянемо докладніше властивості рівнянь, які застосовують для опису динаміки процесів різної природи, представлених часовими рядами. Уточнимо, яке ж рівняння коректно називати авторегресією з ковзним середнім?

Рівняння виду

$$\begin{aligned}
 y(k) &= a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j \varepsilon(k-j) + \varepsilon(k) = \\
 &= a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=0}^q b_j \varepsilon(k-j), \quad (2.1.1)
 \end{aligned}$$

$$b_0 = 1, \quad \sum_{j=1}^q b_j = 1,$$

називають авторегресією з ковзним середнім, якщо корені його характеристичного рівняння лежать всередині одиничного кола на комплексній площині, а $\{\varepsilon(k)\}$ – випадковий процес.

Якщо один або більше коренів характеристичного рівняння, записаного для (2.1.1), дорівнюють одиниці, то послідовність $\{y(k)\}$ називають *інтегрованою* або *процесом з одиничними коренями*, а рівняння (2.1.1) називають авторегресією з інтегрованим ковзним середнім (APIKC). Процеси з одиничними коренями відносяться до класу *нестаціонарних процесів*. Оскільки такі процеси є досить характерними для виробничих технологій, економіки, фінансів, екології та інших галузей, то їм необхідно приділити значну увагу. Нестаціонарні процеси такого типу особливо часто зустрічаються в переходній економіці та відповідній фінансовій діяльності, для яких властива висока нестаціонарна динаміка розвитку. Так, інфляція в

Україні зросла за три роки (1991 – 1993 роки) від 15-20% до 10250%. Таке катастрофічне зростання рівня інфляції супроводжувалось її повільним спадом до рівня 15-20% на протязі наступних 8-10 років. Це приклад процесу з надзвичайно високою динамікою, а також позитивним і негативним трендом на різних етапах його розвитку.

Сильна і слабка стаціонарність

Розрізняють стаціонарність *слабку* і *сильну*. Слабку стаціонарність називають ще:

- стаціонарністю *по коваріації*;
- стаціонарністю *другого порядку*;
- стаціонарністю *в широкому смислі*.

Формально стохастичний процес із скінченним середнім та дисперсією називають стаціонарним по коваріації, якщо для всіх k і $k-s$, ($k = 0, 1, 2, \dots$; $s = 0, 1, 2, \dots$) виконуються наступні умови:

- *математичне сподівання* не залежить від часу:

$$E[y(k)] = E[y(k-s)] = \mu = \text{const} ; \quad (2.1.2)$$

- *дисперсія* залишається постійною для всього часового інтервалу, на якому розглядається процес:

$$E\{[y(k)-\mu]^2\} = E\{[y(k-s)-\mu]^2\} = \sigma_y^2 = \text{const} \quad (2.1.3)$$

або $\text{var}[y(k)] = \text{var}[y(k-s)] = \sigma_y^2 = \text{const} ;$

- *автоковаріація* залишається незмінною в часі для всього часового інтервалу:

$$\begin{aligned} & E\{[y(k)-\mu][y(k-s)-\mu]\} = \\ & = E\{[y(k-j)-\mu][y(k-j-s)-\mu]\} = \gamma(s) = \text{const} , \end{aligned} \quad (2.1.4)$$

або

$$\text{cov}[y(k), y(k-s)] = \text{cov}[y(k-j), y(k-j-s)] = \gamma(s) = \text{const} .$$

Сильна стаціонарність не вимагає щоб математичне сподівання та дисперсія були постійними величинами.

Таким чином, слабка стаціонарність накладає більш строгі обмеження на поведінку випадкового процесу ніж сильна. В подальшому будемо розглядати ті процеси, які є стаціонарними по коваріації, тобто стаціонарні в

слабкому смислі. Відповідно, процес будемо вважати *нестаціонарним*, якщо не виконується хоча б одна з умов (2.1.2)-(2.1.4).

Зазначимо також, що основні статистичні характеристики процесу – математичне сподівання, дисперсію та коваріацію, можна обчислювати на основі конкретних значень часового ряду (*вибіркові характеристики*) або на основі математичної моделі, що описує досліджуваний процес (*теоретичні характеристики*).

При досліженні стаціонарності процесів важливо знати необхідні та достатні умови їх стаціонарності. Це дозволяє визначити конкретні параметри та їх комбінації, які впливають на стаціонарність. Такі знання необхідні для дослідження характеристик нових моделей, визначення областей стійкості розв'язків і т. ін.

2.2 Необхідні та достатні умови стаціонарності процесу AP(1)

Розглянемо стохастичний процес AP(1)

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k), \quad (2.2.1)$$

де $\{\varepsilon(k)\}$ – послідовність білого шуму з нульовим середнім, $E[\varepsilon(k)] = 0$, та скінченою постійною дисперсією $E[\varepsilon^2(k)] = \sigma_\varepsilon^2$; $y(0) = y_0$ – детермінована початкова умова.

Повний розв'язок цього рівняння має вигляд (докладно методика знаходження повних розв'язків різницевих рівнянь розглядається у третьому розділі):

$$y(k) = a_0 \sum_{i=0}^{k-1} a_1^i + a_1^k y_0 + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i). \quad (2.2.2)$$

Запишемо математичне сподівання для цього розв'язку:

$$E[y(k)] = a_0 \sum_{i=0}^{k-1} a_1^i + a_1^k y_0, \quad (2.2.3)$$

а також математичне сподівання для (2.2.2), але при зміні моменту часу на s кроків вперед:

$$E[y(k+s)] = a_0 \sum_{i=0}^{k+s-1} a_1^i + a_1^{k+s} y_0. \quad (2.2.4)$$

Вирази (2.2.3) і (2.2.4) свідчать про те, що $E[y(k)] \neq E[y(k+s)]$, тобто математичне сподівання залежить від часу і послідовність $\{y(k)\}$ є нестационарною на відносно коротких проміжках часу при невеликих значеннях k .

При $|a_1| < 1$ і $k \rightarrow \infty$ можна записати, що

$$\begin{aligned} \lim_{k \rightarrow \infty} [y(k)] &= \lim_{k \rightarrow \infty} \left[a_0 \sum_{i=0}^{k-1} a_1^i + a_1^k y_0 + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i) \right] = \\ &= \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i). \end{aligned} \quad (2.2.5)$$

Таким чином, безумовне математичне сподівання

$$\mu = E[y(k)] = \frac{a_0}{1-a_1}, \quad k \rightarrow \infty, \quad (2.2.6)$$

тобто при $k \rightarrow \infty$ процес буде стаціонарним: $E[y(k)] = E[y(k+s)] = \mu$.

Знайдемо вираз для теоретичної дисперсії при великих значеннях k за допомогою рівняння (2.2.5):

$$\begin{aligned} \text{var}[y(k)] &= E[(y(k) - \mu)^2] = E \left[\left(\sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) \right)^2 \right] = \\ &= E \left\{ [\varepsilon(k) + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + \dots] \times [\varepsilon(k) + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + \dots] \right\} = \\ &\sigma_{\varepsilon}^2 (1 + a_1^2 + a_1^4 + \dots) = \frac{\sigma_{\varepsilon}^2}{1 - a_1^2} = \text{const}, \end{aligned}$$

оскільки $E[\varepsilon(k-i) \varepsilon(k-j)] = 0$ при $i \neq j$.

Запишемо вираз для автоковаріації:

$$\begin{aligned} \text{cov}[y(k) - \mu] &= E[(y(k) - \mu)(y(k-s) - \mu)] = \\ &= E \left\{ [\varepsilon(k) + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + \dots + a_1^s \varepsilon(k-s) + a_1^{s+1} \varepsilon(k-s-1) + \dots] \times \right. \end{aligned}$$

$$\begin{aligned} & \times [\varepsilon(k-s) + a_1 \varepsilon(k-1-s) + a_1^2 \varepsilon(k-2-s) + \dots] = \\ & = \sigma_\varepsilon^2 a_1^s [1 + a_1^2 + a_1^4 + \dots] = \frac{\sigma_\varepsilon^2 a_1^s}{1 - a_1^2} = const \end{aligned}$$

для заданого значення параметра s .

Таким чином, стохастичний процес AP(1) стаціонарний при $|a_1| < 1$ і $k \rightarrow \infty$, тобто стаціональність проявляється на довгих часових інтервалах.

Якщо врахувати однорідний розв'язок, то отримаємо:

$$y(k) = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) + A a_1^k.$$

Тепер математичне сподівання

$$E[y(k)] = \frac{a_0}{1 - a_1} + A \cdot E[a_1^k]$$

залежить від часу і послідовність $\{y(k)\}$ буде стаціонарною тільки в тому випадку, коли $A a_1^k = 0$. Тобто для досягнення стаціонарності необхідно вибрати з початкових умов довільну константу $A = 0$. Величину $A a_1^k$ можна інтерпретувати як величину відхилення від довгострокової рівноваги.

Сформулюємо остаточно умови стаціонарності стохастичного процесу AP(1):

- а) модуль коефіцієнта $|a_1| < 1$ (корінь характеристичного рівняння);
- б) однорідний розв'язок повинен дорівнювати нулю, тобто $A = 0$, або процес починається досить давно в минулому, що забезпечує збіжність $a_1^k \rightarrow 0$.

Умови стаціонарності процесу авторегресії довільного порядку

Отримані умови стаціонарності процесу AP(1) легко розповсюджуються на процес авторегресії довільного порядку AP(p). Однорідний розв'язок для AP(p) має вигляд:

$$y^h(k) = \sum_{i=1}^p A_i \alpha_i^k,$$

а якщо є кратні корені, то

$$y^h(k) = \alpha^k \sum_{i=1}^m A_i k^{i-1} + \sum_{i=m+1}^p A_i \alpha_i^k, \quad (2.2.7)$$

де α – корінь характеристичного рівняння кратності m ; $\alpha_i, i = m+1, \dots, p$ – різні дійсні корені. Очевидно, що середнє, дисперсія та коваріація, обчислені для (2.2.7), будуть залежати від часу.

Таким чином, умовою стаціонарності процесу AP(p) є рівність нулю його однорідного розв'язку.

2.3 Умови стаціонарності часткового розв'язку рівняння АРКС

Розглянемо задачу визначення умов стаціонарності на прикладі. На практиці досить часто використовують модель АРКС(2, 1):

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + \varepsilon(k) + b_1 \varepsilon(k-1), \quad (2.3.1)$$

де $a_0 = 0$, оскільки воно не впливає на стаціональність. Оскільки для досягнення стаціонарності однорідний розв'язок повинен дорівнювати нулю, то знайдемо частковий розв'язок. Скористаємося пробною функцією для часткового розв'язку

$$y_{np}^p(k) = \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i) \quad (2.3.2)$$

і підставимо його в (2.3.1):

$$\begin{aligned} & \alpha_0 \varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots = \\ & = a_1 [\alpha_0 \varepsilon(k-1) + \alpha_1 \varepsilon(k-2) + \dots] + a_2 [\alpha_0 \varepsilon(k-2) + \alpha_1 \varepsilon(k-3) + \dots] + \\ & + \varepsilon(k) + b_1 \varepsilon(k-1). \end{aligned}$$

Для того щоб отримати тотожність, прирівняємо коефіцієнти при одинакових змінних:

$$\begin{aligned}
\alpha_0 &= 1; \\
\alpha_1 &= a_1 \alpha_0 + b_1 \quad \text{або} \quad \alpha_1 = a_1 + b_1; \\
\alpha_2 &= a_1 \alpha_1 + a_2 \alpha_0; \\
&\vdots \\
\alpha_i &= a_1 \alpha_{i-1} + a_2 \alpha_{i-2}, \quad \forall i \geq 2.
\end{aligned}$$

Характеристичне рівняння для $\alpha_i = a_1 \alpha_{i-1} + a_2 \alpha_{i-2}$, $\forall i \geq 2$ відповідає характеристичному рівнянню, записаному для однорідної частини (2.3.1):

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

Таким чином, якщо корені цього характеристичного рівняння знаходяться всередині одиничного кола на комплексній площині, то $\{\alpha_i\}$ – збіжний процес. Це і є умовою стаціонарності неоднорідного розв'язку рівняння АРКС(2, 1).

Статистичні характеристики послідовності $\{y_{np}^p(k) = \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i)\}$

Для перевірки стаціонарності послідовності $\{y_{np}^p(k)\}$ знайдемо математичне сподівання, дисперсію та автоковаріацію для неоднорідного розв'язку, зумовленого наявністю випадкової величини $\varepsilon(k)$:

a) математичне сподівання:

$$E[y^p(k)] = E[\alpha_0 \varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots] = 0,$$

а також

$$E[y^p(k)] = E[y^p(k-s)] = 0, \quad \forall k, s. \quad (2.3.3)$$

б) дисперсія:

$$\begin{aligned}
\text{var}[y^p(k)] &= E[\alpha_0 \varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots]^2 = \\
&= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \alpha_i^2,
\end{aligned} \quad (2.3.4)$$

тобто, $\text{var}[y^p(k)] = \text{var}[y^p(k-s)]$, $\forall k, s$.

в) автоковаріація:

$s = 1$:

$$\begin{aligned}\text{cov}[y^p(k), y^p(k-1)] &= E\{[\varepsilon(k) + \alpha_1\varepsilon(k-1) + \alpha_2\varepsilon(k-2) + \dots] \times \\ &\quad \times [\varepsilon(k-1) + \alpha_1\varepsilon(k-2) + \alpha_2\varepsilon(k-3) + \dots]\} = \\ &= \sigma_\varepsilon^2(\alpha_1 + \alpha_2\alpha_1 + \alpha_3\alpha_2 + \dots).\end{aligned}$$

$s = 2$:

$$\begin{aligned}\text{cov}[y^p(k), y^p(k-2)] &= E\{[\varepsilon(k) + \alpha_1\varepsilon(k-1) + \alpha_2\varepsilon(k-2) + \dots] \times \\ &\quad \times [\varepsilon(k-2) + \alpha_1\varepsilon(k-3) + \alpha_2\varepsilon(k-4) + \dots]\} = \\ &= \sigma_\varepsilon^2(\alpha_2 + \alpha_3\alpha_1 + \alpha_4\alpha_2 + \dots)\end{aligned}$$

і для довільного значення параметра s можна записати:

$$\text{cov}[y(k), y(k-s)] = \sigma_\varepsilon^2(\alpha_s + \alpha_{s+1}\alpha_1 + \alpha_{s+2}\alpha_2 + \dots) = \text{const}, \quad (2.3.5)$$

тобто коваріація постійна і не залежить від часу k .

Таким чином, якщо корені характеристичного рівняння λ_i не лежать всередині одиничного кола на комплексній площині, то послідовність коефіцієнтів $\{\alpha_i\}$ не буде збіжною. Звідси випливає, що послідовність $\{y^p(k)\}$ також буде розбіжною.

Приклад 2.1. Визначимо умови стаціонарності для процесу АРКС(2,1)

$$y(k) = 1,6y(k-1) - 0,9y(k-2) + \varepsilon(k) + 0,5\varepsilon(k-1). \quad (2.3.6)$$

За допомогою методу невизначених коефіцієнтів знайдемо коефіцієнти неоднорідного розв'язку, зумовленого наявністю $\varepsilon(k)$:

$$\begin{aligned}\alpha_0 &= 1; \\ \alpha_1 &= 1,6 + 0,5 = 2,1; \\ &\vdots \\ \alpha_i &= 1,6\alpha_{i-1} - 0,9\alpha_{i-2}, \quad \forall i \geq 2,\end{aligned} \quad (2.3.7)$$

тобто, (2.3.7) – це рівняння авторегресії другого порядку для коефіцієнтів неоднорідного розв'язку. Знайдемо розв'язок характеристичного рівняння для (2.3.6) та (2.3.7):

$$\begin{aligned}\lambda^2 - 1,6\lambda + 0,9 &= 0; \\ \lambda_{1,2} &= \frac{1}{2}(1,6 \pm \sqrt{2,56 - 3,6}) = \frac{1}{2}(1,6 \pm j1,02); \quad j = \sqrt{-1}.\end{aligned}$$

Розв'язок однорідного рівняння (2.3.7) в полярних координатах має

вигляд:

$$\alpha^h(i) = \beta_1 r^i \cos(\theta i + \beta_2),$$

де $r = \sqrt{-a_2} = \sqrt{0,9} = 0,949$; $\cos(\theta) = \frac{a_1}{2(a_2)^{1/2}} = \frac{1,6}{2 \cdot 0,949} = 0,843$, а звідси значення кута в радіанах: $\theta = 0,567$.

Таким чином, можна конкретизувати однорідний розв'язок

$$\alpha^h(i) = \beta_1 (0,949)^i \cos(0,567 i + \beta_2). \quad (2.3.8)$$

Константи β_1 і β_2 знайдемо за допомогою початкових умов: $\alpha_0 = 1$, $\alpha_1 = 2,1$:

$$\begin{cases} 1 = \beta_1 \cos(\beta_2) \\ 2,1 = 0,949 \beta_1 \cos(0,567 + \beta_2) \end{cases}.$$

Звідси $\beta_1 = \frac{1}{\cos(\beta_2)}$; $\cos(\beta_2) - \frac{0,949}{2,1} \cos(0,567 + \beta_2) = 0$. Скориставшись тригонометричною тотожністю

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta),$$

знайдемо, що $\beta_2 = -1,197$. Таким чином, розв'язок, із врахуванням початкових умов, приймає вигляд:

$$\alpha(i) = -\frac{1}{\cos(-1,197)} \cdot 0,949^i \cos(0,567 i - 1,197). \quad (2.3.9)$$

З (2.3.9) випливає, що загальна тенденція зміни коефіцієнтів α_i у часі буде в сторону зменшення їх значень. Деякі значення цих коефіцієнтів наведено в таблиці 6.1.

Таблиця 6.1

i	0	1	2	3	4	5	6	7	8	9	10
α_i	1,0	2,1	2,46	2,046	1,06	-0,146	1,187	-1,786	-1,761	-1,22	-0,37

Якщо продовжити обчислення коефіцієнтів α_i , то побачимо, що цей процес збіжний – затухаючі гармонічні коливання. Звідси випливає, що процес $\{y(k)\}$ також повинен бути збіжним.

2.4 Узагальнення умов стаціонарності на процеси АРКС(p, q) довільного порядку

В даному параграфі розглянемо задачу визначення умов стаціонарності для процесів ковзного середнього та авторегресії довільного порядку.

Стаціонарність процесу $KC(\infty)$

Процес $KC(\infty)$ описується рівнянням

$$x(k) = \sum_{i=0}^{\infty} b_i \varepsilon(k-i), \quad (2.4.1)$$

де $\{\varepsilon(k)\}$ – процес білого шуму із скінченою постійною дисперсією σ_ε^2 . Необхідно встановити чи буде процес $x(k)$ стаціонарним за коваріацією. Запишемо вирази для математичного сподівання, дисперсії та коваріації процесу (2.4.1):

а) математичне сподівання:

$$\begin{aligned} \mu_x &= E[x(k)] = E[\varepsilon(k) + b_1\varepsilon(k-1) + b_2\varepsilon(k-2) + \dots] = 0; \\ \mu_x &= E[x(k)] = E[x(k-s)] = 0; \end{aligned}$$

б) дисперсія:

$$\begin{aligned} \text{var}[x(k)] &= E\{[\varepsilon(k) + b_1\varepsilon(k-1) + b_2\varepsilon(k-2) + \dots]^2\} = \\ &= E[\varepsilon(k)]^2 + b_1^2 E[\varepsilon(k-1)]^2 + b_2^2 E[\varepsilon(k-2)]^2 + \dots = \\ &= \sigma_\varepsilon^2 [1 + b_1^2 + b_2^2 + \dots]. \end{aligned} \quad (2.4.2)$$

Очевидно, що $\text{var}[x(k-s)] = \sigma_\varepsilon^2 [1 + b_1^2 + b_2^2 + \dots]$ і

$$\text{var}[x(k)] = \text{var}[x(k-s)] = \text{const}, \quad \forall k, s.$$

Звідси випливає, що умовою стаціонарності процесу $\{x(k)\}$ є скінченість суми $\sum b_i^2$.

в) автоковаріація:

$$\begin{aligned} \text{cov}[x(k), x(k-s)] &= E\{[\varepsilon(k) + b_1\varepsilon(k-1) + b_2\varepsilon(k-2) + \dots] \times \\ &\quad \times [\varepsilon(k-s) + b_1\varepsilon(k-s-1) + b_2\varepsilon(k-s-2) + \dots]\} = \\ &= \sigma_\varepsilon^2 (b_s + b_1b_{s+1} + b_2b_{s+2} + \dots). \end{aligned} \quad (2.4.3)$$

Припустимо, що сума $(b_s + b_1b_{s+1} + b_2b_{s+2} + \dots)$ скінчена. В такому випадку коваріація залежить тільки від s і не залежить від k . Тобто, скінченість суми і буде однією з умов стаціонарності процесу $\{x(k)\}$.

Таким чином, необхідними і достатніми умовами стаціонарності процесу КС(∞) є скінченість сум коефіцієнтів:

$$\sum b_i^2 \leq (b_s + b_1b_{s+1} + b_2b_{s+2} + \dots). \quad (2.4.4)$$

Оскільки остання сума повинна бути скінченою для всіх значень s , то умову скінченості суми $\sum b_i^2$ можна вважати надлишковою. Звідси можна зробити висновок, що будь-який процес ковзного середнього скінченого порядку буде стаціонарним.

Висновок: будь-який процес ковзного середнього, що має скінчений порядок, є стаціонарним.

Стаціональність процесу AP(p)

Визначимо обмеження стаціонарності на коефіцієнти авторегресії довільного порядку p . Частковий розв'язок рівняння

$$y(k) = a_0 + \sum_{i=0}^p a_i y(k-i) + \varepsilon(k) \quad (2.4.5)$$

має вигляд:

$$y^p(k) = \frac{a_0}{1 - \sum_{i=1}^p a_i} + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i). \quad (2.4.6)$$

По аналогії з рівнянням другого порядку в даному випадку також можна знайти різницеве рівняння для обчислення коефіцієнтів α_i :

$$\alpha_i = a_1 \alpha_{i-1} + a_2 \alpha_{i-2} + \dots + a_p \alpha_{i-p}. \quad (2.4.7)$$

Якщо корені характеристичного рівняння, записаного для (2.4.7), знаходяться всередині одиничного кола на комплексній площині, то послідовність коефіцієнтів $\{\alpha_i\}$ буде збіжною, тобто, сума $\sum \alpha_i^2$ буде скінченою. Звідси випливає, що для перевірки стаціонарності послідовності $\{y(k)\}$ можна скористатись рівнянням (2.4.6). Для цього знайдемо вирази для математичного сподівання, дисперсії та коваріації:

а) математичне сподівання:

$$E[y(k)] = E[y(k-s)] = \frac{a_0}{1 - \sum a_i} = \text{const}.$$

Ця величина буде скінченою, якщо $1 - \sum a_i \neq 0$, тобто це умова того, що корені відповідного характеристичного рівняння знаходяться всередині одиничного кола.

б) Дисперсія:

$$\begin{aligned} \text{var}[y(k)] &= E\{[\varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots]^2\} = \\ &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \alpha_i^2. \\ \text{var}[y(k-s)] &= E\{[\varepsilon(k-s) + \alpha_1 \varepsilon(k-s-1) + \alpha_2 \varepsilon(k-s-2) + \dots]^2\} = \\ &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \alpha_i^2. \end{aligned}$$

Таким чином, дисперсія буде скінченою, якщо сума $\sum \alpha_i^2$ є скінченою.

в) Автоковаріація:

$$\begin{aligned} \gamma_s &= \text{cov}[y(k), y(k-s)] = E\{[\varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots] \times \\ &\quad \times [\varepsilon(k-s) + \alpha_1 \varepsilon(k-s-1) + \alpha_2 \varepsilon(k-s-2) + \dots]\} = \\ &= \sigma_\varepsilon^2 (\alpha_s + \alpha_1 \alpha_{s+1} + \alpha_2 \alpha_{s+2} + \dots) = \text{const} \end{aligned}$$

і не залежить від часу. Таким чином, процес AP(p) буде стаціонарним, якщо

відповідні суми коефіцієнтів будуть збіжними.

Процес АРКС(p, q) в цілому

Розглянемо тепер процес АРКС(p, q) в цілому:

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=0}^q b_j \varepsilon(k-j), \quad b_0 = 0. \quad (2.4.8)$$

Запишемо неоднорідний розв'язок за допомогою лагового оператора:

$$y(k) = \frac{a_0}{1 - \sum_{i=1}^p a_i} + \frac{\varepsilon(k)}{1 - \sum_{i=1}^p a_i L^i} + \frac{b_1 \varepsilon(k-1)}{1 - \sum_{i=1}^p a_i L^i} + \dots \quad (2.4.9)$$

Перший член в правій частині (2.4.9) є константою, а всі інші будуть стаціонарними, якщо корені поліномів $(1 - \sum a_i L^i)$ лежать всередині одиничного кола.

Таким чином, якщо послідовність $\{x(k)\}$, яка описується рівнянням (2.4.1), є стаціонарною, то і послідовність $\{y(k)\}$ буде стаціонарною при умові, що корені характеристичного рівняння для авторегресійної частини знаходяться всередині одиничного кола.

Висновок: якщо послідовність $\{x(k)\}$, яка описується рівнянням (2.4.1), стаціонарна, то і послідовність $\{y(k)\}$ буде стаціонарною при умові, що корені характеристичного рівняння для авторегресійної частини знаходяться всередині одиничного кола на комплексній площині.

2.5 Теоретична автокореляційна функція

В попередньому параграфі ми встановили, що для визначення стаціонарності процесів необхідно обчислювати три основні статистичні характеристики: математичне сподівання, дисперсію і коваріацію. Ці

параметри можна обчислювати безпосередньо за вибіркою, тобто знаходити *вибіркові характеристики*, або за допомогою математичної моделі – *теоретичні характеристики*.

Розглянемо розрахунок теоретичних статистичних характеристик для моделі AP(1):

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k), \quad E[\varepsilon(k)] = 0,$$

$$E[\varepsilon(k)\varepsilon(k-j)] = \begin{cases} \sigma_\varepsilon^2, & j=0, \\ 0, & j \geq 1; \end{cases} \quad E[\varepsilon(k)\varepsilon(y-j)] = \begin{cases} \sigma_\varepsilon^2, & j=0, \\ 0, & j \geq 1. \end{cases}$$

Дисперсія: $\text{var}[y(k)] = \text{var}[a_1 y(k-1)] + \text{var}[\varepsilon(k)].$ Оскільки для стаціонарного процесу, при $|a_1| < 1$, виконується умова, що $\text{var}[y(k)] = \text{var}[y(k-1)]$, то $\text{var}[y(k)](1 - a_1^2) = \text{var}[\varepsilon(k)].$ Звідси знайдемо вираз для дисперсії:

$$\sigma_y^2 = \text{var}[y(k)] = \frac{\sigma_\varepsilon^2}{1 - a_1^2} = \gamma_0, \quad |a_1| < 1. \quad (2.5.1)$$

Нагадаємо, що вибіркова дисперсія часового ряду $\{y(k)\}$ обчислюється за виразом:

$$\sigma_y^2 = \frac{1}{N-1} \sum_{k=1}^N [y(k) - \bar{y}]^2,$$

де \bar{y} – середнє значення ряду.

Коваріація: оскільки для стаціонарного процесу справедлива рівність: $\text{cov}[y(k), y(k-s)] = \text{cov}[y(k), y(k+s)]$, то зручніше користуватись останнім визначенням. Запишемо рівняння AP(1) для моменту $(k+1)$:

$$y(k+1) = a_1 y(k) + \varepsilon(k+1) = a_1 [a_1 y(k-1) + \varepsilon(k)] + \varepsilon(k+1),$$

$$\begin{aligned} \gamma(1) &= \text{cov}[y(k), y(k+1)] = \\ &= E\{[a_1 y(k-1) + \varepsilon(k)][a_1^2 y(k-1) + a_1 \varepsilon(k) + \varepsilon(k+1)]\} = \\ &= a_1^3 \text{var}[y(k-1)] + a_1 \text{var}[\varepsilon(k)] = a_1^3 \frac{\sigma_y^2}{1 - a_1^2} + a_1 \sigma_\varepsilon^2. \end{aligned}$$

Покладаючи $\sigma_y^2 \approx \sigma_\varepsilon^2$ (що не входить в протиріччя з визначенням

випадкових процесів), отримаємо:

$$\gamma(1) = \text{cov}[y(k), y(k+1)] = \frac{a_1^3 \sigma_\varepsilon^2 + a_1 \sigma_\varepsilon^2 - a_1^3 \sigma_\varepsilon^2}{1 - a_1^2} = \frac{a_1 \sigma_\varepsilon^2}{1 - a_1^2}. \quad (2.5.2)$$

Для довільного значення s можна тепер записати:

$$\gamma(s) = E\{[y(k) - \mu_y], [y(k-s) - \mu_y]\} = \frac{\sigma_\varepsilon^2 a_1^s}{1 - a_1^2}. \quad (2.5.3)$$

Автокореляційну функцію (АКФ) можна знайти шляхом нормування коваріаційної функції відносно $\gamma(0) = \gamma_0$:

$$\rho(s) = \frac{\gamma(s)}{\gamma_0}; \quad \rho_0 = 1, \quad \rho(1) = a_1, \quad \rho(2) = a_1^2, \dots, \rho(s) = a_1^s.$$

Якщо $0 < a_1 < 1$, то АКФ експоненціально спадає, а при $-1 < a_1 < 0$ АКФ має коливальний спадаючий характер. Порівняння теоретичної та вибіркової АКФ також дозволяє визначити адекватність моделі процесу. Крім того, теоретична АКФ полегшує оцінювання стаціонарності процесів.

2.6 Обчислення кореляційних функцій за допомогою рівнянь Юла-Уокера

АКФ процесу авторегресії другого порядку AP(2)

Систему рівнянь Юла-Уокера отримують шляхом множення обох частин рівняння, що описує динаміку процесу, на $y(k-s)$, $s = 0, 1, 2, \dots$. Помножимо ліву і праву частини рівняння

$$\begin{aligned} y(k) &= a_1 y(k-1) + a_2 y(k-2) + \varepsilon(k), \\ \{\varepsilon(k)\} &\sim N(0, \sigma_\varepsilon^2), \end{aligned} \quad (2.6.1)$$

на $y(k-s)$, $s = 0, 1, 2, \dots$ і знайдемо математичне сподівання обох частин. Зазначимо, що перетин a_0 не впливає на кореляційну функцію, а тому його виключено з рівняння (2.6.1).

Система рівнянь Юла-Уокера в даному випадку приймає вигляд:

$$\begin{aligned}
s=0: \quad & E[y(k)y(k)] = a_1E[y(k-1)y(k)] + a_2E[y(k-2)y(k)] + E[\varepsilon(k)y(k)], \\
s=1: \quad & E[y(k)y(k-1)] = a_1E[y(k-1)y(k-1)] + a_2E[y(k-2)y(k-1)] + \\
& \quad + E[\varepsilon(k)y(k-1)], \\
s=2: \quad & E[y(k)y(k-2)] = a_1E[y(k-1)y(k-2)] + a_2E[y(k-2)y(k-2)] + \\
& \quad + E[\varepsilon(k)y(k-2)], \\
& \vdots \\
\exists s: \quad & E[y(k)y(k-s)] = a_1E[y(k-1)y(k-s)] + a_2E[y(k-2)y(k-s)] + \\
& \quad + E[\varepsilon(k)y(k-s)].
\end{aligned} \tag{2.6.2}$$

За визначенням відомо, що коваріація

$$E[y(k)y(k-s)] = E[y(k-s)y(k)] = \gamma_s,$$

а дисперсія ряду $\{y(k)\}$ визначається дисперсією випадкової змінної $\varepsilon(k)$, тобто, $E[\varepsilon(k)y(k)] \approx \sigma_\varepsilon^2$. Крім того, приймемо, що $b_0 = 1$ при $\varepsilon(k)$, і $\varepsilon(k)$, не повинно бути корелюваним з $y(k-s)$, $s \geq 1$ або $E[\varepsilon(k)y(k-s)] = 0$ при $s \geq 1$.

Запишемо систему (2.6.2) через коваріації $\gamma_0, \gamma_1, \dots, \gamma_s$:

$$\left\{
\begin{array}{l}
\gamma_0 = a_1\gamma_1 + a_2\gamma_2 + \sigma_\varepsilon^2 \\
\gamma_1 = a_1\gamma_0 + a_2\gamma_1 \\
\vdots \\
\gamma_s = a_1\gamma_{s-1} + a_2\gamma_{s-2}
\end{array}
\right. \tag{2.6.3}$$

і поділимо два останніх рівняння на γ_0 :

$$\left\{
\begin{array}{l}
\rho_1 = a_1\rho_0 + a_2\rho_1, \\
\rho_s = a_1\rho_{s-1} + a_2\rho_{s-2}.
\end{array}
\right. \tag{2.6.4}$$

Оскільки $\rho_0 = 1$, то $\rho_1 = a_1/(1-a_2)$. Всі інші значення дискретної кореляційної функції можна знайти за допомогою рекурсивної формули:

$$\begin{aligned}
\rho_s &= a_1\rho_{s-1} + a_2\rho_{s-2} \\
\text{або} \quad \rho(s) &= a_1\rho(s-1) + a_2\rho(s-2).
\end{aligned} \tag{2.6.5}$$

Так,

$$\rho(2) = \frac{a_1}{1-a_2} + a_2, \quad \rho(3) = a_1 \left(\frac{a_1^2}{1-a_2} + a_2 \right) + \frac{a_2 a_1}{1-a_2}.$$

АКФ процесу ковзного середнього 1-го порядку КС(1)

Для процесу КС(1) рівняння Юла-Уокера мають вигляд:

$$\begin{aligned} \gamma_0 &= E[y(k)y(k)] = E\{[\varepsilon(k) + b_1\varepsilon(k-1)][\varepsilon(k) + b_1\varepsilon(k-1)]\} = \\ &= \sigma_\varepsilon^2 + b_1^2 \sigma_\varepsilon^2 = (1+b_1^2)\sigma_\varepsilon^2, \\ \gamma_1 &= E[y(k)y(k-1)] = E\{[\varepsilon(k) + b_1\varepsilon(k-1)][\varepsilon(k-1) + b_1\varepsilon(k-2)]\} = \\ &= b_1 \sigma_\varepsilon^2, \\ \gamma_2 &= E[y(k)y(k-2)] = E\{[\varepsilon(k) + b_1\varepsilon(k-1)][\varepsilon(k-2) + b_1\varepsilon(k-3)]\} = 0 \\ i \quad \gamma_s &= 0. \end{aligned}$$

Поділимо γ_0 , γ_1 на γ_0 і отримаємо:

$$\rho_0 = 1, \quad \rho_1 = \frac{b_1}{1+b_1^2}, \quad \rho_s = 0, \quad \forall s > 1,$$

тобто, АКФ процесу КС(1) має тільки два значення, відмінних від нуля.

АКФ процесу APKC(1,1)

Для рівняння

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k) + b_1 \varepsilon(k-1)$$

необхідно знайти коваріаційну функцію γ_s і АКФ ρ_s . За визначенням,

$$\begin{aligned} \gamma_s &= E\{[y(k) - \mu_y][y(k-s) - \mu_y]\}, \\ E[\varepsilon(k)y(k-s)] &= 0, \quad s \geq 1, \\ E[\varepsilon(k)y(k)] &= \sigma_\varepsilon^2, \end{aligned}$$

$$\begin{aligned} E[\varepsilon(k-1) y(k)] &= E\{\varepsilon(k-1)[a_1 y(k-1) + \varepsilon(k) + b_1 \varepsilon(k-1)]\} = \\ &= a_1 \sigma_\varepsilon^2 + b_1 \sigma_\varepsilon^2 = \sigma_\varepsilon^2 (a_1 + b_1). \end{aligned}$$

Рівняння Юла-Уокера:

$$\begin{aligned} \gamma_0 &= E[y(k) y(k)] = a_1 E[y(k) y(k-1)] + \\ &\quad + E[y(k) \varepsilon(k)] + b_1 E[y(k) \varepsilon(k-1)] = \\ &= a_1 \gamma_1 + \sigma_\varepsilon^2 + b_1 (a_1 + b_1) \sigma_\varepsilon^2, \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E[y(k) y(k-1)] = a_1 E[y(k-1) y(k-1)] + \\ &\quad + E[y(k-1) \varepsilon(k)] + b_1 E[y(k-1) \varepsilon(k-1)] = \\ &= a_1 \gamma_0 + b_1 \sigma_\varepsilon^2. \end{aligned}$$

$$\begin{aligned} \gamma_2 &= E[y(k) y(k-2)] = a_1 E[y(k-1) y(k-2)] + \\ &\quad + E[y(k-2) \varepsilon(k)] + b_1 E[y(k-2) \varepsilon(k-1)] = \\ &= a_1 \gamma_1. \end{aligned}$$

⋮

$$\begin{aligned} \gamma_s &= E[y(k) y(k-s)] = a_1 E[y(k-1) y(k-s)] + \\ &\quad + E[y(k-s) \varepsilon(k)] + b_1 E[y(k-s) \varepsilon(k-1)] = \\ &= a_1 \gamma_{s-1}. \end{aligned}$$

У скороченій формі:

$$\begin{aligned} \gamma_0 &= a_1 \gamma_1 + \sigma_\varepsilon^2 + b_1 (a_1 + b_1) \sigma_\varepsilon^2, \\ \gamma_1 &= a_1 \gamma_0 + b_1 \sigma_\varepsilon^2, \\ \gamma_2 &= a_1 \gamma_1, \\ \gamma_3 &= a_1 \gamma_2, \\ &\vdots \\ \gamma_s &= a_1 \gamma_{s-1}. \end{aligned}$$

Розв'язуючи цю систему рівнянь, отримаємо:

$$\gamma_0 = \frac{1 + b_1^2 + 2a_1 b_1}{1 - a_1^2} \cdot \sigma_\varepsilon^2; \quad \gamma_1 = \frac{(1 + a_1 b_1)(a_1 + b_1)}{1 + b_1^2 + 2a_1 b_1},$$

і коефіцієнти кореляції:

$$\rho_0 = 1, \quad \rho_1 = \frac{(1+a_1b_1)(a_1+b_1)}{1+b_1^2 + 2a_1b_1}, \quad \rho_s = a_1\rho_{s-1}, \quad \forall s > 1.$$

Оскільки $\rho_s = a_1\rho_{s-1}$, то збіжність процесу залежить від значення a_1 . Якщо $0 < a_1 < 1$, то збіжність гладка; а якщо $-1 < a_1 < 0$, то збіжність носить коливальний характер.

АКФ процесу APKC(2,1)

Для рівняння

$$y(k) = a_0 + a_1y(k-1) + a_2y(k-2) + \varepsilon(k) + b_1\varepsilon(k-1)$$

необхідно знайти коваріаційну функцію γ_s і АКФ ρ_s . За означенням,

$$\begin{aligned} \gamma_s &= E\{[y(k)-\mu_y][y(k-s)-\mu_y]\}, \\ E[\varepsilon(k)y(k-s)] &= 0, \quad s \geq 1, \\ E[\varepsilon(k)y(k)] &= \sigma_\varepsilon^2, \end{aligned}$$

$$\begin{aligned} E[\varepsilon(k-1)y(k)] &= E\{\varepsilon(k-1)[a_1y(k-1) + \varepsilon(k) + b_1\varepsilon(k-1)]\} = \\ &= a_1\sigma_\varepsilon^2 + b_1\sigma_\varepsilon^2 = \sigma_\varepsilon^2(a_1 + b_1). \end{aligned}$$

Рівняння Юла-Уокера:

$$\begin{aligned} \gamma_0 &= E[y(k)y(k)] = a_1E[y(k)y(k-1)] + a_2E[y(k)y(k-2)] + \\ &\quad + E[y(k)\varepsilon(k)] + b_1E[y(k)\varepsilon(k-1)] = \\ &= a_1\gamma_1 + a_2\gamma_2 + \sigma_\varepsilon^2 + b_1(a_1 + b_1)\sigma_\varepsilon^2, \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E[y(k)y(k-1)] = a_1E[y(k-1)y(k-1)] + a_2E[y(k-1)y(k-2)] + \\ &\quad + E[y(k-1)\varepsilon(k)] + b_1E[y(k-1)\varepsilon(k-1)] = \\ &= a_1\gamma_0 + a_2\gamma_1 + b_1\sigma_\varepsilon^2. \end{aligned}$$

$$\begin{aligned} \gamma_2 &= E[y(k)y(k-2)] = a_1E[y(k-1)y(k-2)] + a_2E[y(k-2)y(k-2)] + \\ &\quad + E[y(k-2)\varepsilon(k)] + b_1E[y(k-2)\varepsilon(k-1)] = \\ &= a_1\gamma_1 + a_2\gamma_0. \end{aligned}$$

$$\begin{aligned}\gamma_3 &= E[y(k)y(k-3)] = a_1 E[y(k-1)y(k-3)] + a_2 E[y(k-2)y(k-3)] + \\ &\quad + E[y(k-3)\varepsilon(k)] + b_1 E[y(k-3)\varepsilon(k-1)] = \\ &= a_1\gamma_2 + a_2\gamma_1.\end{aligned}$$

⋮

$$\gamma_s = a_1\gamma_{s-1} + a_2\gamma_{s-2}.$$

Перейдемо до скороченої форми запису:

$$\begin{aligned}\gamma_0 &= a_1\gamma_1 + a_2\gamma_2 + \sigma_\varepsilon^2[1 + b_1(a_1 + b_1)], \\ \gamma_1 &= a_1\gamma_0 + a_2\gamma_1 + b_1\sigma_\varepsilon^2, \\ \gamma_2 &= a_1\gamma_1 + a_2\gamma_0, \\ \gamma_3 &= a_1\gamma_2 + a_2\gamma_1, \\ \gamma_4 &= a_1\gamma_3 + a_2\gamma_2, \\ &\vdots \\ \gamma_s &= a_1\gamma_{s-1} + a_2\gamma_{s-2}.\end{aligned}$$

Розв'язуючи цю систему рівнянь, знайдемо, що:

$$\gamma_0 = \frac{a_1(1+a_2)c_3 + c}{1 - [a_1c_2(1+a_2) + a_2^2]}, \quad \gamma_1 = \frac{a_1\gamma_0 + c_1}{1 - a_2},$$

де $c = \sigma_\varepsilon^2[1 + b_1(a_1 + b_1)]$; $c_1 = b_1\sigma_\varepsilon^2$; $c_2 = a_1/(1 - a_2)$; $c_3 = c_1/(1 - a_2)$, а всі інші значення γ_s можна знайти за наведеною вище рекурсивною формулою.

В загальному випадку для процесу APKC(p, q) коефіцієнти кореляції можна знайти за допомогою наступної рекурсивної формули:

$$\rho_i = a_1\rho_{i-1} + a_2\rho_{i-2} + \dots + a_p\rho_{i-p}.$$

2.7 Часткова автокореляційна функція

Перед тим як розглядати часткову автокореляційну функцію визначимо деякі особливості звичайної АКФ. Так, АКФ процесу AP(1) обчислюється за виразами:

$$\rho_0 = 1, \rho(1) = a_1, \rho(2) = a_1^2, \dots, \rho(s) = a_1^s,$$

тобто значення $y(k)$ і $y(k-2)$ корельовані, хоча $y(k-2)$ в моделі відсутнє. З наведеного виразу також видно, що коефіцієнт кореляції ρ_2 між значеннями $y(k)$ і $y(k-2)$ дорівнює добутку коефіцієнтів кореляції між $y(k)$ і $y(k-1)$ та між $y(k-1)$ і $y(k-2)$:

$$\rho_2 = \rho_1 \cdot \rho_1 = \rho_1^2.$$

Очевидно, що такі “непрямі” кореляції наявні в АКФ будь-якого процесу авторегресії. В цілому це призводить до того, що АКФ не досить чітко відображає порядок авторегресії.

На відміну від АКФ, часткова АКФ (ЧАКФ) виключає вплив проміжних коефіцієнтів кореляції, тобто на значення коефіцієнта часткової кореляції ϕ_{ss} між $y(k)$ і $y(k-s)$ не впливають величини $y(k-1), \dots, y(k-s+1)$ (рис. 2.1).

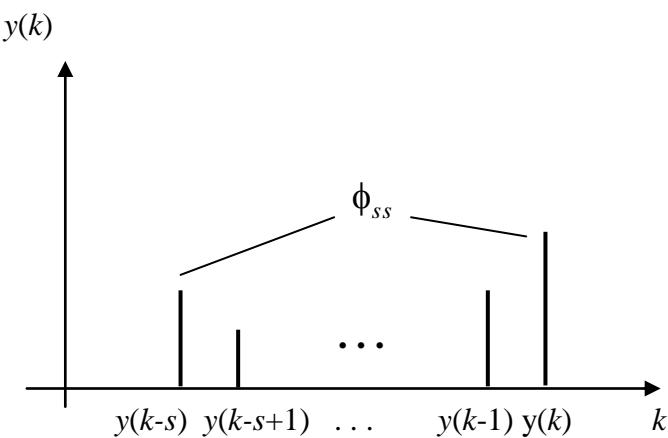


Рис. 2.1. Коефіцієнт часткової кореляції між $y(k)$ і $y(k-s)$.

Для процесу AR(1) коефіцієнт ЧАКФ між $y(k)$ і $y(k-2)$ дорівнює нулю за визначенням. Звідси випливає, що в загальному випадку для процесу AR(p) немає прямої кореляції між значеннями $y(k)$ і $y(k-s)$ при $s > p$: $\phi_{ss} = 0, \forall s > p$. Ця властивість ЧАКФ є дуже корисною при розв'язку задач ідентифікації моделей процесів AR(p).

Однак, це не відноситься до процесів ковзного середнього. Так, для процесу CC(1)

$$y(k) = \varepsilon(k) + b_1 \varepsilon(k-1)$$

можна записати $y(k) = (1 + b_1 L) \varepsilon(k)$ або, скориставшись властивістю операторів зсуву, отримаємо:

$$\varepsilon(k) = \frac{y(k)}{1 + b_1 L} = y(k) - b_1 y(k-1) + b_1^2 y(k-2) - b_1^3 y(k-3) + \dots, \quad |b_1| < 1.$$

З цього виразу випливає, що значення $y(k)$ буде корельованим із всіма іншими, затриманими в часі значеннями і ЧАКФ цього процесу не спадає різко до нуля.

Коефіцієнти ЧАКФ можна обчислити безпосередньо за допомогою значень часового ряду. Процедуру обчислень можна представити у вигляді двох кроків:

1. Сформувати вибірку з відхилень від середнього:

$$\{y'(k)\} = \{y(k)\} - \mu_y.$$

2. Оцінити рівняння необхідного порядку

$$y'(k) = \phi_{11} y'(k-1) + e(k), \quad (2.7.1)$$

де $e(k)$ – похибка моделі, яка може і не бути білим шумом. Оскільки рівняння (2.7.1) не містить інших значень крім $y'(k)$ ті $y'(k-1)$, то ϕ_{11} є коефіцієнтом часткової кореляції між цими значеннями.

По аналогії можна сформувати та оцінити рівняння другого порядку

$$y'(k) = \phi_{21} y'(k-1) + \phi_{22} y'(k-2) + e(k),$$

де ϕ_{22} – коефіцієнт часткової кореляції між $y'(k)$ і $y'(k-2)$, який не враховує впливу значення $y'(k-1)$.

Однак на практиці ЧАКФ обчислюють за наступними виразами:

$$\phi_{11} = \rho_1, \quad \phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}; \quad \phi_{ss} = \frac{\rho_s - \sum_{j=1}^{s-1} \phi_{s-1,j} \rho_{s-j}}{1 - \sum_{j=1}^{s-1} \phi_{s-1,j} \rho_j}, \quad (2.7.2)$$

де $\phi_{sj} = \phi_{s-1,j} - \phi_{ss} \phi_{s-1,s-j}$, $j = 1, 2, 3, \dots, s-1$.

Приклад обчислення ЧАКФ

Обчислити АКФ і ЧАКФ для процесу АРКС(1,1):

$$y(k) = -0,7y(k-1) + \varepsilon(k) - 0,7\varepsilon(k-1).$$

а) Коефіцієнти АКФ обчислимо за допомогою отриманих вище виразів:
 $\rho_0 = 1$;

$$\rho_1 = \frac{(1+a_1b_1)(a_1+b_1)}{1+b_1^2+2a_1b_1} = \frac{(1+0,49)(-0,7-0,7)}{1+0,49+2\cdot0,49} = -0,8445.$$

Тепер можна скористатись рекурсивною формулою $\rho_i = -0,7\rho_{i-1}$:

$$\begin{aligned} \rho_2 &= 0,591; \quad \rho_3 = -0,414; \quad \rho_4 = 0,290; \quad \rho_5 = -0,203; \quad \rho_6 = 0,142; \\ \rho_7 &= -0,010; \quad \rho_8 = 0,070; \quad \rho_9 = -0,049. \end{aligned}$$

б) коефіцієнти ЧАКФ:

$$\phi_{11} = \rho_1 = -0,8445;$$

$$\phi_{22} = (\rho_2 - \rho_1^2)/(1 - \rho_1^2) = \frac{0,591 - (-0,8445)^2}{1 - (-0,8445)^2} = -0,425;$$

$$\phi_{21} = \phi_{11} - \phi_{22}\phi_{11} = -1,204;$$

$$\phi_{33} = \frac{\rho_3 - \sum_{j=1}^2 \phi_{2,j} \rho_{3-j}}{1 - \sum_{j=1}^2 \phi_{2,j} \rho_j} = -0,262; \quad \phi_{44} = \frac{\rho_4 - \sum_{j=1}^3 \phi_{3,j} \rho_{4-j}}{1 - \sum_{j=1}^3 \phi_{3,j} \rho_j} = -0,173,$$

де $\phi_{3j} = \phi_{2j} - \phi_{33}\phi_{2,2-j}$, а звідси $\phi_{3,1} = -1,315$; $\phi_{3,2} = -0,74$. Інші коефіцієнти мають такі значення:

$$\phi_{55} = -0,117; \quad \phi_{66} = -0,0681; \quad \phi_{77} = -0,056; \quad \phi_{88} = -0,039.$$

В загальному випадку ЧАКФ стаціонарного процесу АРКС(p,q) повинна збігатись до нуля, починаючи з p -го значення, а швидкість збіжності залежить від значень коефіцієнтів полінома

$$1 + b_1L + b_2L^2 + \dots + b_qL^q = 0.$$

Висновки щодо стаціонарних процесів

1. АКФ процесу АРКС(p,q) починає збігатись до нуля при значеннях зсуву $s = q$. Коефіцієнти АКФ задовольняють різницевому рівнянню:

$$\rho_i = a_1\rho_{i-1} + a_2\rho_{i-2} + \dots + a_p\rho_{i-p}.$$

Оскільки корені характеристичного рівняння знаходяться всередині

одиничного кола на комплексній площині, то коефіцієнти АКФ будуть збігатись до нуля при значеннях зсуву $s \geq q$.

2. ЧАКФ процесу $\text{APKC}(p,q)$ починає збігатись до нуля при $s \geq p$. Починаючи із значення $s = p$ коефіцієнти ЧАКФ за своїм характером відповідають коефіцієнтам АКФ.

Ці властивості кореляційних функцій дозволяють вибрati порядок моделі $\text{APKC}(p,q)$. Так, для деяких конкретних стаціонарних процесів АКФ і ЧАКФ мають такі значення:

- *білий шум*: всі $\rho_s = 0$, $\phi_{ss=0}$;
- *процес AP(1)* при $-1 < a_1 < 0$: $\rho_s = a_1^s$ – експоненціально спадають; $\phi_{11} = \rho_1$; $\phi_{ss} = 0$, $s > 1$;
- *процес AP(p)*: коефіцієнти ρ_s збігаються до нуля, можливі коливання; значення ϕ_{ss} суттєво відрізняються від нуля при $s \leq p$ і $\phi_{ss} = 0$, $s > p$;
- *процес KC(1)* при $0 < b_1 < 1$: ρ_1 має від'ємне значення і $\rho_s = 0$, $s > 1$; $\phi_{ss} < 0$ і мають коливальну збіжність;
- *процес KC(1)* при $-1 < b_1 < 0$: ρ_1 має додатне значення і $\rho_s = 0$, $s > 1$; $\phi_{ss} > 0$ і мають коливальну збіжність;
- *процес APKC(1,1)* при $0 < a_1 < 1$: ρ_s мають гладку збіжність і $\text{sign}(\rho_1) = \text{sign}(a_1 + b_1)$; $\phi_{11} = \rho_1$, далі коливальна збіжність;
- *процес APKC(1,1)* при $-1 < a_1 < 0$: ρ_s мають коливальну збіжність і $\text{sign}(\rho_1) = \text{sign}(a_1 + b_1)$; $\phi_{11} = \rho_1$, далі гладка збіжність; $\text{sign}(\phi_{ss}) = \text{sign}(\phi_{11})$;
- *процес APKC(p,q)*: ρ_s мають гладку або коливальну збіжність, починаючи із значення $s > q$; ϕ_{ss} мають гладку або коливальну збіжність, починаючи з $s > p$.

Означення значущості коефіцієнтів АКФ

Для того щоб коректно визначити порядок моделі $\text{APKC}(p,q)$, необхідно користуватись відповідними статистиками. Для стаціонарного процесу значення r_s мають нормальній розподiл. Оскiльки при визначенні порядку ми користуємося вибiрковими статистичними характеристиками, то для визначення значущості r_s (тобто, суттєвої вiдмiнностi вiд нуля значень АКФ) можна скористатись статистикою Бокса-Пірса або Лiнга-Бокса.

Статистика Бокса-Пірса або Q -статистика розраховується за виразом:

$$Q = N \sum_{i=1}^s r_i^2.$$

Якщо досліджуваний процес АРКС стаціонарний, то статистика Q має розподіл χ^2 з s степенями свободи. Це досить наближена статистика, але її використовують для визначення стаціонарності *довгих вибірок* даних.

Для коротких вибірок запропонована статистика Лунга-Бокса:

$$Q_{LB} = N(N+2) \sum_{i=1}^s r_i^2 / (N-i),$$

яку використовують також при визначенні порядку процесу. Оскільки значення Q пропорціональне $\sum r_i^2$, то менші значенні Q відповідають меншій ступені корельованості досліджуваного процесу.

Запитання і вправи до другого розділу

1. Дайте визначення процесу авторегресії з ковзним середнім? Поясніть можливу природу випадкового процесу $\{\varepsilon(k)\}$ у правій частині рівняння.
2. Які процеси називають інтегрованими? Наведіть приклади інтегрованих процесів.
3. Дайте визначення слабкої стаціонарності? Чим відрізняється слабка стаціональність від сильної?
4. В чому полягає різниця між теоретичними і вибірковими статистичними характеристиками? Як можна скористатись теоретичними статистичними характеристиками?
5. Поясніть різницю між кореляцією та коваріацією? Чи можна користуватись коваріацією з тією ж метою, що використовують кореляцію?
6. Обчисліть коваріаційну функцію $\text{cov}[x(k), x(k-s)]$ для процесу ковзного середнього:

$$x(k) = 0,15\gamma(k-1) + 0,25\gamma(k-2) + 0,35\gamma(k-3) + 0,25\gamma(k-4),$$

де $\gamma(k)$ – випадковий процес білого шуму: $E[\gamma(k)\gamma(l)] = \begin{cases} \sigma_\gamma^2, & k = l; \\ 0, & k \neq l. \end{cases}$

7. Для авторегресійного процесу другого порядку

$$y(k) = a_0 + a_2 y(k-2) + \varepsilon(k)$$

визначте:

- | | | | |
|-----|--------|-------------|-------------|
| (а) | умовні | математичні | сподівання: |
|-----|--------|-------------|-------------|
- $E_{k-2}[y(k)]; E_{k-1}[y(k)]; E_k[y(k+2)];$
- (б) коваріації: $\text{cov}[y(k), y(k-1)]; \text{cov}[y(k), y(k-2)];$
- (в) значення автокореляційної функції: $\rho(1), \rho(2), \rho(3).$
- (г) значення часткової автокореляційної функції ϕ_{11} і $\phi_{22}.$
8. Побудуйте функцію прогнозування $\hat{y}(k+s) = E_k[y(k+s)]$ для процесу АРКС(1,2). Знайдіть вираз для похибок оцінок прогнозів $f(s) = y(k+s) - \hat{y}(k+s)$ і теоретичну автокореляційну функцію для послідовності $\{f(s)\}.$
9. Поясніть умови стаціонарності процесу АР(1)? Наведіть приклад фінансового чи економічного процесу, для якого було б доцільно виконати умови стаціонарності.
10. Яка на вашу думку загальна мета аналізу стаціонарності процесів?
11. Для чого будують систему рівнянь Юла-Уокера?
12. Як можна інтерпретувати однорідний розв'язок Aa_1^k при аналізі стаціонарності процесів АРКС?
13. Яка умова стаціонарності процесу ковзного середнього довільного порядку?
14. Сформулюйте умови стаціонарності для процесу АРКС довільного порядку?
15. При якому значенні параметра зсуву s ЧАКФ процесу АРКС(p,q) починає збігатись до нуля? Починаючи із якого значення коефіцієнти ЧАКФ за своїм характером відповідають коефіцієнтам АКФ?
16. Які статистики використовують для аналізу значущості АКФ?

Розділ 3

ЗАСТОСУВАННЯ РІЗНИЦЕВИХ РІВНЯНЬ ДО ОПИСАННЯ СТАТИСТИЧНИХ ДАНИХ

3.1 Загальні відомості про різницеві рівняння

При використанні дискретних рівнянь незалежну змінну, час t , замінюють дискретним часом, тобто, покладають $t = kT_s$, де T_s – період дискретизації вимірів, який у технічних системах приймає значення від десятків мікросекунд до десятків секунд і навіть хвилин, а при моделюванні фінансово-економічних процесів – від декількох хвилин до одного року, в залежності від того, які статистичні дані можна отримати. Період дискретизації, як правило, нормують до одиниці і незалежною змінною залишається k (дискретний час), яке приймає цілі значення від 0 до ∞ . При цьому для кожної прикладної задачі дискретна одиниця часу має відповідне фактичне значення.

Завдяки простоті структури та наявності надійних методів оцінювання параметрів різницеві рівняння (РР) знайшли надзвичайно широке застосування при моделюванні процесів в техніці, економіці та фінансах, екології, біології та інших прикладних і наукових галузях. Простим прикладом різницевого рівняння є стохастичне рівняння авторегресії першого порядку з одиничним коефіцієнтом (окремий випадок) при затриманому в часі значенні основної змінної:

$$y(k) = y(k-1) + \varepsilon(k), \quad (3.1.1)$$

де $y(k)$ – основна змінна; $\varepsilon(k)$ – випадкова величина, яка відображає вплив різноманітних невимірюваних факторів на основну змінну. В першу чергу це випадкові збурення, що діють на процес. Частіше всього припускають, що випадкова величина має нормальній розподіл: $\{\varepsilon(k)\} \sim N_n(0, \sigma_\varepsilon^2)$, тобто вона має нульове середнє та скінчену дисперсію σ_ε^2 . Наприклад, у випадку коли $y(k)$ – ціни на деякі біржові акції в k -й день, то $\varepsilon(k)$ характеризує коливання ціни під впливом багатьох факторів, які неможливо ввести в

модель, оскільки неможливо отримані їх виміри, або ж ці впливи носять якісний характер. Як правило, у наведеній моделі процес $\{\varepsilon(k)\}$ має

наступні обмеження: $E[\varepsilon(k)] = 0$, $E[\varepsilon(k)\varepsilon(l)] = \begin{cases} \sigma_\varepsilon^2, & k = l, \\ 0, & k \neq l. \end{cases}$. Більш докладно

роль випадкової змінної буде розглядатись окремо у кожному конкретному випадку.

За допомогою рівняння (3.1.1) описують, наприклад, ціну акції на біржі в момент часу, що відповідає аргументу k . Його називають ще рівнянням, яке описує *процес випадкового кроку* (випадкове блукання або *random walk*). Таку назву воно отримало з тієї причини, що приріст значення основної змінної визначається, фактично, випадковим процесом. Воно може бути записане також у формі першої різниці

$$\Delta y(k) = \varepsilon(k),$$

де $\Delta y(k) = y(k) - y(k - 1)$. Саме можливості використання різниць перших і вищих порядків для опису часових рядів зумовило використання назви *різницеві рівняння*.

Для рівняння випадкового кроку можна записати однорідне

$$y(k) - y(k - 1) = 0,$$

яке має характеристичне рівняння: $\alpha - 1 = 0$ (як отримати характеристичне рівняння в загальному випадку, ми розглянемо нижче в даному розділі). Таким чином, дане характеристичне рівняння має один корінь $\alpha = 1$. Якщо характеристичне рівняння має хоча б один одиничний корінь, то кажуть, що відповідне йому різницеве рівняння описує *процес з одиничним коренем*. Нижче покажемо, що процеси з одиничними коренями – це *процеси з трендами* або *інтегровані* процеси. В подальшому під трендом будемо розуміти загальний довгостроковий напрям розвитку процесу; фактично він співпадає з поточним середнім значенням.

Більш загальною формою різницевого рівняння (3.1.1) є наступна:

$$\Delta y(k) = \alpha_0 + \alpha_1 y(k - 1) + \varepsilon(k), \quad (3.1.2)$$

але для того, щоб воно відповідало процесу випадкового кроку необхідно покласти $\alpha_0 = \alpha_1 = 0$, інакше це рівняння вже не буде відповідати своєму визначенню. Зазначимо, що порядок авторегресійної частини різницевого рівняння визначається числом минулих (затриманих) вимірів залежної змінної, які використовуються в правій частині для пояснення її зміни в часі.

Різницеві рівняння, у правій частині яких наявні минулі (затримані)

виміри основної змінної, називають *авторегресійними* (AP), тобто регресія змінної на саму себе. Рівняння авторегресії n -го порядку має вигляд:

$$y(k) = a_0 + \sum_{i=1}^n a_i y(k-i) + \varepsilon(k). \quad (3.1.3)$$

Якщо для процесу, що моделюється, можна виявити вхідну змінну, то вона вона записується в правій частині

$$y(k) = a_0 + \sum_{i=1}^n a_i y(k-i) + \sum_{j=1}^q b_j x(k-j) + \varepsilon(k). \quad (3.1.3)$$

Якщо $x(k)$ – випадковий процес, то таке рівняння називають авторегресією із ковзним середнім (APKC). Формально, у випадку ковзного середнього необхідно також виконати умову: $\sum_{j=1}^q b_j = 1$.

Характеристичне рівняння, записане для (3.1.3), може мати одиничні корені, тобто, один або більше коренів характеристичного рівняння можуть приймати значення «1». Такі процеси називають процесами авторегресії з інтегрованим ковзним середнім – APIKC(p, d, q), де d – число одиничних коренів характеристичного рівняння. Процеси цього класу *нестаціонарні* – вони мають тренд, порядок якого визначається числом одиничних коренів. Якщо $d = 1$, то тренд лінійний; якщо $d = 2$, то тренд квадратичний і т.д. Таким чином, *процеси з трендом, інтегровані процеси і процеси з одиничними коренями* – це різні назви процесів, що мають тренд.

Застосування перших та різниць вищих порядків

Разом із основною змінною для описання процесів використовують *перші та другі різниці*, наприклад:

$$\begin{aligned}\Delta y(k) &= y(k) - y(k-1), \\ \Delta y(k+1) &= y(k+1) - y(k), \\ \Delta y(k+2) &= y(k+2) - y(k+1).\end{aligned}$$

Наведені перші різниці відображають швидкість зміни основної змінної, що відповідає першій похідній для рівнянь, записаних у неперервному часі, тобто для дифференціальних рівнянь.

Обчислення перших різниць приводить до видалення лінійного тренду з процесу. Наприклад, нехай тренд описується лінійним рівнянням:

$$y(k) = a_0 + a_1 k.$$

Перша різниця для цього процесу:

$$\Delta y(k) = y(k) - y(k-1) = a_0 + a_1 k - a_0 - a_1 \cdot (k-1) = a_1,$$

тобто, після дискретного диференціювання залишилась константа.

Другі різниці відображають швидкість зміни у часі перших різниць (тобто прискорення) і записуються так:

$$\begin{aligned} \Delta^2 y(k) &= \Delta(\Delta y(k)) = \Delta[y(k) - y(k-1)] = [y(k) - y(k-1)] - [y(k-1) - y(k-2)] = \\ &= y(k) - 2y(k-1) + y(k-2), \end{aligned}$$

$$\Delta^2 y(k+1) = \Delta(\Delta y(k+1)) = y(k+1) - 2y(k) + y(k-1).$$

Останній вираз називають різницевою схемою другої похідної. На практиці другі різниці використовують досить рідко, а різниці вищого порядку не використовуються. Застосування других різниць до процесу приводить до видалення квадратичного тренду, що можна легко проілюструвати на прикладі описання тренду поліномом другого порядку від k . Наприклад, нехай:

$$y(k) = a_0 + a_1 k + a_2 k^2.$$

Перші різниці:

$$\begin{aligned} \Delta y(k) &= a_0 + a_1 k + a_2 k^2 - [a_0 + a_1 \cdot (k-1) + a_2 \cdot (k-1)^2] = \\ &= a_0 + a_1 k + a_2 k^2 - a_0 - a_1 k + a_1 - a_2 k^2 + 2a_2 k - a_2 = \\ &= a_1 + 2a_2 k - a_2. \end{aligned}$$

Другі різниці:

$$\begin{aligned} \Delta^2 y(k) &= a_1 + 2a_2 k - a_2 - [a_1 + 2a_2 \cdot (k-1) - a_2] = \\ &= 2a_2. \end{aligned}$$

Редукована та структурована форми різницевого рівняння. Для ілюстрації рівнянь такого типу розглянемо модель Самуельсона для валового внутрішнього продукту (ВВП), споживання та інвестицій:

$$y(k) = c(k) + I(k), \quad (3.1.4)$$

$$c(k) = \alpha y(k-1) + \varepsilon_c(k), \quad 0 < \alpha < 1, \quad (3.1.5)$$

$$I(k) = \beta [c(k) - c(k-1)] + \varepsilon_I(k), \quad \beta > 0, \quad (3.1.6)$$

де $y(k), c(k), I(k)$ – ВВП, споживання та інвестиції в момент k відповідно. В цій моделі $y(k), c(k), I(k)$ – ендогенні змінні, а $\varepsilon_c(k), \varepsilon_I(k)$ – стохастичні змінні з нульовим середнім, які характеризують вплив збурень на споживання та інвестиції, тобто вплив тих змінних, які не введені в модель. Третє рівняння ілюструє принцип прискорення (акселерації), оскільки в нього введена швидкість зміни споживання $[c(k) - c(k-1)]$.

Рівняння (3.1.6) називають **структураним рівнянням**, оскільки в ньому ендогенна змінна $I(k)$ залежить від поточного (а не минулого) значення іншої ендогенної змінної $c(k)$.

Редукованою формою різницевого рівняння називають таку, в якій основна змінна (в правій частині) залежить від минулих вимірів цієї ж змінної, минулих вимірів інших ендогенних змінних, поточних та минулих вимірів екзогенних змінних та збурень.

Таким чином, рівняння для споживання вже представлене у редукованій формі. Рівняння для інвестицій не відповідає редукованій формі, оскільки в правій частині присутнє поточне значення споживання. Для того щоб отримати редуковану форму для інвестицій, підставимо (3.1.5) в рівняння для інвестицій і отримаємо

$$\begin{aligned} I(k) &= \beta [\alpha y(k-1) + \varepsilon_c(k) - c(k-1)] + \varepsilon_I(k) = \\ &= \alpha \beta y(k-1) - \beta c(k-1) + \beta \varepsilon_c(k) + \varepsilon_I(k). \end{aligned}$$

Необхідно зазначити, що отримана редукована форма для інвестицій не єдина. Якщо змінну дискретного часу k зменшити на одиницю, то рівняння (3.1.3) можна переписати у вигляді: $c(k-1) = \alpha y(k-2) + \varepsilon_c(k-1)$. За допомогою отриманого рівняння запишемо редуковану форму для інвестицій у вигляді

$$\begin{aligned} I(k) &= \alpha \beta y(k-1) - \beta [\alpha y(k-2) + \varepsilon_c(k-1)] + \beta \varepsilon_c(k) + \varepsilon_I(k) = \\ &= \alpha \beta [y(k-1) - y(k-2)] + \beta [\varepsilon_c(k) - \varepsilon_c(k-1)] + \varepsilon_I(k). \quad (3.1.7) \end{aligned}$$

Аналогічно можна отримати редуковану форму для ВНП, якщо підставити (3.1.5) і (3.1.7) в рівняння (3.1.4):

$$\begin{aligned} y(k) &= \alpha y(k-1) + \varepsilon_c(k) + \alpha \beta [y(k-1) - y(k-2)] + \beta [\varepsilon_c(k) - \varepsilon_c(k-1)] + \varepsilon_I(k) = \\ &= \alpha(1 + \beta) y(k-1) - \alpha \beta y(k-2) + (1 + \beta) \varepsilon_c(k) + \varepsilon_I(k) - \beta \varepsilon_c(k-1). \quad (3.1.8) \end{aligned}$$

Рівняння (3.1.8) називають редукованою формою стосовно однієї змінної, оскільки в правій частині присутні тільки минулі виміри основної змінної та збурення. Така форма зручна для прогнозування, оскільки для цього необхідно мати виміри тільки однієї змінної. Нижче будуть розглянуті підходи до прогнозування, які ґрунтуються на розв'язку різницевих рівнянь. Вони дозволяють обчислювати прогноз на довільне число кроків вперед.

Необхідно також зазначити, що різницеві рівняння, які описують процеси перехідного періоду, можна використати для створення систем оптимального управління цими процесами, при розробці інформаційних систем підтримки прийняття оптимальних (чи субоптимальних) рішень, при створенні експертних систем та ін. Позитивною стороною використання РР є те, що їхні параметри легко обновляються при поступанні нових вимірів, також можна досить просто змінити порядок і структуру математичної моделі. При цьому під **структурою моделі** будемо розуміти *кількість рівнянь* (вимірність) моделі, *їх порядок, наявність нелінійностей* та їхній тип, *наявність запізнення (лагу)* відносно входу та його оцінка, а також *збурення процесу і їх тип*.

3.2 Ітераційний метод знаходження розв'язків різницевих рівнянь

Поняття розв'язку різницевого рівняння

Розв'язки різницевих рівнянь (так само як і диференціальних) необхідні для аналізу асимптотичної поведінки процесу, для формування на їх основі функцій прогнозування, а також для порівняння характеристик різних процесів між собою.

Розв'язок різницевого рівняння представляє собою *функцію часу, вхідної змінної, початкових умов та збурення*. Існують різні підходи до розв'язку РР.

Як простий приклад знаходження розв'язку розглянемо процес, який описується першою різницею

$$\Delta y(k) = 2,5, \quad (3.2.1)$$

тобто, приріст змінної за один період дискретизації складає 1,5 одиниці у вибраному масштабі. Замінимо першу різницю відповідними значеннями змінної:

$$y(k) = y(k-1) + 2,5$$

і задамо початкову умову $y(0) = y_0 = 0,25$. Тепер запишемо значення $y(k)$ для кількох моментів часу, починаючи з $k=1$:

$$y(1) = y_0 + 2,5,$$

$$y(2) = y(1) + 2,5 = y_0 + 2,5 + 2,5,$$

$$y(3) = y(2) + 2,5 = y_0 + 2,5 + 2,5 + 2,5 = y_0 + 3 \cdot 2,5$$

і для довільного k :

$$y(k) = y_0 + 2,5 \cdot k = 2,5k + 0,25. \quad (3.2.2)$$

Отриманий розв'язок для (3.2.1) справедливий для будь-якого значення k . Для того щоб це довести, підставимо отриманий розв'язок в рівняння (3.2.1):

$$2,5k + 0,25 = 2,5(k-1) + 0,25 + 2,5$$

або

$$0,25 = 0,25.$$

Таким чином, знайдений розв'язок справедливий для $\forall k, k = 0, \dots, \infty$. Він представляє собою рівняння прямої, яка перетинає вісь ординат в точці $y_0 = 0,25$. В зв'язку з цим значення $y_0 = a_0$ називають ще *перетином*.

Ітераційний підхід до знаходження розв'язку рівняння першого порядку

Розглянемо знаходження розв'язку РР першого порядку ітераційним методом. Нехай рівняння першого порядку має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k) \quad (3.2.3)$$

з початковою умовою y_0 . Для моментів часу $k=1,2,3$ можна записати, що

$$y(1) = a_0 + a_1 y(0) + \varepsilon(1),$$

$$\begin{aligned} y(2) &= a_0 + a_1 y(1) + \varepsilon(2) = a_0 + a_1 [a_0 + a_1 y(0) + \varepsilon(1)] + \varepsilon(2) = \\ &= a_0 + a_0 a_1 + (a_1)^2 y(0) + a_1 \varepsilon(1) + \varepsilon(2). \end{aligned}$$

$$y(3) = a_0 + a_1 y(2) + \varepsilon(3) =$$

$$= a_0 [1 + a_1 + (a_1)^2] + (a_1)^3 y(0) + a_1^2 \varepsilon(1) + a_1 \varepsilon(2) + \varepsilon(3).$$

Тепер можна записати розв'язок рівняння (3.2.3) для довільного моменту k :

$$y(k) = a_0 \sum_{i=0}^{k-1} a_1^i + a_1^k y_0 + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i). \quad (3.2.4).$$

Розв'язок (3.2.4) також можна отримати, якщо розпочати ітерації в зворотному часі, тобто

$$\begin{aligned} y(k) &= a_0 + a_1 [a_0 + a_1 y(k-2) + \varepsilon(k-1)] + \varepsilon(k) = \\ &= a_0 + (1 + a_1) + a_1^2 [a_0 + a_1 y(k-3) + \varepsilon(k-2)] + \varepsilon(k) + a_1 \varepsilon(k-1). \end{aligned}$$

Продовження ітераційного процесу приведе до розв'язку (3.2.4).

Повний розв'язок (3.2.4) складається з двох частин. Перша частина залежить від початкової умови $a_1^k y_0$, а друга від наявності в рівнянні перетину a_0 і випадкової змінної $\varepsilon(k)$. Відповідно, першу частину називають однорідним розв'язком або розв'язком однорідного рівняння і позначають так:

$$y^h(k) = a_1^k y_0 = A \alpha^k,$$

де індекс h – *homogeneous* (однорідний). В більш загальному вигляді однорідний розв'язок записують так: $y^h(k) = A \alpha^k$, де A – довільна константа, яка визначається за допомогою початкових умов; α – корінь характеристичного рівняння, який дорівнює a_1 у випадку рівняння $AP(1)$.

Неоднорідний або частковий розв'язок позначають так:

$$y^p(k) = a_0 \sum_{i=0}^{k-1} a_1^i + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i),$$

де верхній індекс p – *partial* (частковий). Повний розв'язок представляє собою суму однорідного і часткового розв'язків, тобто,

$$y(k) = y^h(k) + y^p(k).$$

Якщо початкова умова не задана, то замість y_0 в (3.2.4) можна підставити $y(0) = a_0 + a_1 y(-1) + \varepsilon(0)$ (випадок не дуже реальний, але теоретично можливий):

$$y(k) = a_0 \sum_{i=0}^{k-1} a_1^i + a_1^k [a_0 + a_1 y(-1) + \varepsilon(0)] + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i) =$$

$$= a_0 \sum_{i=0}^k a_1^i + \sum_{i=0}^k a_1^i \varepsilon(k-i) + a_1^{k+1} y(-1). \quad (3.2.5)$$

Продовжуючи ітерації в зворотному напрямі на m кроків, отримаємо

$$y(k) = a_0 \sum_{i=0}^{k+m} a_1^i + \sum_{i=0}^{k+m} a_1^i \varepsilon(k-i) + a_1^{k+m-1} y(-m-1). \quad (3.2.6)$$

При $|a_1| < 1$ член a_1^{k+m+1} збігається до нуля, якщо $m \rightarrow \infty$, а нескіченна сума $[1 + a_1 + a_1^2 + \dots]$ збігається до виразу $1/(1 - a_1)$ як сума членів геометричної прогресії. Таким чином, розв'язок (3.2.6) на нескінченності має вигляд

$$y(k) = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i), \quad |a_1| < 0, \quad k \rightarrow \infty. \quad (3.2.7)$$

Для того щоб перевірити правильність розв'язку, необхідно підставити (3.2.7) в (3.2.3). В результаті отримаємо тотожність. Однак, (3.2.7) – це не єдиний розв'язок для (3.2.3). Розв'язком є також наступне рівняння:

$$y(k) = \frac{a_0}{1 - a_1} + A a_1^k + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i), \quad |a_1| < 0, \quad (3.2.8)$$

де A – довільна константа. Підставимо (3.2.8) в (3.2.3) і отримаємо тотожність: що свідчить про правильність розв'язку:

$$\frac{a_0}{1 - a_1} + A a_1^k + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) = a_0 + a_1 \left[\frac{a_0}{1 - a_1} + A a_1^{k-1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-1-i) \right] + \varepsilon(k).$$

З наведеного вище розв'язку рівняння першого порядку видно, що коефіцієнт a_1 відіграє важливу роль в поведінці розв'язку. Якщо $|a_1| > 1$, то розв'язок буде розбіжним. Збіжність чи розбіжність буде мати коливальний характер, якщо $a_1 < 0$; це зумовлено тим, що $a_1^k > 0$ при парному k і $a_1^k < 0$ при непарному k .

Альтернативний підхід до знаходження розв'язків

Існує **альтернативний підхід до знаходження розв'язків різницевих рівнянь**, який полягає у знаходженні окремо розв'язку однорідного рівняння та часткового розв'язку. Наприклад, розв'язок однорідного рівняння виду

$$y(k) = a_1 y(k-1) \quad \text{або} \quad y(k) = a_1 y(k-1) + a_2 y(k-2)$$

називають *однорідним розв'язком*. Тривіальним розв'язком однорідного рівняння є $y(1) = y(2) = \dots = 0$. Якщо покласти $a_0 = 0$ і всі значення випадкової змінної $\{\varepsilon(k)\} = 0$, то із (3.1.9) випливає, що розв'язком рівняння буде також $y(k) = a_1^k y_0$. Можна показати, що $y(k) = A a_1^k$, де A – довільна константа, також буде розв'язком однорідного рівняння. Підставимо останнє значення в однорідне рівняння і отримаємо тотожність:

$$A(a_1)^k = a_1 A(a_1)^{k-1}. \quad (3.2.9)$$

Із наведеного вище аналізу розв'язку рівняння першого порядку, а також з (3.2.9) випливає, що однорідний розв'язок має наступні властивості:

1. Якщо $|a_1| < 1$, то вираз $(a_1)^k$ збігається до нуля при $k \rightarrow \infty$. Збіжність має гладкий характер, якщо $0 < a_1 < 1$, і коливальний характер при $-1 < a_1 < 0$.

2. Якщо $|a_1| > 1$, то однорідний розв'язок буде нестійким (тобто, спостерігається розбіжність). Якщо $a_1 > 1$, то однорідний розв'язок наближається до нескінченності при $k \rightarrow \infty$ без коливань, а якщо $a_1 < -1$, то розбіжність носить коливальний характер.

3. Якщо $a_1 = 1$, то однорідному рівнянню $y(k) = a_1 y(k-1)$ задовольняє будь-яка довільна константа A . Якщо ж $a_1 = -1$, то розв'язок називають *метастійким* — $(a_1)^k = 1$ для парних значень k , і $(a_1)^k = -1$ при непарних значеннях k .

Рівняння (3.2.7) називають *частковим* (або *неоднорідним*) *розв'язком* різницевого рівнянням. Ця частина загального розв'язку зумовлена наявністю в правій частині рівняння членів a_0 та $\varepsilon(k)$. Процес знаходження часткового розв'язку буде розглянуто нижче, а зараз наведемо загальну методику розв'язання різницевого рівняння, яка логічно випливає з отриманого ітераційним методом повного розв'язку рівняння $AP(1)$.

Загальна методика знаходження розв'язку різницевого рівняння

Користуючись результатами знайденого розв'язку рівняння першого порядку, отриманими вище, можна сформулювати загальну методику знаходження розв'язків різницевих рівнянь типу АР та АРКС:

Крок 1: Для моделі АРКС(p, q) сформувати однорідне рівняння та знайти всі його однорідні розв'язки (їх число дорівнює p).

Крок 2: Знайти всі складові часткового розв'язку, зумовлені константами, детермінованими та випадковими функціями у правій частині рівняння.

Крок 3: Записати загальний розв'язок у вигляді суми однорідного та часткового розв'язків.

Крок 4: За допомогою початкових умов знайти значення довільних констант.

Розглянемо приклад знаходження повного розв'язку рівняння авторегресії другого порядку.

Приклад 3.1. Знайдемо загальний розв'язок РР другого порядку $y(k) = 2,1 + 0,8y(k-1) - 0,15y(k-2)$.

Сформуємо однорідне рівняння:

$$y(k) - 0,8y(k-1) + 0,15y(k-2) = 0.$$

Оскільки це рівняння другого порядку, то воно має два однорідних розв'язки, які знаходять за допомогою розв'язку характеристичного рівняння, записаного для однорідного (більш детально ця методика буде розглянута нижче). Характеристичне рівняння має вигляд:

$$\lambda^2 - 0,8\lambda + 0,15 = 0.$$

Це квадратне рівняння має два корені: $\lambda_1 = 0,5$; $\lambda_2 = 0,3$. Таким чином, два однорідних розв'язки можна записати як $y_1^h(k) = A_1(0,5)^k$; $y_2^h(k) = A_2(0,3)^k$.

Правильність знаходження однорідних розв'язків можна перевірити шляхом підстановки в однорідне рівняння. Наприклад:

$$A_1(0,5)^k - 0,8A_1(0,5)^{k-1} + 0,15A_1(0,3)^{k-2} = 0.$$

Якщо розділити всі члени на $(0,3)^{k-2}$, то отримаємо

$$(0,3)^2 - 0,9(0,3) + 0,15 = 0,09 - 0,27 + 0,15 = 0.$$

Другий однорідний розв'язок перевіряється аналогічно.

Частковий розв'язок для детермінованого збурення знайдемо так:

$$y^P = 2,1 + 0,8y^P - 0,15y^P.$$

Звідси $y^P = 6,0$, де верхній індекс $p=partial$ (частковий). Загальна методика визначення часткового розв'язку буде наведена нижче. Тепер об'єднаємо однорідний та частковий розв'язок:

$$y(k) = 6,0 + A_1(0,5)^k + A_2(0,3)^k,$$

де A_1, A_2 – довільні константи, які знайдемо за допомогою початкових умов. Оскільки маємо дві невідомих константи, то необхідно мати дві умови. Нехай $y(0) = 1,0$; $y(1) = 2,0$. Тепер можемо записати систему із двох рівнянь для знаходження констант:

$$k=0: 1,0 = 6,0 + A_1 + A_2,$$

$$k=1: 2,0 = 6,0 + 0,5A_1 + 0,3A_2$$

Розв'язуючи систему, знайдемо, що $A_1 = -12,5$; $A_2 = 7,5$, і розв'язок приймає вигляд

$$y(k) = 6,0 - 12,5(0,5)^k + 7,5(0,3)^k.$$

Таким чином, знайдений розв'язок має середнє значення (його називають ще зміщенням) 6,5, а два інші члени збігаються до нуля при $k \rightarrow \infty$.

3.3 Знаходження однорідних та часткових розв'язків

3.3.1 Розв'язування однорідних рівнянь

Однорідне рівняння другого порядку

При моделюванні економічних процесів найчастіше використовують РР 1-3 порядку, але зовсім не виключена можливість використання рівнянь вищих порядків. У зв'язку з цим виникає необхідність знаходження розв'язків однорідних рівнянь різних порядків. Оскільки розв'язок рівняння другого порядку дає можливість зробити деякі узагальнення на рівняння вищих порядків, то розглянемо спочатку рівняння другого порядку:

$$y(k) - a_1y(k-1) - a_2y(k-2) = 0. \quad (3.3.1)$$

По аналогії з рівнянням першого порядку виберемо пробний однорідний розв'язок у вигляді $y^h(k) = A\alpha^k$. Підставимо пробний розв'язок в рівняння (3.3.1):

$$A\alpha^k - a_1 A\alpha^{k-1} - a_2 A\alpha^{k-2} = 0 \quad (3.3.2)$$

і поділимо обидві частини отриманого рівняння на $A\alpha^{k-2}$. В результаті отримаємо так зване *характеристичне рівняння* відносно α :

$$\alpha^2 - a_1\alpha - a_2 = 0. \quad (3.3.3)$$

Запишемо корені цього квадратного рівняння:

$$\alpha_{1,2} = \frac{a_1 \pm \sqrt{a_1^2 + 4a_2}}{2} = \frac{a_1 \pm \sqrt{d}}{2}, \quad (3.3.4)$$

де дискримінант $d = \sqrt{a_1^2 + 4a_2}$. Кожний корінь характеристичного рівняння дає один однорідний розв'язок. Розв'язком є також лінійна комбінація

$$A_1\alpha_1^k + A_2\alpha_2^k, \quad (3.3.5)$$

що легко перевірити підстановкою (3.3.5) в (3.3.1):

$$A_1\alpha_1^k + A_2\alpha_2^k = a_1(A_1\alpha_1^{k-1} + A_2\alpha_2^{k-1}) + a_2(A_1\alpha_1^{k-2} + A_2\alpha_2^{k-2}).$$

Перенесемо праву частину вліво і згрупуємо члени:

$$A_1(\alpha_1^k - a_1\alpha_1^{k-1} - a_2\alpha_1^{k-2}) + A_2(\alpha_2^k - a_1\alpha_2^{k-1} - a_2\alpha_2^{k-2}) = 0.$$

Оскільки α_1, α_2 – це корені характеристичного рівняння (3.3.3), то обидва члени в дужках дорівнюють нулю. Таким чином, повний однорідний розв'язок РР другого порядку має вигляд:

$$y^h(k) = A_1\alpha_1^k + A_2\alpha_2^k. \quad (3.3.6)$$

Оскільки *рівняння другого порядку* – це основа для аналізу розв'язків рівнянь вищих порядків, в літературі розглядають три можливих випадки знаходження їх розв'язку в залежності від значення дискримінанта:

$d > 0$, $d = 0$ і $d < 0$. Крім того, рівняння другого порядку – досить поширенна модель економічних, фінансових, технічних та інших процесів.

A. Дискримінант додатний: $a_1^2 + 4a_2 > 0$. В такому випадку корені будуть різними дійсними числами. Однорідний розв'язок має вигляд:

$$y^h(k) = A_1 \alpha_1^k + A_2 \alpha_2^k.$$

Очевидно, що при $|\alpha_1| > 1$ або $|\alpha_2| > 1$ розв'язок буде розбіжним.

Приклад 3.2. Розглянемо процес другого порядку

$$y(k) = 0,95y(k-1) + 0,35y(k-2),$$

тобто, $a_1 = 0,95$; $a_2 = 0,35$. Для однорідного рівняння $y(k) - a_1 y(k-1) - a_2 y(k-2) = 0$ характеристичним є наступне: $\alpha^2 - 0,95\alpha - 0,35 = 0$. Дискримінант $d = a_1^2 + 4a_2 = 0,95^2 + 4 \cdot 0,35 = 2,3025$. Корені характеристичного рівняння:

$$\alpha_1 = \frac{1}{2}(a_1 + \sqrt{d}) = \frac{1}{2}(0,95 + 1,517) = 1,233; \quad \alpha_2 = \frac{1}{2}(0,95 - 1,517) = -0,284.$$

Таким чином, однорідний розв'язок має вигляд:

$$y^h(k) = A_1 \cdot 1,233^k + A_2 \cdot (-0,284)^k,$$

що свідчить про розбіжність розв'язку.

B. Дискримінант $a_1^2 + 4a_2 = 0$ (випадок кратних коренів). При $d = 0$ отримаємо два одинакових корені: $\alpha_1 = \alpha_2 = a_1 / 2$. Таким чином, однорідним розв'язком є $A \cdot (a_1 / 2)^k$. Однак, при $d = 0$ існує також однорідний розв'язок вигляду: $A \cdot k(a_1 / 2)^k$. Покажемо, що це так шляхом його підстановки в рівняння:

$$k(a_1 / 2)^k - a_1[(k-1)(a_1 / 2)^{k-1}] - a_2[(k-2)(a_1 / 2)^{k-2}] = 0.$$

Для зручності аналізу поділимо обидві частини цього рівняння на $(a_1 / 2)^{k-2}$ і

отримаємо:

$$-[(a_1^2/4) + a_2]k + [(a_1^2/2) + 2a_2] = 0.$$

Оскільки $a_1^2 + 4a_2 = 0$, то обидва вирази в дужках будуть дорівнювати нулю і, таким чином, $k(a_1/2)^k$ є розв'язком рівняння другого порядку при $d = 0$. Тепер можемо записати однорідний розв'язок для даного випадку як

$$y^h(k) = A_1 \left(\frac{a_1}{2} \right)^k + A_2 k \left(\frac{a_1}{2} \right)^k.$$

Цей розв'язок буде розбіжним, якщо $|a_1| > 2$. Якщо $|a_1| < 2$, то складова розв'язку $A_1(a_1/2)^k$ буде збіжною, але поведінка складової $A_2 k(a_1/2)^k$ буде невизначеною, тому що в ній присутній множник k . Знайдемо границю, до якої прямує другий однорідний розв'язок при $k \rightarrow \infty$. Введемо позначення:

$$\beta = \frac{1}{a_1/2} = \frac{1}{\alpha}.$$

Тепер скористаємось правилом Лопіталя:

$$\lim_{k \rightarrow \infty} k \alpha^k = \lim_{k \rightarrow \infty} \frac{k}{\beta^k} = \lim_{k \rightarrow \infty} \frac{1}{\beta^k \ln \beta} = 0,$$

оскільки $|\beta| > 1$ при $|\alpha| < 1$.

Якщо виконати аналіз розв'язку в даному випадку, то виявиться, що він має немонотонний характер – деякий сплеск на перших 20-30 кроках з подальшою збіжністю до нуля при $k \rightarrow \infty$.

При $-2 < a_1 < 0$ розв'язок має коливальний характер з досить великою амплітудою на перших 20-30 кроках (додатна амплітуда при парних k і від'ємна при непарних). Однак він також збігається до нуля при $k \rightarrow \infty$, тобто в даному випадку процес збіжний, незважаючи на наявність в розв'язку дискретного часу k в явній формі.

В. Дискримінант $a_1^2 + 4a_2 < 0$. В даному випадку характеристичне рівняння має два комплексно спряжених корені. Оскільки $a_1^2 \geq 0$ завжди, то комплексні корені можуть з'явитись тільки у випадку, коли $a_2 < 0$.

Таким чином, корені характеристичного рівняння можна записати у вигляді:

$$\alpha_1 = \frac{1}{2}(a_1 + i\sqrt{d}), \quad \alpha_2 = \frac{1}{2}(a_1 - i\sqrt{d}), \quad \text{де } i = \sqrt{-1}.$$

При знаходженні розв'язку однорідного рівняння скористаємось наступними тригонометричними тотожностями:

$$\begin{aligned}\sin(\theta_1 + \theta_2) &= \sin(\theta_1)\cos(\theta_2) + \cos(\theta_1)\sin(\theta_2); \\ \cos(\theta_1 + \theta_2) &= \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2),\end{aligned}$$

які при $\theta_1 = \theta_2$ спрощуються до вигляду:

$$\begin{aligned}\sin(2\theta) &= 2\sin(\theta)\cos(\theta); \\ \cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta).\end{aligned}$$

Нагадаємо, що комплексне число $a + ib$ можна представити точкою на комплексній площині (рис.3.1) і перейти до представлення коренів характеристичного рівняння в полярних координатах.

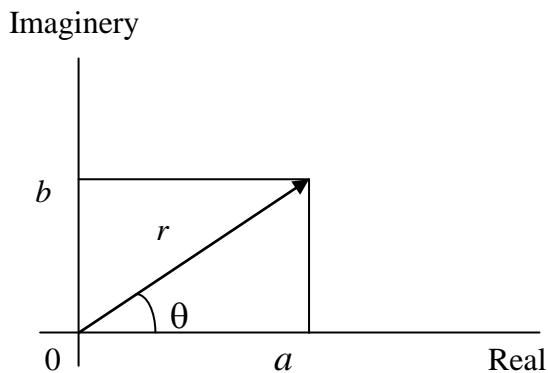


Рис.3.1. Представлення комплексного числа на площині

Запишемо комплексне число через тригонометричні функції кута θ і отримаємо його представлення в полярних координатах:

$$a = r \cos(\theta), \quad b = r \sin(\theta).$$

По відношенню до виразів для коренів однорідного рівняння α_1, α_2 можна записати, що $a = a_1/2, b = \sqrt{d}/2$. Таким чином, для коренів α_1, α_2 можна записати вирази в полярних координатах:

$$\begin{aligned}\alpha_1 &= a + ib = r [\cos(\theta) + i \sin(\theta)], \\ \alpha_2 &= a - ib = r [\cos(\theta) - i \sin(\theta)].\end{aligned}$$

Для того щоб сформувати розв'язок, необхідно отримати вирази для

α_1^k і α_2^k . Спочатку запишемо вираз для α_1^2 із врахуванням того, що $i^2 = -1$:

$$\begin{aligned}\alpha_1^2 &= \{r[\cos(\theta) + i\sin(\theta)]\} \{r[\cos(\theta) + i\sin(\theta)]\} = \\ &= r^2 [\cos(\theta)\cos(\theta) - \sin(\theta)\sin(\theta) + 2i\sin(\theta)\cos(\theta)] = \\ &= r^2 [\cos(2\theta) + i\sin(2\theta)].\end{aligned}$$

Якщо продовжити процес піднесення до степеня, то отримаємо такі вирази для коренів:

$$\alpha_1^k = r^k [\cos(k\theta) + i\sin(k\theta)], \quad \alpha_2^k = r^k [\cos(k\theta) - i\sin(k\theta)].$$

Оскільки однорідний розв'язок $y^h(k)$ – це дійсне число, а α_1, α_2 – комплексні числа, то довільні константи A_1, A_2 – повинні бути спряженими комплексними числами в наступній формі:

$$A_1 = B [\cos(\varphi) + i\sin(\varphi)], \quad A_2 = B [\cos(\varphi) - i\sin(\varphi)],$$

де B, φ – довільні дійсні числа.

Тепер можна обчислити добутки $A_1\alpha_1^k$ і $A_2\alpha_2^k$. Так,

$$\begin{aligned}A_1\alpha_1^k &= B[\cos(\varphi) + i\sin(\varphi)]r^k [\cos(k\theta) + i\sin(k\theta)] = \\ &= Br^k [\cos(\varphi)\cos(k\theta) - \sin(\varphi)\sin(k\theta) + i\cos(k\theta)\sin(\varphi) + i\sin(k\theta)\cos(\varphi)].\end{aligned}$$

Використовуючи наведені вище тригонометричні тотожності, отримаємо:

$$A_1\alpha_1^k = Br^k [\cos(k\theta + \varphi) + i\sin(k\theta + \varphi)].$$

По аналогії можна отримати вираз для другого добутку:

$$A_2\alpha_2^k = Br^k [\cos(k\theta + \varphi) - i\sin(k\theta + \varphi)].$$

Запишемо однорідний розв'язок як суму двох розв'язків:

$$\begin{aligned}y^h(k) &= Br^k [\cos(k\theta + \varphi) + i\sin(k\theta + \varphi)] + Br^k [\cos(k\theta + \varphi) - i\sin(k\theta + \varphi)] = \\ &= 2Br^k \cos(k\theta + \varphi).\end{aligned}$$

Якщо покласти $2B = \beta_1$ і $\theta = \omega$, то розв'язок прийме вигляд:

$$y^h(k) = \beta_1 r^k \cos(\omega k + \varphi),$$

де β_1, φ – довільні константи; $r = \sqrt{(-a_2)}$; значення ω вибирається із умови:

$$\cos(\omega) = \frac{a_1}{2(-a_2)^{1/2}} = \frac{a_1}{2r}.$$

З фізичної точки зору величина ω має смисл кругової частоти, тобто, $\omega = 2\pi f$ (f – лінійна частота в герцах), а φ – початкова фаза коливань розв'язку. Тобто, однорідний розв'язок має, в даному випадку, форму затухаючих або розбіжних коливань в залежності від значення r .

Наведений вище вираз для r можна знайти так:

$$r = \sqrt{a^2 + b^2} = \sqrt{\left(\frac{a_1}{2}\right)^2 + \left(\frac{\sqrt{|d|}}{2}\right)^2} = \sqrt{\frac{a_1^2 + [-(a_1^2 + 4a_2)]}{4}} = \sqrt{-a_2}.$$

Розглянемо приклад знаходження однорідного розв'язку у випадку комплексних коренів.

Приклад 3.3. Нехай рівняння другого порядку має вигляд:

$$y(k) = 1,6y(k-1) - 0,8y(k-2).$$

Характеристичне рівняння для нього:

$$\alpha^2 - 1,6\alpha + 0,8 = 0.$$

Дискримінант $d = a_1^2 + 4a_2 = -0,64$. Корені характеристичного рівняння і параметри однорідного розв'язку:

$$\alpha_1 = 0,5(1,6 + i0,8), \quad \alpha_2 = 0,5(1,6 - i0,8).$$

$$r = (-a_2)^{1/2} = (0,80)^{1/2} = 0,894.$$

$$\cos(\omega) = \frac{a_1}{2(-a_2)^{1/2}} = \frac{1,6}{2 \cdot 0,894} = 0,894.$$

Умови стійкості розв'язку однорідного рівняння другого порядку

В першому випадку (**A – різні дійсні корені**) для забезпечення стійкості розв'язку необхідно виконати умову:

$$-1 < \alpha_i < 1, \quad i = 1, 2 \quad \text{або} \quad -1 < \frac{1}{2} \left(a_1 \pm \sqrt{a_1^2 + 4a_2} \right) < 1.$$

Більший корінь характеристичного рівняння буде меншим одиниці, якщо

$$a_1 + (a_1^2 + 4a_2)^{1/2} < 2 \quad \text{або} \quad (a_1^2 + 4a_2)^{1/2} < 2 - a_1.$$

Таким чином, можна записати, що $a_1^2 + 4a_2 < 4 - 4a_1 + a_1^2$ або умовою стійкості є

$$a_1 + a_2 < 1.$$

Менший корінь $\alpha_2 = 0,5(a_1 - \sqrt{d})$ буде більшим -1 , якщо

$$a_1 - (a_1^2 + 4a_2)^{1/2} > -2 \quad \text{або} \quad 2 + a_1 > (a_1^2 + 4a_2)^{1/2}.$$

Звідси отримаємо: $4 + 4a_1 + a_1^2 > a_1^2 + 4a_2$, а умова стійкості має вигляд:

$$a_2 - a_1 < 1.$$

В другому випадку (**Б** – кратні корені) $a_1^2 + 4a_2 = 0$. Умовою стійкості розв’язку є $|a_1| < 2$.

В третьому випадку (**В** – комплексні корені) $a_1^2 + 4a_2 < 0$. Умова стійкості має вигляд: $r = (-a_2)^{1/2} < 1$ або $-a_2 < 1$, при $a_2 < 0$.

Загальною умовою стійкості розв’язку для всіх трьох випадків є така: корені характеристичного рівняння повинні лежати всередині одиничного кола на комплексній площині.

У випадку різних та кратних дійсних коренів вони будуть лежати на горизонтальній осі комплексної площини, але всередині кола одиничного радіуса. В третьому випадку корені розташовуються в довільних точках площини, але для забезпечення стійкості розв’язку вони також повинні лежати всередині одиничного кола.

Структура розв’язку однорідного рівняння вищого порядку

Розв'язки однорідних рівнянь вищих порядків знаходять по аналогії з рівнянням другого порядку. Розглянемо однорідне рівняння n -го порядку:

$$y(k) - \sum_{i=1}^n a_i y(k-i) = 0.$$

Підставимо в нього однорідний розв'язок загального виду $y^h(k) = A\alpha^k$:

$$A\alpha^k - a_1 A\alpha^{k-1} - a_2 A\alpha^{k-2} - \dots - a_n A\alpha^{k-n} = 0$$

і розділимо отримане рівняння на $A\alpha^{k-n}$. В результаті отримаємо характеристичне рівняння n -го порядку:

$$\alpha^n - a_1 \alpha^{n-1} - a_2 \alpha^{n-2} - \dots - a_n = 0.$$

Воно має n розв'язків, тобто n коренів α_i , $i = 1, \dots, n$. Згідно з отриманими вище результатами лінійна комбінація n коренів також буде розв'язком. Якщо серед знайдених коренів буде корінь кратності m , то розв'язок матиме вигляд:

$$y^h(k) = A_1 \alpha^k + A_2 k \alpha^k + A_3 k^2 \alpha^k + \dots + A_m k^{m-1} \alpha^k + A_{m+1} \alpha_{m+1}^k + \dots + A_n \alpha_n^k.$$

Очевидно, що у випадку наявності комплексно-спряженого кореня розв'язок буде мати відповідну гармонічну складову виду $y^h(k) = \beta_1 r^k \cos(\omega k + \varphi)$. Очевидно, що для рівнянь вищих порядків також необхідно аналізувати стійкість розв'язків, тому розглянемо спочатку стійкість (збіжність) розв'язку рівняння другого порядку.

Умови стійкості розв'язку однорідного рівняння другого порядку

На основі умов стійкості розв'язку однорідного рівняння другого порядку можна сформулювати умови визначення стійкості розв'язку однорідного рівняння n -го порядку:

1. Необхідною умовою того, щоб корені характеристичного рівняння лежали в одиничному колі ϵ

$$\sum_{i=1}^n a_i < 1.$$

2. Оскільки значення коефіцієнтів a_i можуть бути додатними та від'ємними, то достатньою умовою того, що корені характеристичного рівняння будуть лежати всередині одиничного кола є

$$\sum_{i=1}^n |a_i| < 1.$$

3. Щонайменше один корінь характеристичного рівняння дорівнює одиниці, якщо виконується наступна рівність:

$$\sum_{i=1}^n a_i = 1.$$

Будь-яку послідовність (часовий ряд), яка містить один або більше одиничних коренів відповідного характеристичного рівняння, називають *процесом з одиничним коренем* (або процес з трендом).

4. Умови стійкості розв'язку рівняння третього порядку мають вигляд:

$$\begin{aligned} 1 - a_1 - a_2 - a_3 &> 0, \\ 1 + a_1 - a_2 + a_3 &> 0, \\ 1 - a_1 a_3 + a_2 - a_3^2 &> 0, \\ 3 + a_1 + a_2 - 3a_3 &> 0 \quad \text{або} \quad 3 - a_1 + a_2 + 3a_3 > 0. \end{aligned}$$

При цьому, якщо перші три умови виконуються, то остання є надлишковою.

На практиці досить часто зустрічаються випадки, коли виникають труднощі з визначенням точних значень коренів характеристичного рівняння. Однак, в більшості випадків при дослідженні розв'язків достатньо знати якісну картину їх поведінки, для чого необхідно знати чи знаходяться корені характеристичного рівняння всередині одиничного кола. Тому для дослідження розв'язків можна застосувати *теорему Шура*, яка визначає необхідні та достатні умови стійкості.

Теорема Шура формулюється так: якщо для характеристичного рівняння n -го порядку всі n детермінантів (які наведені нижче) є додатними, то дійсні частини всіх характеристичних коренів будуть меншими одиниці по модулю, тобто:

$$|\operatorname{Re}(\alpha_i)| < 1, \quad i = 1, \dots, n.$$

$$\Delta_1 = \begin{vmatrix} 1 & -a_n \\ -a_n & 1 \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} 1 & 0 & -a_n & -a_{n-1} \\ -a_1 & 1 & 0 & -a_n \\ -a_n & 0 & 1 & -a_1 \\ -a_{n-1} & -a_n & 0 & 1 \end{vmatrix};$$

$$\Delta_3 = \begin{vmatrix} 1 & 0 & 0 & -a_n & -a_{n-1} & -a_{n-2} \\ -a_1 & 1 & 0 & 0 & -a_n & -a_{n-1} \\ -a_2 & -a_1 & 1 & 0 & 0 & -a_n \\ -a_n & 0 & 0 & 1 & -a_1 & -a_2 \\ -a_{n-1} & -a_n & 0 & 0 & 1 & -a_1 \\ -a_{n-2} & -a_{n-1} & -a_n & 0 & 0 & 1 \end{vmatrix};$$

$$\Delta_n = \begin{vmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 & -a_n & -a_{n-1} & \cdot & \cdot & \cdot & -a_1 \\ -a_1 & 1 & 0 & \cdot & \cdot & 0 & 0 & -a_n & \cdot & \cdot & \cdot & -a_2 \\ -a_2 & -a_1 & 1 & \cdot & \cdot & 0 & 0 & 0 & -a_n & \cdot & \cdot & -a_3 \\ \cdot & \cdot \\ -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdot & \cdot & 1 & 0 & 0 & 0 & \cdot & \cdot & -a_n \\ -a_n & 0 & 0 & \cdot & \cdot & 0 & 1 & -a_1 & -a_2 & \cdot & \cdot & -a_{n-1} \\ -a_{n-1} & -a_n & 0 & \cdot & \cdot & 0 & 0 & 1 & \cdot & \cdot & \cdot & -a_{n-2} \\ \cdot & \cdot \\ -a_2 & -a_3 & -a_4 & \cdot & \cdot & 0 & 0 & 0 & \cdot & \cdot & 1 & -a_1 \\ -a_1 & -a_2 & -a_3 & \cdot & \cdot & -a_n & 0 & 0 & \cdot & \cdot & \cdot & 1 \end{vmatrix}.$$

Зазначимо, що кожний з наведених детермінантів можна розглядати як такий, що складається з чотирьох блоків. Кожний блок Δ_i представляє собою трикутну матрицю вимірності $i \times i$. Лівий верхній блок має одиничну головну діагональ і всі нулі над нею. В цьому блоці індекс збільшується на одиницю по мірі руху вниз по стовпчику від діагоналі. Лівий нижній блок представляє собою транспонований верхній правий.

На головній діагоналі верхнього правого блоку знаходяться

коефіцієнти a_n , а під цією діагоналлю – нулі. Якщо рухатись вверх по стовпчиках від діагоналі, то індекс зменшується на одиницю. Нижній лівий блок представляє собою транспозицію верхнього правого. Значення коефіцієнта a_0 прийнято за одиницю. Все ці особливості полегшують побудову наведених вище детермінантів.

Якщо прийняти до уваги те, що на практиці досить рідко використовуються рівняння, порядок яких складає більше 3, то обчислення детермінантів реалізується досить просто. Крім того, можна легко побудувати універсальну програму для обчислення детермінантів довільного порядку. Разом з тим необхідно зауважити, що надзвичайно простими в користуванні є наведені вище чотири умови визначення стійкості розв'язку однорідного рівняння n -го порядку, які дозволяють оперативно встановити якісний характер розв'язку.

3.3.2 Знаходження часткових розв'язків

Наявність часткового (неоднорідного) розв'язку зумовлена наявністю у рівнянні констант, а також детермінованих і випадкових змінних і функцій. Вигляд часткового розв'язку залежить від типу збурюючої функції в правій частині різницевого рівняння. Методика знаходження часткових розв'язків складається з таких кроків:

- формування пробної функції;
- підстановка пробної функції у рівняння і визначення невідомих коефіцієнтів пробної функції (наприклад, за методом невизначених коефіцієнтів);
- аналіз характеру та збіжності отриманого розв'язку.

Розглянемо декілька характерних випадків.

Випадок 1

У правій частині різницевого рівняння РР наявні тільки постійна складова (зміщення) a_0 і авторегресійна (AP) частина. Тобто рівняння має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n). \quad (3.3.7)$$

При подачі константи на вхід лінійного процесу на виході також повинна бути константа. Тобто можна передбачити, що частковим розв'язком цього рівняння повинна бути деяка константа, що можна записати так: $y(k) = y(k-1) = \dots = c$. Підставимо цей пробний розв'язок в (3.3.7) і

отримаємо:

$$c = a_0 + a_1 c + a_2 c + \dots + a_n c.$$

Звідси

$$c = \frac{a_0}{1 - a_1 - a_2 - \dots - a_n}. \quad (3.3.8)$$

Оскільки знаменник не повинен дорівнювати нулю, то константа c є частковим розв'язком і

$$y^h = \frac{a_0}{1 - a_1 - a_2 - \dots - a_n} = \frac{a_0}{1 - \sum_{i=1}^n a_i}. \quad (3.3.9)$$

Якщо $\sum_{i=0}^n a_i = 1$, то маємо особливий випадок, коли процес $\{y(k)\}$

називають *процесом з одиничним коренем*. Оскільки в цьому випадку процес $\{y(k)\}$ розбіжний, то частковий пробний розв'язок у вигляді константи не підходить для використання. Очевидно, що цю розбіжність (*лінійний* або *квазілінійний тренд*) необхідно описати функцією часу. У простому випадку (наявний один одиничний корінь характеристичного рівняння) можна скористатися пробним розв'язком $y^p(k) = ck$. Підставимо цей пробний розв'язок в (3.3.7) і отримаємо:

$$ck = a_0 + (k-1)a_1 c + (k-2)a_2 c + \dots + (k-n)a_n c$$

або

$$(1 - a_1 - a_2 - \dots - a_n)ck = a_0 - c(a_1 + 2a_2 + 3a_3 + \dots + na_n).$$

Оскільки $1 - a_1 - a_2 - \dots - a_n = 0$, то

$$c = \frac{a_0}{a_1 + 2a_2 + 3a_3 + \dots + na_n}.$$

Наприклад, для рівняння

$$y(k) = 2,5 + 0,75 y(k-1) + 0,25 y(k-2)$$

$a_1 + a_2 = 0,75 + 0,25 = 1$ і частковий розв'язок має вигляд ck , де $c = 2,5 / (0,75 + 2 \cdot 0,25) = 2,0$. Якщо пробний розв'язок ck не приводить до успішного знаходження розв'язку, то необхідно взяти іншу пробну функцію.

Наприклад, можна скористатися нелінійними функціями виду $y^p(k) = ck^2, ck^3, \dots, ck^n$. Для рівняння довільного порядку одна з цих функцій обов'язково буде частковим розв'язком.

Випадок 2

Процес збурюється експонентою $x(k) = bd^{rk}$, де b, d, r – константи. Наприклад, для економічних процесів коефіцієнт r можна розглядати як фактор зростання. Розглянемо послідовність знаходження часткового розв'язку на прикладі рівняння першого порядку:

$$y(k) = a_0 + a_1 y(k-1) + bd^{rk}. \quad (3.3.10)$$

У правій частині цього рівняння є константа a_0 . Тому можна очікувати, що в частковому розв'язку також буде наявна константа. Для лінійного процесу можна також передбачити, що при подачі на його вхід сигналу bd^{rk} на виході з'явиться реакція такої ж форми, але з деяким коефіцієнтом передачі. Тому логічно записати пробний розв'язок у формі:

$$y^p(k) = c_0 + c_1 d^{rk},$$

де c_0, c_1 – невідомі константи, які знайдемо за методом невизначених коефіцієнтів. Підставимо пробний розв'язок в рівняння (3.3.10) і в результаті отримаємо:

$$c_0 + c_1 d^{rk} = a_0 + a_1(c_0 + c_1 d^{r(k-1)}) + bd^{rk}. \quad (3.3.11)$$

Перенесемо в ліву частину члени, які мають невідомі константи c_0, c_1 і об'єднаємо їх наступним чином:

$$c_0 + c_1 \frac{d^{rk}}{1-a_1} \left(\frac{d^r - a_1}{d^r} \right) = \frac{a_0}{1-a_1} + \frac{b}{1-a_1} d^{rk}.$$

Звідси знайдемо, що

$$c_0 = \frac{a_0}{1-a_1}; \quad c_1 \frac{d^{rk}}{1-a_1} \frac{d^r - a_1}{d^r} = \frac{b}{1-a_1} d^{rk},$$

або

$$c_1 = \frac{bd^r}{d^r - a_1}.$$

Тепер запишемо частковий розв'язок для рівняння (3.3.10):

$$y^p(k) = \frac{a_0}{1-a_1} + \frac{bd^r}{d^r - a_1} d^{rk}. \quad (3.3.12)$$

Якщо $|d^r| < 1$, то (3.3.12) збігається до значення $a_0/(1-a_1)$ при $k \rightarrow \infty$. Якщо $a_1 = 1$, то при знаходженні часткового розв'язку необхідно взяти пробну функцію $c_0 = ck$. А якщо $a_1 = d^r$, то необхідно завдати пробну функцію у вигляді:

$$c_1 = \frac{kbd^r}{d^r - a_1}.$$

Методика знаходження часткового розв'язку для рівнянь вищих порядків залишається такою самою. Спочатку необхідно коректно вибрати пробний розв'язок у відповідністю із структурою рівняння, а потім застосувати метод невизначених коефіцієнтів для знаходження невідомих параметрів.

Випадок 3

На вхід процесу подається *функція* у формі $x(k) = bk^d$, де b – константа; d – ціле додатне число. Різницеве рівняння має вигляд:

$$y(k) = a_0 + \sum_{i=1}^n a_i y(k-i) + bk^d. \quad (3.3.13)$$

В даному випадку $y(k)$ залежить від k^d , значення $y(k-1)$ залежить від $(k-1)^d$, значення $y(k-2)$ залежить від $(k-2)^d$ і так далі. Тому пробну функцію для часткового розв'язку можна вибрати у формі:

$$y^p(k) = c_0 + c_1 k + c_2 k^2 + \dots + c_d k^d.$$

Для того щоб знайти невідомі коефіцієнти c_i , підставимо пробний

розв'язок в рівняння другого порядку:

$$c_0 + c_1 k = a_0 + a_1 [c_0 + (k-1)c_1] + a_2 [c_0 + (k-2)c_1] + bk. \quad (3.3.14)$$

Значення коефіцієнтів c_0, c_1 виберемо такими, щоб (3.3.14) залишалось тотожністю для всіх можливих значень k . Згрупуємо всі константи і всі члени, які мають множник k :

$$c_0 - a_0 - a_1 c_0 + a_1 c_1 - a_2 c_0 + 2a_2 c_1 = a_1 k c_1 + a_2 k c_1 + b k - c_1 k.$$

Звідси отримаємо значення невідомих констант за *методом невизначених коефіцієнтів*:

$$\begin{aligned} c_1 &= b / (1 - a_1 - a_2), \\ c_0 &= [a_0 - (2a_2 + a_1)c_1] / (1 - a_1 - a_2) \end{aligned}$$

або

$$c_0 = \frac{a_0}{1 - a_1 - a_2} - \frac{b(2a_2 + a_1)}{(1 - a_1 - a_2)^2}.$$

Таким чином, частковий розв'язок включає в себе лінійний часовий тренд. Якщо $a_1 + a_2 = 1$, то пробний розв'язок необхідно домножити на k . Методика залишається незмінною для рівнянь вищого порядку.

Приклад 3.4. Розглянемо детермінований процес авторегресії другого порядку:

$$\begin{aligned} y(k) &= 2,5 + a_1(k-1) + a_2 y(k-2), \\ a_1 &= 0,75; \quad a_2 = 0,25; \quad y(0) = 1; \quad y(1) = 2. \end{aligned}$$

Оскільки сума коефіцієнтів моделі $\sum_{i=1}^2 a_i = 0,75 + 0,25 = 1$, то за пробну функцію часткового розв'язку вибираємо $y_{\text{проб}}^p(k) = ck$, де

$$c = \frac{a_0}{\sum_{i=1}^2 a_i} = \frac{2,5}{0,75 + 2 \cdot 0,25} = \frac{2,5}{1,25} = 2,0.$$

Таким чином, частковий розв'язок має вигляд: $y^p(k) = 2,0k$.

Характеристичне рівняння $\alpha^2 - 0,75\alpha - 0,25 = 0$ має корені: $\alpha_1 = 1$, $\alpha_2 = -0,25$. Запишемо повний розв'язок

$$y(k) = A_1 + A_2(-0,25)^k + 2,0k.$$

Використаємо початкові умови для знаходження довільних констант:

$$k=0: 1 = A_1 + A_2,$$

$$k=1: 2 = A_1 - 0,25 A_2 + 2,0.$$

Розв'зуючи цю систему двох рівнянь, знайдемо: $A_1 = 0,2$; $A_2 = 0,8$.

Приклад 3.5. Необхідно повний розв'язок рівняння:

$$y(k) = a_0 + 2y(k-1) - y(k-2).$$

Характеристичне рівняння $\alpha^2 - 2\alpha + 1 = 0$ має корені: $\alpha_1 = \alpha_2 = 1$, тобто, маємо випадок кратних коренів. Оскільки пробна функція $y_{npo\delta}^p(k) = ck$ не приводить до успіху, то вибираємо за пробну $y_{npo\delta}^p(k) = ck^2$. В результаті підстановки пробного розв'язку в вихідне рівняння знайдемо:

$$c = \frac{a_0}{k^2 - 2(k-1)^2 + (k-2)^2} = \frac{a_0}{2}.$$

Тепер запишемо повний розв'язок:

$$y(k) = A_1 + A_2 k + \frac{a_0}{2} k^2.$$

Випадок 4

Застосування методу невизначених коефіцієнтів для знаходження часткового розв'язку стохастичного різницевого рівняння (у правій частині наявний випадковий процес).

Процес авторегресії першого порядку

Стохастична авторегресія першого порядку описується рівнянням

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k).$$

Природа процесу $\{y(k)\}$ в даному випадку така, що частковий розв'язок залежить в даному випадку тільки від постійного члена, часу та послідовності $\{\varepsilon(k)\}$. Враховуючи те, що часова змінна у збурюючій функції задана неявно, то змінна k може входити в частковий розв'язок тільки в тому випадку, коли корінь характеристичного рівняння дорівнює 1.

Для ілюстрації методу скористаємося такою пробною функцією:

$$y(k) = b_0 + b_1 k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i), \quad (3.3.15)$$

де b_0, b_1, α_i – невідомі коефіцієнти. Для того щоб знайти значення цих

коєфіцієнтів, підставимо пробний розв'язок (3.3.15) в РР першого порядку:

$$\begin{aligned} b_0 + b_1 k + \alpha_0 \varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots &= \\ = a_0 + a_1 [b_0 + b_1(k-1) + \alpha_0 \varepsilon(k-1) + \alpha_1 \varepsilon(k-2) + \dots] + \varepsilon(k). \end{aligned}$$

Після групування подібних членів отримаємо:

$$(b_0 - a_0 - a_1 b_0 + a_1 b_1) + b_1(1 - a_1)k + (\alpha_0 - 1)\varepsilon(k) + (\alpha_1 - a_1 \alpha_0)\varepsilon(k-1) + (\alpha_2 - a_1 \alpha_1)\varepsilon(k-2) + (\alpha_3 - a_1 \alpha_2)\varepsilon(k-3) + \dots = 0. \quad (3.3.16)$$

Отримане рівняння (3.3.16) повинне бути справедливим при всіх можливих значеннях k і всіх можливих значеннях $\varepsilon(k)$. Звідси випливає, що повинні виконуватись такі умови:

$$\begin{aligned} \alpha_0 - 1 &= 0, \\ \alpha_1 - a_1 \alpha_0 &= 0, \\ \alpha_2 - a_1 \alpha_1 &= 0, \\ &\vdots \\ b_0 - a_0 - a_1 b_0 + a_1 b_1 &= 0, \\ b_1 - a_1 b_1 &= 0. \end{aligned} \quad (3.3.17)$$

Зазначимо, що першою групою умов (перші три рядки) можна скористатись для знаходження розв'язку відносно α_i в рекурсивній формі. Використовуючи умови (3.3.17), знайдемо:

$$\alpha_0 = 1, \quad \alpha_1 = a_1, \quad \alpha_2 = a_1 \alpha_1 = a_1^2, \quad \alpha_i = a_1^i.$$

Якщо $a_1 \neq 1$, то $b_1 = 0$, $b_0 = a_0 / (1 - a_1)$ і частковий розв'язок має вигляд:

$$y(k) = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i).$$

Такий же розв'язок був отриманий вище для РР першого порядку методом ітерацій. Загальний розв'язок отримаємо в результаті об'єднання часткового та однорідного розв'язків:

$$y(k) = \frac{a_0}{1 - a_1} + A a_1^k + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i). \quad (3.3.18)$$

Якщо відома початкова умова $y(0) = y_0$, то можна знайти значення довільної константи із рівняння:

$$y_0 = \frac{a_0}{1 - a_1} + A + \sum_{i=0}^{\infty} a_1^i \varepsilon(-i).$$

Таким чином загальний розв'язок можна записати так:

$$y(k) = \frac{a_0}{1 - a_1} + a_1^k \left[y_0 - \frac{a_0}{1 - a_1} - \sum_{i=0}^{\infty} a_1^i \varepsilon(-i) \right] + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i),$$

або

$$y(k) = \frac{a_0}{1 - a_1} + a_1^k \left(y_0 - \frac{a_0}{1 - a_1} \right) + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i). \quad (3.3.19)$$

Якщо $a_1 = 1$, то b_0 може бути довільною константою, а $b_1 = a_0$. В цьому випадку можливою формою часткового розв'язку може бути функція, яка включає добуток $a_0 k$, оскільки з'являється одиничний корінь. Така функція має вигляд:

$$y(k) = b_0 + a_0 k + \sum_{i=0}^{\infty} \varepsilon(k-i). \quad (3.3.20)$$

Таку функцію називають «неналежною», тому що сума елементів послідовності $\{\varepsilon(k)\}$ може бути нескінченною. Для початкової умови y_0 запишемо

$$y_0 = b_0 + \sum_{i=0}^{\infty} \varepsilon(-i) \quad \text{або} \quad b_0 = y_0 - \sum_{i=0}^{\infty} \varepsilon(-i).$$

Із врахуванням цього результату (3.3.20) приймає вигляд:

$$y(k) = y_0 + a_0 k + \sum_{i=0}^k \varepsilon(k-i).$$

Стохастичний процес ARKC(1,1)

Як розширення попереднього випадку, розглянемо тепер стохастичне рівняння ARKC(1,1):

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k) + \beta_1 \varepsilon(k-1). \quad (3.3.21)$$

Розв'язок цього рівняння буде включати константу, зважену послідовність $\{\varepsilon(k)\}$ та час k в першій степені (тому що рівняння першого порядку). Як і у попередньому прикладі, k потрібно включати в пробний розв'язок, якщо корінь характеристичного рівняння дорівнює одиниці. Таким чином, скористаємося пробним розв'язком виду:

$$y_{np}(k) = b_0 + b_1 k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i).$$

Підставимо пробний розв'язок в (3.3.21) і отримаємо:

$$b_0 + b_1 k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i) = a_0 + a_1 \left[b_0 + b_1(k-1) + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i-1) \right] + \varepsilon(k) + \beta_1 \varepsilon(k-1)$$

або

$$\begin{aligned} b_0 + b_1 k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i) &= a_0 + a_1 b_0 + a_1 b_1 k - a_1 b_1 + a_1 \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i-1) + \\ &\quad + \varepsilon(k) + \beta_1 \varepsilon(k-1). \end{aligned}$$

Прирівнюючи коефіцієнти при однакових змінних $\varepsilon(k), \varepsilon(k-1), \varepsilon(k-2), \dots$, отримаємо наступні значення невідомих коефіцієнтів розв'язку:

$$\alpha_0 = 1;$$

$$\alpha_1 = a_1 \alpha_0 + \beta_1 = a_1 + \beta_1;$$

$$\alpha_2 = a_1 \alpha_1 = a_1(a_1 + \beta_1);$$

$$\alpha_3 = a_1 \alpha_2 = a_1^2(a_1 + \beta_1);$$

⋮

$$\alpha_i = a_1 \alpha_{i-1} \quad \text{або} \quad \alpha_i = a_1^{i-1}(a_1 + \beta_1).$$

Прирівнюючи константи та члени, які містять час k , отримаємо:

$$b_0 = a_0 + a_1 b_0 - a_1 b_1;$$

$$b_1 = a_1 b_1.$$

Для даного прикладу також необхідно розглянути два випадки: $a_1 \neq 1$ і $a_1 = 1$. Якщо $a_1 \neq 1$, то це означає, що тренду немає і $b_1 = 0$. Звідси випливає, що

$$b_0 = \frac{a_0}{1-a_1},$$

а частковий розв'язок має вигляд:

$$y(k) = \frac{a_0}{1-a_1} + \varepsilon(k) + (a_1 + \beta_1) \sum_{i=1}^{\infty} a_1^{i-1} \varepsilon(k-i).$$

Загальний розв'язок включає ще однорідний:

$$y(k) = \frac{a_0}{1-a_1} + A a_1^k + \varepsilon(k) + (a_1 + \beta_1) \sum_{i=1}^{\infty} a_1^{i-1} \varepsilon(k-i).$$

Якщо $a_1 = 1$, то з виразу $b_0 = a_0 + a_1 b_0 - a_1 b_1$ випливає, що $b_1 = a_0$, а коефіцієнт b_0 є довільною константою, яка може бути визначена за допомогою початкових умов. Таким чином, «неналежна» форма розв'язку для даного випадку має вигляд:

$$y(k) = b_0 + a_0 k + \varepsilon(k) + (1 + \beta_1) \sum_{i=1}^{\infty} \varepsilon(k-i). \quad (3.3.22)$$

Покладаючи початкову умову $y(0) = y_0$ при $k = 0$, отримаємо:

$$y_0 = b_0 + \varepsilon_0 + (1 + \beta_1) \sum_{i=1}^{\infty} \varepsilon(-i),$$

що дає змогу визначити b_0 за виразом: $b_0 = y_0 - \varepsilon(0) - (1 + \beta_1) \sum_{i=1}^{\infty} \varepsilon(-i)$.

Підставимо цей вираз в (3.3.22):

$$\begin{aligned} y(k) &= y_0 - \varepsilon(0) - (1 + \beta_1) \sum_{i=1}^{\infty} \varepsilon(-i) + a_0 k + \varepsilon(k) + (1 + \beta_1) \sum_{i=1}^{\infty} \varepsilon(k-i) = \\ &= y_0 + a_0 k - \varepsilon(0) - (1 + \beta_1) \varepsilon(-1) - (1 + \beta_1) \varepsilon(-2) - (1 + \beta_1) \varepsilon(-3) - \dots + \varepsilon(k) + \\ &\quad + (1 + \beta_1) \varepsilon(k-1) + (1 + \beta_1) \varepsilon(k-2) + \dots + (1 + \beta_1) \varepsilon(0) + (1 + \beta_1) \varepsilon(-1) + \dots = \\ &= y_0 + a_0 k + \varepsilon(k) + (1 + \beta_1) \sum_{i=1}^{k-1} \varepsilon(k-i). \end{aligned}$$

Стохастичний процес AP(2)

Оскільки на практиці досить часто зустрічаються моделі другого порядку, то розглянемо знаходження розв'язку для стохастичної авторегресії AP(2). Це рівняння має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + \varepsilon(k). \quad (3.3.23)$$

Оскільки це рівняння другого порядку, то його характеристичне рівняння може мати два одиничних корені, а тому максимальний порядок полінома для описання можливого тренду повинен складати 2. Тобто пробний розв'язок запишемо так:

$$y_{np}(k) = b_0 + b_1 k + b_2 k^2 + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i),$$

де b_0, b_1, b_2, α_i – невідомі коефіцієнти розв'язку рівняння (3.3.23). Підстановка пробного розв'язку в рівняння (3.3.23) приводить до такого виразу:

$$\begin{aligned} & b_0 + b_1 k + b_2 k^2 + \alpha_0 \varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \dots = \\ & = a_0 + a_1 [b_0 + b_1(k-1) + b_2(k-1)^2 + \alpha_0 \varepsilon(k-1) + \alpha_1 \varepsilon(k-2) + \alpha_2 \varepsilon(k-3) + \dots] + \\ & + a_2 [b_0 + b_1(k-2) + b_2(k-2)^2 + \alpha_0 \varepsilon(k-2) + \alpha_1 \varepsilon(k-3) + \alpha_2 \varepsilon(k-4) + \dots] + \varepsilon(k). \end{aligned} \quad (3.3.24)$$

Із умов, які приводять наведений вираз до тотожності при довільних реалізаціях процесу $\{\varepsilon(k)\}$, отримуємо такі значення невизначених коефіцієнтів:

$$\begin{aligned} \alpha_0 &= 1; \\ \alpha_1 &= a_1 \alpha_0 = a_1; \\ \alpha_2 &= a_1 \alpha_1 + a_2 \alpha_0 = a_1^2 + a_2; \\ \alpha_3 &= a_1 \alpha_2 + a_2 \alpha_1 = a_1^3 + 2a_1 a_2; \\ &\vdots \end{aligned}$$

З наведених виразів для обчислення коефіцієнтів α_i видно, що

починаючи з α_2 , всі коефіцієнти можна визначити за допомогою різницевого рівняння другого порядку:

$$\alpha_j = a_1\alpha_{j-1} + a_2\alpha_{j-2}.$$

Оскільки α_0 і α_1 відомі, то всі наступні коефіцієнти можна визначити ітераційно.

Властивості збіжності невідомих коефіцієнтів можна визначити так само, як і властивості збіжності розв'язку однорідного рівняння другого порядку, тобто:

1. Для збіжності розв'язку однорідного рівняння другого порядку необхідно виконати наступні умови: $|a_2| < 1$, $a_1 + a_2 < 1$ і $a_2 - a_1 < 1$. Збіжність також передбачає, що вплив минулих значень послідовності $\{\varepsilon(k)\}$ на поточне значення $y(k)$ зменшується із плином часу.
2. Якщо коефіцієнти збігаються до деяких скінчених усталених значень, то ця збіжність має безпосередній прямий характер при $(a_1^2 + 4a_2) > 0$. Якщо $(a_1^2 + 4a_2) < 0$, то збіжність буде мати гармонічну форму, а при $(a_1^2 + 4a_2) = 0$ спочатку спостерігається швидке збільшення цих значень з подальшою збіжністю.

Розглянемо члени тотожності (3.3.24), які не зв'язані з випадковою послідовністю $\{\varepsilon(k)\}$:

$$\begin{aligned} & b_2(1-a_1-a_2)k^2 + [b_1(1-a_1-a_2) + 2b_2(a_1+2a_2)]k + \\ & + [b_0(1-a_1-a_2) - a_0 + a_1(b_1-b_2) + 2a_2(b_1-2b_2)] = 0. \end{aligned} \quad (3.3.25)$$

Рівність лівої частини цього виразу нулю повинна виконуватись для $\forall k$.

Спочатку розглянемо випадок, коли $a_1 + a_2 \neq 1$. Оскільки в такому випадку $1 - a_1 - a_2 \neq 0$, то для збіжності необхідно покласти $b_2 = 0$. Оскільки для забезпечення збіжності коефіцієнт при k повинен дорівнювати нулю, то необхідно також покласти $b_1 = 0$. Якщо $b_1 = b_2 = 0$, то можна знайти, що

$$b_0 = \frac{a_0}{1 - a_1 - a_2}.$$

Якщо ж $a_1 + a_2 = 1$, то значення коефіцієнтів b_0, b_1, b_2 будуть залежати

від конкретних значень коефіцієнтів a_0, a_1, a_2 .

Таким чином, можна зробити наступний висновок:

Умовами збіжності часткового розв'язку є умови стійкості розв'язку однорідного рівняння. Якщо корені характеристичного рівняння є одиничними, то в частковому розв'язку з'являється поліноміальний тренд, а порядок полінома дорівнює числу одиничних коренів.

Примітка: Якщо складові збурюючої функції в правій частині рівняння, тобто $x(k)$, мають такий же вигляд як і складові $y^h(k)$, то пробна функція часткового розв'язку змінюється. Всі складові $y^p(k)$, які відповідають таким же складовим $y^h(k)$, необхідно помножити на k в такій найменшій степені, щоб їх тотожність порушувалась. Відповідний приклад буде розглянуто в наступному параграфі.

Процес AP(2) з випадковою та детермінованою функціями на вході

Розглянемо процес з випадковою та детермінованою функціями на вході:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + x(k) + \varepsilon(k), \quad (3.3.26)$$

де $a_0 = 0,3$; $a_1 = -0,5$; $a_2 = 1,5$; $x(k) = c_1 e^{c_2(k+1)}$; $\varepsilon(k)$ – центрований випадковий процес типу білого шуму; c_1, c_2 – константи.

Характеристичне рівняння $\alpha^2 + 0,5\alpha - 1,5 = 0$ має корені: $\alpha_1 = 1$ і $\alpha_2 = -1,5$. Тобто, це процес з одиничним коренем, який має тренд. Таким чином, однорідний розв'язок має вигляд:

$$y^h(k) = A_1 + A_2 (-1,5)^k.$$

Підставимо пробну функцію

$$y_{np}^p(k) = b_0 + b_1 k + b_2 k^2 + b_3 e^{c_2(k+1)} + \sum_{i=1}^{\infty} \alpha_i \varepsilon(k-i)$$

в рівняння (3.3.26):

$$\begin{aligned}
b_0 + b_1 k + b_2 k^2 + b_3 e^{c_2(k+1)} + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i) = \\
= a_0 + a_1 [b_0 + b_1(k-1) + b_2(k-1)^2 + b_3 e^{c_2 k} + \alpha_0 \varepsilon(k-1) + \alpha_1 \varepsilon(k-2) + \dots] + \\
+ a_2 [b_0 + b_1(k-2) + b_2(k-2)^2 + b_3 e^{c_2(k-1)} + \alpha_0 \varepsilon(k-2) + \alpha_2 \varepsilon(k-3) + \dots] + \\
+ c_1 e^{c_2(k+1)}.
\end{aligned}$$

Прирівняємо коефіцієнти в лівій і правій частині при однакових змінних:

$$b_0 = a_0 + a_1 b_0 - a_1 b_1 + a_1 b_2 + a_2 b_0 - a_2 b_1 + 4 a_2 b_2; \quad (3.3.27)$$

Коефіцієнти при k :

$$b_1 = a_1 b_1 - 2 a_1 b_2 + a_2 b_1 - 4 a_2 b_2 \quad \text{або} \quad b_1(1 - a_1 - a_2) = -2 b_2(a_1 + 2 a_2).$$

Оскільки $1 - a_1 - a_2 = 1 + 0,5 - 1,5 = 0$, а $a_1 + 2 a_2 = -0,5 + 5 = 2,5$, то необхідно покласти $b_2 = 0$.

При $b_2 = 0$ із (3.3.27) можна записати, що

$$b_0 = a_0 + a_1 b_0 - a_1 b_1 + a_2 b_0 - a_2 b_1 \quad \text{або} \quad b_0(1 - a_1 - a_2) = a_0 - b_1(a_1 + a_2).$$

$$\text{Звідси } b_1 = \frac{a_0}{a_1 + a_2} = \frac{0,3}{-0,5 + 1,5} = 0,3.$$

Прирівняємо ще коефіцієнти при $e^{c_2 k}$:

$$b_3 e^{c_2} = a_1 b_3 + a_2 b_3 e^{-c_2} + c_1 e^{c_2},$$

і визначимо константу b_3 :

$$b_3 = \frac{c_1 e^{c_2}}{e^{c_2} - a_1 - a_2 e^{-c_2}} \quad \text{або} \quad b_3 = \frac{c_1}{1 - a_1 e^{-c_2} - a_2 e^{-2c_2}} = \frac{c_1}{1 + 0,5 e^{-c_2} - 1,5 e^{-2c_2}}.$$

Залишився невизначенним коефіцієнт b_0 , який можна знайти за допомогою початкових умов.

Запишемо повний розв'язок:

$$y(k) = b_0 + A_1 + A_2(-1,5)^k + 0,3k + \frac{c_1}{1 + 0,5e^{-c_2} - 1,5e^{-2c_2}} e^{c_2(k+1)} + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i),$$

де $\alpha_i = -0,5\alpha_{i-1} + 1,5\alpha_{i-2}$. Очевидно, що цей процес має тренд, на який накладаються коливання у вигляді адитивних складових $A_2(-1,5)^k$ та $\sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i)$.

3.4 Приклади знаходження повних розв'язків різницевих рівнянь

Приклад 3.4.1. Модель процесу інфляції.

В роботі [10] для моделювання процесу інфляції запропоновано різницеве рівняння наступного вигляду:

$$m(k) = \alpha + p(k) - \beta[p^e(k+1) - p(k)], \quad \beta > 0,$$

де $m(k)$ – логарифм номінального забезпечення грошовою масою в момент k ; $p(k)$ – логарифм поточного рівня цін; $p^e(k+1)$ – логарифм рівня цін, які очікуються в момент часу $k+1$. Очевидно, що наведене рівняння може бути використане тільки для нормованих безрозмірних змінних.

Перепишемо наведену вище модель у вигляді:

$$m(k) - p(k) = \alpha - \beta[p^e(k+1) - p(k)].$$

Тепер можна стверджувати, що потреба в грошах $[m(k) - p(k)]$ знаходиться у від'ємному взаємозв'язку із приростом цін $[p^e(k) - p(k)]$, тобто відносне перевищення цін над об'ємом грошової маси в лівій частині рівняння вимагає знаку мінус перед приростом цін в правій частині.

Звичайно, що наведена вище модель є дуже спрощеною для перехідного періоду, тому що для її побудови використано тільки дві змінні. Фактично, на інфляцію в перехідний період впливають і інші фактори, такі як значне падіння рівня виробництва, перекачування валюти за кордон (в той час як вона повинна використовуватися для збільшення внутрішніх інвестицій), рівень тіньової економіки і т.п. Така модель буде набагато складнішою за своєю структурою, але більш адекватною процесу інфляції.

Для визначення впливу на процес інфляції майбутніх збурень отримаємо розв'язок рівняння відносно їх майбутніх значень. Нехай забезпечення грошовою масою має наступний простий вигляд:

$$m(k) = m_a + \varepsilon(k),$$

де m_a – середнє значення випуску грошової маси; $\varepsilon(k)$ – некорельзований випадковий процес з нульовим середнім, $E[\varepsilon(k)] = 0$. Покладемо також, що

$$p^e(k+1) = p(k+1),$$

тобто прогнозовані ціни відповідають дійсним значенням цін в момент $k+1$.

Таким чином, рівняння рівноваги грошової маси та рівня цін приймає вигляд:

$$m_a + \varepsilon(k) - p(k) = \alpha - \beta[p(k+1) - p(k)]$$

або

$$p(k+1) - (1 + \frac{1}{\beta})p(k) = -(m_a - \alpha)/\beta - \varepsilon(k)/\beta.$$

Знайдемо розв'язок цього рівняння. Для знаходження часткового розв'язку скористаємося методом невизначених коефіцієнтів. Оскільки перед змінною $p(k)$ стоїть коефіцієнт $a_1 = 1 + 1/\beta > 1$, то будемо шукати розв'язок відносно майбутніх відліків збурень. Пробний розв'язок записано у вигляді:

$$p_{npo\delta}(k) = b_0 + \sum_{i=0}^{\infty} \gamma_i \varepsilon(k+i)$$

і підставимо його у наведене вище рівняння для того щоб знайти невідомі коефіцієнти b_0, γ_i :

$$b_0 + \sum_{i=0}^{\infty} \gamma_i \varepsilon(k+1+i) - (1 + 1/\beta) \left[b_0 + \sum_{i=0}^{\infty} \gamma_i \varepsilon(k+i) \right] = [\alpha - m_a - \varepsilon(k)]/\beta.$$

Для того щоб останнє рівняння було тотожністю для всіх можливих ненульових значень $\{\varepsilon(k)\}$ необхідно покласти (прирівняти коефіцієнти при одинакових значеннях $\varepsilon(k)$):

- константи: $b_0 - b_0(1 + \beta)/\beta = (\alpha - m_a)/\beta \Rightarrow b_0 = m_a - \alpha$,
- коефіцієнти при $\varepsilon(k)$: $-\gamma_0(1 + \beta)/\beta = -1/\beta \Rightarrow \gamma_0 = 1/(1 + \beta)$,
- коефіцієнти при $\varepsilon(k+1)$: $\gamma_0 - \gamma_1(1 + \beta)/\beta = 0 \Rightarrow \gamma_1 = \beta/(1 + \beta)^2$,
- для довільного моменту часу: $\gamma(i) - \gamma(i-1)(1 + \beta)/\beta = \beta^i/(1 + \beta)^{i+1}$.

Частковий розв'язок в компактній формі матиме вигляд:

$$p^P(k) = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(k+i).$$

Для рівняння першого порядку

$$p(k+1) - \left(1 + \frac{1}{\beta}\right)p(k) = 0$$

однорідний розв'язок запишемо так:

$$p^h(k) = A \left(1 + \frac{1}{\beta}\right)^k.$$

Таким чином, загальний розв'язок, як сума часткового та однорідного, приймає наступний вигляд:

$$p(k) = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(k+i) + A \left(1 + \frac{1}{\beta}\right)^k.$$

Очевидно, що $\lim_{i \rightarrow \infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} = 0$, тобто частковий розв'язок не розбігається,

але однорідний буде зростати при збільшенні k . Для того щоб уникнути розбіжності однорідного розв'язку, необхідно покласти $A=0$.

Скористаємося початковими умовами для визначення умов збіжності однорідного розв'язку:

$$p_0 = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(i) + A,$$

або

$$A = p_0 - m_a + \alpha - \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(i).$$

Таким чином, для того щоб $A=0$ необхідно мати наступну початкову умову:

$$p_0 = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(i).$$

В цьому рівнянні $(m_a - \alpha)$ – детермінована умова довгострокової “рівноваги”, яка зустрічається також в інших моделях. Якщо розв’язок початкового рівняння збігається, то він повинен збігатися до цієї детермінованої частини. Складова розв’язку

$$\frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(k+1)$$

ілюструє вплив збурень на величину грошової маси в обороті. Якщо прийняти до уваги те, що

$$\left| \frac{\beta}{1+\beta} \right| < 1,$$

то вплив цієї складової на поведінку розв’язку в цілому буде помітним тільки на короткому початковому проміжку часу, оскільки коефіцієнти зменшуватимуться зростом значення степені $i + 1$.

Із сказаного можна зробити висновок, що частковий розв’язок характеризує поведінку (рівновагу) головної змінної $p(k)$ на короткому та більш значному інтервалах часу. Однорідний розв’язок можна розглядати як порушення рівноваги в початковий період часу.

Оскільки частковий розв’язок

$$p^P(k) = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(k+i)$$

представляє собою рівняння рівноваги між ціною та грошовою масою для k -го періоду часу, то початкове значення

$$p_0 = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(i)$$

є рівноважним значенням ціни для нульового періоду. Таким чином, умова

$$A(1 + 1/\beta)^k = 0$$

є умовою нульового відхилення від рівноваги (ціни \leftrightarrow гроші) в початковий період часу.

Якщо на послідовність $\{p(k)\}$ накласти обмеження, що вона не повинна розбігатися, то загальний розв'язок (його називають в такому вигляді ще імпульсною характеристикою – частковий розв'язок без константи) матиме наступний вигляд:

$$p(k) = m_a - \alpha + \frac{1}{\beta} \sum_{i=0}^{\infty} \left(\frac{\beta}{1+\beta} \right)^{i+1} \varepsilon(k+i).$$

Відзначимо, що в кожний момент часу k ціна пропорціональна середньому значенню грошової маси в обороті. Це можна легко показати, оскільки всі змінні задані в логарифмічному масштабі і часткова похідна

$$\frac{\partial p(k)}{\partial t} = 1.$$

Короткочасні зміни поставок грошової маси в оборот поводять себе таким чином. Наслідком останнього розв'язку є те, що зростання грошової маси в майбутньому, яке представлене значеннями $\varepsilon(k+1)$, призводять до зростання цін в поточний період часу. Але основна ідея полягає в тому, що збільшення випуску грошової маси в майбутньому буде сприяти також збільшенню цін в майбутньому, а звідси випливає, що гроші невигідно тримати без руху, вони повинні бути в обороті.

Приклад 3.4.2. Знайти загальний розв'язок рівняння

$$y(k) = -2y(k-1) - y(k-2) + k(-1)^k. \quad (3.4.1)$$

Характеристичне рівняння:

$$\lambda^2 + 2\lambda + 1 = 0$$

і його корені: $\lambda_{1,2} = \frac{1}{2}(-2 \pm \sqrt{4-4})$, тобто, $\lambda_1 = \lambda_2 = -1$. Запишемо однорідний розв'язок із врахуванням наявності кратного кореня

$$y^h(k) = A_1(-1)^k + A_2 k(-1)^k.$$

Якщо складові збурюючої функції в правій частині рівняння (3.4.1), тобто, $x(k) = k(-1)^k$ мають такий же вигляд як і складові $y^h(k)$, то пробна функція часткового розв'язку змінюється. Всі складові $y^p(k)$, які відповідають таким же складовим $y^h(k)$, необхідно помножити на k в такій найменшій степені, щоб їх тотожність порушувалась. Таким чином, пробна функція для часткового розв'язку:

$$y_{\text{проб}}^p(k) = [c_2 k + c_3 k^2] k(-1)^k = [c_2 k^2 + c_3 k^3] (-1)^k. \quad (3.4.2)$$

Підставимо (3.4.2) в (3.4.1) і знайдемо невідомі коефіцієнти c_2, c_3 за допомогою методу невизначених коефіцієнтів. Таким чином,

$$\begin{aligned}[c_2k^2 + c_3k^3](-1)^k &= -2[c_2(k-1)^2 + c_3(k-1)^3](-1)^{(k-1)} - \\ &\quad -[c_2(k-2)^2 + c_3(k-2)^3](-1)^{(k-2)} + k(-1)^k\end{aligned}$$

або

$$\begin{aligned}[c_2k^2 + c_3k^3](-1)^k &= 2[c_2(k-1)^2 + c_3(k-1)^3](-1)^k - \\ &\quad -[c_2(k-2)^2 + c_3(k-2)^3](-1)^k + k(-1)^k.\end{aligned}$$

Звідси знайдемо значення невідомих коефіцієнтів $c_2 = 1/2$; $c_3 = 1/6$ і повний розв'язок

$$y(k) = \left[c_0 + c_1k + \frac{1}{2}k^2 + \frac{1}{6}k^3 \right] (-1)^k.$$

Приклад 3.4.3. Для порівняння дискретної моделі з неперервною розглянемо розв'язок диференціального рівняння другого порядку, яке описує вимушенні коливання

$$\ddot{y}(t) + 2h\dot{y}(t) + r^2 y(t) = x(t), \quad (3.4.3)$$

де h – коефіцієнт, що характеризує властивості опору середовища (демпфірування); r – коефіцієнт жорсткості. Нехай збурююча сила змінюється за гармонічним законом $x(t) = M \sin(\omega t)$, а опір середовища покладемо для простоти нульовим, $h = 0$. Тепер рівняння (3.4.3) приймає вигляд:

$$\ddot{y}(t) + r^2 y(t) = M \sin(\omega t). \quad (3.4.4)$$

Власні коливання тіла визначаються однорідним диференціальним рівнянням

$$\ddot{y}(t) + r^2 y(t) = 0 \quad (3.4.5)$$

і описуються рівнянням гармонічних коливань

$$y(t) = A \sin(rt + \varphi).$$

Необхідно знайти частковий розв'язок рівняння (3.4.3). Вигляд часткового розв'язку залежить від того, чи буде число $a + ib = i\omega$, $i = \sqrt{-1}$, коренем характеристичного рівняння для рівняння (3.4.5). Оскільки характеристичне рівняння

$$\lambda^2 + r^2 = 0$$

має корені $\lambda_{1,2} = \pm ir$, то аналіз зводиться до випадку, коли частота впливу збурюючої сили співпадає з частотою власних коливань тіла (*резонанс*) і

коли не співпадає (*відсутність резонансу*).

Розглянемо нерезонансний випадок ($\omega \neq r$). Згідно теорії диференціальних рівнянь частковий розв'язок рівняння (3.4.4) необхідно шукати за допомогою пробної функції:

$$y^P(t) = A \cos(\omega t) + B \sin(\omega t), \quad (3.4.6)$$

де A і B – коефіцієнти, які необхідно визначити.

З останнього рівняння отримаємо:

$$\begin{aligned} y^P(t) &= A \cos(\omega t) + B \sin(\omega t), \\ \dot{y}^P(t) &= -A\omega \sin(\omega t) + B\omega \cos(\omega t), \\ \ddot{y}^P(t) &= -A\omega^2 \cos(\omega t) - B\omega^2 \sin(\omega t). \end{aligned}$$

Підставляючи другу похідну в (3.4.4), отримаємо

$$A(r^2 - \omega^2) \cos(\omega t) + B(r^2 - \omega^2) \sin(\omega t) = M \sin(\omega t).$$

З початкових умов знайдемо невідомі коефіцієнти:

$$\left. \begin{aligned} A(r^2 - \omega^2) &= 0 \\ B(r^2 - \omega^2) &= M \end{aligned} \right\} \Rightarrow A = 0, \quad B = \frac{M}{r^2 - \omega^2}.$$

Частковий розв'язок

$$y^P(t) = \frac{M}{r^2 - \omega^2} \sin(\omega t). \quad (3.4.7)$$

Загальний розв'язок приймає вигляд:

$$y^P(t) = \frac{M}{r^2 - \omega^2} \sin(\omega t) + A \sin(rt + \varphi). \quad (3.4.8)$$

Таким чином, коливання, що описуються рівнянням (3.4.8), визначаються лінійною комбінацією гармонічних коливань різної частоти.

Резонансні коливання. Явище резонансу виникає у випадку, коли частота зміни збурюючої сили співпадає з частотою власних коливань тіла ($\omega = r$). Рівняння коливань має вигляд:

$$\ddot{y}(t) + r^2 y(t) = M \sin(rt). \quad (3.4.9)$$

В даному випадку частковий розв'язок буде містити множник t :

$$y^P(t) = t[A \cos(rt) + B \sin(rt)] \quad (3.4.10)$$

і отримуємо наступний результат:

$$\begin{aligned} y^P(t) &= t[A \cos(rt) + B \sin(rt)], \\ \dot{y}^P(t) &= A \cos(rt) + B \sin(rt) + t[-Ar \sin(rt) + Br \cos(rt)], \\ \ddot{y}^P(t) &= 2[-Arsin(rt) + Br\cos(rt)] + t[-Ar^2 \cos(rt) - Br^2 \sin(rt)] \end{aligned}$$

а звідси

$$\begin{aligned} -2A\sin(rt) + 2Br\cos(rt) &= M\sin(rt); \\ \begin{cases} -2Ar = M \\ 2Br = 0 \end{cases} &\Rightarrow A = -\frac{M}{2r}, \quad B = 0, \end{aligned}$$

а тому неоднорідний розв'язок має вигляд:

$$y^p(t) = -\frac{M}{2r}t \cos(rt). \quad (3.4.11)$$

Частковий розв'язок (3.4.11) представляє собою вимушені коливання з необмеженою (роздільною) амплітудою. Гармонічні коливання знаходяться між прямими:

$$y(t) = \frac{M}{2r}t \quad \text{i} \quad y(t) = -\frac{M}{2r}t.$$

Загальний розв'язок

$$y(t) = -\frac{M}{2r}t \cos(rt) + A \sin(rt + \varphi)$$

представляє собою лінійну комбінацію коливань з необмеженою амплітудою та гармонічних коливань з постійною амплітудою тієї ж частоти.

Приклад 3.4.4. Розглянемо дискретне рівняння з гармонічним збуренням

$$y(k) = a_0 + a_1 y(k-1) + A \sin(\omega k),$$

де A і ω – відомі значення амплітуди і кругової частоти вхідного сигналу.

Виберемо пробну функцію у вигляді:

$$y_{prob}^p(k) = c_0 + c_1 k + c_2 \sin(\omega k + \varphi),$$

де φ – зсув фази сигналу при проходженні через лінійну систему.

Скористаємося тригонометричною тотожністю

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$$

і підставимо пробну функцію в рівняння процесу:

$$\begin{aligned} c_0 + c_1 k + c_2 [\sin(\omega k)\cos(\varphi) + \cos(\omega k)\sin(\varphi)] &= a_0 + a_1 \{c_0 + c_1(k-1) + \\ &+ c_2 [\sin(\omega(k-1))\cos(\varphi) + \cos(\omega(k-1))\sin(\varphi)]\} + A \sin(\omega k). \end{aligned}$$

Введемо позначення: $\cos(\varphi) = b_1$, $\sin(\varphi) = b_2$, $\sin(\omega) = b_3$, $\cos(\omega) = b_4$ і в результаті отримаємо тотожності:

$$\begin{aligned} c_0 &= a_0 + a_1 c_0 - a_1 c_1, \\ c_2 b_1 &= c_2 a_1 b_1 b_4 + c_2 a_1 b_2 b_3 + A, \\ c_2 b_2 &= -c_2 a_1 b_1 b_3 + c_2 a_1 b_2 b_4. \end{aligned}$$

Якщо $a_1 \neq 1$, то $c_1 = 0$ і з першої тотожності знайдемо, що $c_0 = a_0 / (1 - a_1)$. За допомогою другої можна визначити коефіцієнт c_2 :

$$c_2 = \frac{A}{b_1 - a_1 b_1 b_4 - a_1 b_2 b_3}.$$

Оскільки коефіцієнти b_3 і b_4 відомі, а $b_1^2 + b_2^2 = 1$, то з третьої тотожності можна знайти b_1 і b_2 і фазовий кут φ .

Таким чином, при $a_1 \neq 1$ розв'язок має вигляд:

$$y(k) = \frac{a_0}{1-a_1} + B a_1^k + \frac{A}{b_1 - a_1 b_1 b_4 - a_1 b_2 b_3} \sin(\omega k + \varphi).$$

Приклад 3.4.5. Знайдемо повний розв'язок рівняння третього порядку АРКС(3, 2):

$$\begin{aligned} y(k) = 0,5 + y(k-1) + 0,25y(k-2) - 0,25y(k-3) + \varepsilon(k) - \\ - 0,125\varepsilon(k-1) + 0,125\varepsilon(k-2). \end{aligned}$$

Запишемо характеристичне рівняння

$$\lambda^3 - \lambda^2 - 0,25\lambda + 0,25 = 0$$

і знайдемо його корені: $\lambda_1 = 1; \lambda_2 = -0,5; \lambda_3 = 0,5$ (процес з одиничним коренем). Однорідний розв'язок має вигляд:

$$y^h(k) = A_1 + A_2(-0,5)^k + A_3(0,5)^k.$$

Для знаходження часткового розв'язку виберемо пробну функцію у вигляді:

$$y_{npo\delta}^p(k) = b_0 + b_1 k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i)$$

і підставимо її в рівняння процесу, скориставшись загальними позначеннями для коефіцієнтів:

$$\begin{aligned} b_0 + b_1 k + \alpha_0 \varepsilon(k) + \alpha_1 \varepsilon(k-1) + \alpha_2 \varepsilon(k-2) + \alpha_3 \varepsilon(k-3) + \dots = \\ = a_0 + a_1 [b_0 + b_1(k-1) + \alpha_0 \varepsilon(k-1) + \\ + \alpha_1 \varepsilon(k-2) + \alpha_2 \varepsilon(k-3) + \alpha_3 \varepsilon(k-4) + \dots] + \\ + a_2 [b_0 + b_1(k-2) + \alpha_0 \varepsilon(k-2) + \alpha_1 \varepsilon(k-3) + \\ + \alpha_2 \varepsilon(k-4) + \alpha_3 \varepsilon(k-5) + \dots] + a_3 [b_0 + b_1(k-3) + \\ + \alpha_0 \varepsilon(k-3) + \alpha_1 \varepsilon(k-4) + \alpha_2 \varepsilon(k-5) + \alpha_3 \varepsilon(k-6) + \dots] + \varepsilon(k) + \beta_1 \varepsilon(k-1) + \\ + \beta_2 \varepsilon(k-2). \end{aligned}$$

Прирівняємо константи в лівій і правій частині, а також коефіцієнти при одинакових змінних:

$$\begin{aligned} b_0 &= a_0 + a_1 b_0 - a_1 b_1 + a_2 b_0 - 2a_2 b_1 + a_3 b_0 - 3a_3 b_1; \\ \text{при } k : \quad b_1 &= a_1 b_1 + a_2 b_1 + a_3 b_1; \end{aligned}$$

при $\varepsilon(k)$: $\alpha_0 = 1$;
 при $\varepsilon(k-1)$: $\alpha_1 = a_1\alpha_0 + \beta_1$;
 при $\varepsilon(k-2)$: $\alpha_2 = a_1\alpha_1 + a_2\alpha_0$;
 при $\varepsilon(k-3)$: $\alpha_3 = a_1\alpha_2 + a_2\alpha_1 + a_3\alpha_0$.

Таким чином, $b_1 \approx 0,67$, а коефіцієнт b_0 необхідно визначити за допомогою початкових умов. Повний розв'язок:

$$y(k) = A + A_2(-0,5)^k + A_3(0,5)^k + 0,67k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i),$$

де $A = A_1 + b_0$, $\alpha_i = \alpha_{i-1} + 0,25\alpha_{i-2} - 0,25\alpha_{i-3}$.

Запитання і вправи до третього розділу

1. Поясніть, для чого необхідно знаходити розв'язки різницевих рівнянь? Чи можна скористатись моделлю АРКС(p, q) для асимптотичного аналізу поведінки процесу?
2. Який процес називають процесом з одиничними коренями? Які синоніми цієї назви? Дайте визначення тренду?
3. На яку частину розв'язку впливають початкові умови? В чому проявляється вплив початкових умов?
4. Що характеризують перші і другі різниці? Запишіть рівняння АР(2) через перші різниці; для розв'язання якої задачі можна скористатись цією моделлю?
5. Який процес називають метастійким?
6. Яким чином можна звести однорідний розв'язок до нуля? Що дасть на практиці зведення однорідного розв'язку до нуля?
7. У якому випадку неоднорідний розв'язок дорівнює нулю?
8. Знайдіть однорідні розв'язки рівнянь:

$$\begin{aligned} y(k) &= 0,35 - 0,75 y(k-1) + 0,25 y(k-2) + \varepsilon(k); \\ y(k) &= a_0 + 1,5y(k-1) - 0,5y(k-2) + \varepsilon(k); \\ y(k) &= a_0 + y(k-2) + \varepsilon(k); \\ y(k) &= a_0 + 2y(k-1) - y(k-2) + \varepsilon(k); \\ y(k) &= a_0 + y(k-1) + 0,25y(k-2) - 0,25y(k-3) + \varepsilon(k). \end{aligned}$$

9. Запишіть загальний вигляд однорідного рівняння для випадку, коли характеристичне рівняння містить трикратний корінь? Чи буде такий розв'язок стійким (збіжним), дайте пояснення?
10. У яких випадках однорідний розв'язок містить гармонічну складову? Чим визначається амплітуда, частота і фаза коливань в даному випадку?
11. Що таке кругова частота? Чим зручне використання кругової частоти у порівнянні з лінійною при аналізі поведінки процесів?
12. Сформулюйте умови збіжності однорідного розв'язку через значення коренів характеристичного рівняння; чи є ці умови достатніми?
13. Сформулюйте необхідні і достатні умови збіжності однорідного розв'язку через коефіцієнти рівняння авторегресії; яким чином можна отримати ці умови?
14. Чи досить знати необхідні і достатні умови збіжності однорідного розв'язку рівняння AP(2) для того, щоб узагальнити ці умови на рівняннявищих порядків?
15. Чи відомі вам інші методи аналізу стійкості розв'язків крім методів, які ґрунтуються на коренях характеристичного рівняння та коефіцієнтах математичної моделі?
16. Які переваги має графічний аналіз збіжності розв'язків, чи можна його застосувати на практиці?
17. В чому полягає подібність та відмінності розв'язків однорідних різницевих та диференціальних рівнянь?
18. Запишіть у загальному вигляді розв'язок наведеного нижче рівняння:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + a_3 y(k-3) + b_1 \varepsilon(k-1) + \varepsilon(k).$$
 Нехай характеристичне рівняння має, в даному випадку, один дійсний і два кратних корені.
19. Нарисуйте графік однорідного розв'язку для випадку двократного кореня характеристичного рівняння?

Розділ 4

МЕТОДИКА ПОБУДОВИ МАТЕМАТИЧНИХ МОДЕЛЕЙ ЧАСОВИХ РЯДІВ

Розглянемо методику побудови математичних моделей, яка більше орієнтована на побудову моделей фінансово-економічних, соціальних та екологічних процесів, для яких, як правило, набагато складніше поставити експеримент та отримати інформативні експериментальні дані в достатньому об'ємі. В наведеному вигляді ця методика також може бути успішно застосована до побудови моделей динаміки технічних систем та технологічних процесів із врахуванням необхідності планування та виконання експерименту для збору необхідних даних.

Основи методики побудови моделей часових рядів запропоновані Боксом і Дженкінсом в роботі [7]. Модифікована авторами методика побудови математичної моделі процесу, складається з таких кроків:

- Виконання аналізу процесу, для якого будується модель, на основі спеціальних літературних джерел, експертних оцінок протікання процесу, візуального дослідження вимірюваних і вихідних змінних, представлених часовими рядами або часовими перерізами, та іншої доступної інформації.
- Попередня обробка експериментальних даних.
- Аналіз часових рядів на можливу наявність нелінійностей за допомогою множини статистичних критеріїв.
- Вибір (формування) структури моделей-кандидатів, для чого необхідно виконати такі дії: (1) обчислити та виконати аналіз кореляційної матриці для часових рядів залежної та незалежних змінних з метою визначення тих екзогенних змінних, які необхідно включити в модель; (2) обчислити автокореляційну (АКФ) та часткову автокореляційну функції (ЧАКФ) залежної змінної з метою визначення оцінки порядку авторегресійної частини моделі; (3) оцінити характеристики інших елементів структури математичної моделі, що буде розглянуто нижче.
- Вибрати метод (методи) оцінювання параметрів математичних моделей вибраних структур і оцінити вектори їх параметрів. Найчастіше це

метод найменших квадратів (МНК), метод максимальної правдоподібності (ММП) та їх модифікації (рекурсивні та нелінійні).

– Вибрати кращу з оцінених моделей-кандидатів за допомогою множини статистичних критеріїв адекватності (якості) моделі. Застосувати модель до розв'язання основної задачі – прогнозування, керування, поглибленого дослідження процесу та остаточно встановити її придатність.

Тепер розглянемо докладно кожний з етапів побудови моделі.

4.1 Аналіз процесу

Аналіз процесу – це надзвичайно важливий етап, коректне виконання якого потребує досвіду дослідження реальних процесів різної природи. Ігнорування цього етапу призводить до неможливості побудови моделі високого ступеня адекватності процесу та її придатності для розв'язання згаданих вище задач. Аналіз процесу спрямовується на розв'язання таких задач:

- визначення числа входів і виходів, тобто визначення вимірності процесу; як правило, вимірність визначається кількістю виходів процесу, кожний із яких описують окремим рівнянням;
- встановлення логічних зв'язків між змінними та аналіз можливостей їх спільного математичного описання (коректного об'єднання в одному математичному виразі); для цього необхідно використати всю наявну інформацію про процес з літератури, звітів та від експертів;
- визначення кількості зовнішніх збурень та їх типу (детерміноване чи стохастичне) та попереднє встановлення можливості їх статистичного описання за допомогою розподілів випадкових величин;
- встановлення можливості декомпозиції процесу на окремі підпроцеси, які є простішими як з точки зору їх функціонування, так і з точки зору математичного опису; декомпозиція – це досить складний процес, який ґрунтуються на спеціальних математичних методах;
- якщо процес має ієрархічну структуру (верхній та нижній рівень функціонування), то необхідно чітко розмежувати ці рівні, визначити функції кожного з них і встановити які типи зв'язків існують між ними; наприклад, технологічні процеси часто можна розмежувати на два і більше рівнів, які зв'язані між собою інформаційними потоками, за логічними ознаками і т. ін.;
- використання знань із спеціальної літератури стосовно особливостей функціонування процесу, відомих законів та закономірностей його протікання, виявлення існуючих моделей

процесу та досвіду його теоретичного чи експериментального дослідження;

- при наявності розроблених моделей досліджуваного процесу необхідно встановити їх недоліки та переваги, а також визначити можливість подальшого використання (модифікації); аналіз і використання існуючих моделей надає можливість суттєво скоротити час та інші витрати на побудову та використання моделі.

Отриману інформацію максимально використовують для попереднього оцінювання структури моделі або декількох моделей-кандидатів, параметри яких оцінюють за допомогою експериментальних даних. В процесі виконання аналізу функціонування досліджуваного процесу доцільно використовувати та порівнювати інформацію з різних джерел. Це особливо стосується фінансово-економічних процесів, щодо яких може надходити інформація з суттєвими протиріччями, пропусками і помилками.

4.2 Попередня обробка даних

Процес попередньої обробки експериментальних (статистичних) даних складається, як правило, з таких операцій:

- нормування та візуальну перевірку даних і, при необхідності, їх корегування; нормування даних означає їх логарифмування або приведення до зручного діапазону їх зміни, наприклад, від 0 до 1; від -1 до +1; від +10 до -10 і т. ін.;
- корегування даних полягає у заповненні пропусків та зменшенні викидів (екстремальних імпульсних значень), що виходять за основний діапазон значень змінних;
- бутстреп аналіз з метою збільшення об'ємів вибірок;
- заміна некоректних вимірів інтерпольованими або усередненими даними;
- формування перших або різниць вищих порядків, які необхідні для аналізу відповідних складових процесу, представленого часовом рядом;
- ортогональні перетворення та цифрова фільтрація даних з метою вилучення шумових складових.

Поширеним методом нормування даних є їх логарифмування з наступним формуванням додаткових часових рядів з перших чи других різниць. Нагадаємо, що перші різниці представляють собою наближений дискретний аналог першої похідної, а другі різниці – другої похідної. Використання різниць дає можливість будувати моделі для швидкості та прискорення основної змінної. Часто із значень ряду віднімають його середнє для того щоб отримати можливість працювати з відхиленнями, а не

повними значеннями змінних. Такий підхід застосовують, наприклад, при побудові моделей у просторі станів з їх подальшим використанням для оптимальної фільтрації або оптимального керування процесом. Застосування того чи іншого методу підготовки даних для моделювання визначається в кожному випадку по-своєму.

Обробка екстремальних (аномальних) значень

Хоча виявлення та обробка екстремальних значень – це велика окрема тема для дослідження, розглянемо деякі можливості щодо розв'язання цієї проблеми. В подальшому будемо вважати дані *аномальними*, якщо вони виникли внаслідок впливу значних похибок вимірювань або похибок, пов'язаних з некоректним збором статистичних даних. Якщо можна встановити факт наявності аномальних даних, то їх просто видаляють з розподілу.

Екстремальні значення – це правильно виміряні (зібрани) дані, які характеризують фактичні раптові (стрибкоподібні) зміни процесу. Підхід до розв'язання задачі дослідження екстремальних значень спостережень залежить від поставленої мети. Якщо дослідника цікавить тільки факт наявності таких значень (наприклад, з метою виявлення умов, що приводять до появи екстремальних значень), то досить мати надійний критерій для виявлення таких спостережень.

Якщо ж ставиться задача виявлення і виключення екстремальних значень (наприклад, з метою покращення оцінок статистичних параметрів і моделей), то виникає задача – як правильно виконати обробку даних. Так, спираючись на критерій для визначення екстремальних значень, можна визначити величину зміщення оцінок параметрів.

Критерії аналізу екстремальних значень застосовують з метою:

- вирівняти спостереження перед аналізом (як правило, суттєво зменшити) великі значення;
- переконатись, що дані містять аномальні значення, що свідчить про необхідність перегляду процедури отримання даних;
- виділити спостереження, які є цікавими з точки зору їх аномальності та, по можливості, описати встановлений ефект математично.

Класичний підхід до виявлення аномальних спостережень полягає в тому, що вибіркові спостереження розглядають як випадкові, нормальню розподілені величини. При цьому для аналізу (виявлення екстремальних значень) формується статистика (статистичний тест, який ґрунтуються на статистичних даних), яка є чутливою до різких відхилень такого типу. Необхідно встановити розподіл цієї статистики при нульовій гіпотезі, що всі спостереження належать нормальній сукупності, а потім відхилити цю гіпотезу, якщо виявиться, що обчислена статистика їй протирічить.

Розглянемо можливий критерій відкидання екстремальних значень. Нехай дана деяка вибірка $\{x_1, x_2, \dots, x_N\}$, $N \geq 3$, яка, за припущенням, є випадковою для випадкової змінної X з нормальним розподілом: $\{X\} \sim (\mu_x, \sigma_x^2)$. Позначимо відхилення від середнього через

$$\tilde{x}_i = x_i - \bar{x}, \quad i = 1, 2, \dots, N, \quad \text{де} \quad \bar{x} = \frac{1}{N} \sum_{k=1}^N x(k).$$

Якщо виділити одне значення із спостережень, то вибіркове середнє для спостережень, що залишились, визначається як

$$\sum_{\substack{k=1 \\ k \neq i}}^N \frac{x_k}{N-1} = \bar{x} - \frac{\tilde{x}_i}{N-1}. \quad (4.2.1)$$

Якщо виділити декілька значень x_1, x_2, \dots, x_r , то вибіркове середнє дорівнює:

$$\bar{x} - \frac{\tilde{x}_1 + \tilde{x}_2 + \dots + \tilde{x}_r}{N-r}. \quad (4.2.2)$$

Позначимо максимальне відхилення через $\tilde{x}_m = x_m - \bar{x}$. Тепер правило визначення екстремального значення можна сформулювати так: *при заданому значенні c спостереження x_m відкидається, якщо $|\tilde{x}_m| > c S_x$* , де S_x – вибіркове стандартне відхилення змінної X .

Якщо вибірка має досить великий об'єм, то значення x_m видаляється і аналіз продовжується. Величина константи c може змінюватись із зміною довжини вибірки; вона зв'язана неявно з t -статистикою:

$$\sqrt{\frac{N c^2 (\nu + \nu_0 - 1)}{\nu (\nu + \nu_0 - \frac{N c^2}{\nu})}} \approx t_{1-\alpha/2}^{\nu_0 + \nu - 1}, \quad (4.2.3)$$

де $\nu = N - 1$; α – рівень значущості; ν_0 – будь-яке інше число додаткових ступенів свободи, яке зв'язане з оцінюванням σ_x^2 за вибіркою, об'єм якої не дорівнює N ($\nu_0 = 0$, якщо такої інформації немає). Також існує наблизений вираз для c через розподіл F :

$$c \approx \left(\frac{v}{N} \right)^{1/2} \left(\frac{3F_{1-q}}{1 + [3F_{1-q} - 1]/(v + v_0)} \right)^{1/2}, \quad (4.2.4)$$

де $q = \Delta\hat{\sigma}_x^2 \frac{v}{N}$; $\Delta\hat{\sigma}_x^2$ – очікуваний приріст дисперсії внаслідок появи екстремальних значень. При використанні (4.2.4) значення c визначається додатнім значенням квадратного кореня.

Виразом (4.2.4) можна скористатись наступним чином: якщо з ряду значення не видалялись, то допустимий (очікуваний) відносний приріст дисперсії (“премію”) $\Delta\hat{\sigma}_x^2$ необхідно помножити на v/N і, таким чином, отримаємо q . За його допомогою знайдемо відповідну верхню процентну точку для відношення дисперсій F_{1-q} при трьох і $v + v_0 - 1$ ступенях свободи. За виразом (4.2.4) обчислимо значення c і застосуємо критерій до x_m . Очікуваний відносний приріст дисперсії (“премію”) залежить від того, наскільки ймовірною є поява екстремальних значень, наприклад, можна прийняти невеликий відносний приріст $\Delta\hat{\sigma}_x^2 = 0,01 \div 0,03$.

Наприклад, якщо $N = 4$, $v = 3$ і $v/N = 0,75$, то при $\Delta\hat{\sigma}_x^2 = 0,02$ маємо: $q = 0,02 \cdot 0,75 = 0,015$. При 3-х ступенях свободи $F_{1-q} = F_{1-0,05} = 9,28$. Тепер знайдемо значення c :

$$c = (0,75)^{1/2} \left(\frac{3F_{0,95}}{1 + (3F_{0,95} - 1)/3} \right)^{1/2} = 1,449.$$

Спостереження x_m необхідно видалити, якщо $|\tilde{x}_m| > 1,449 S_x$. Можливі інші підходи до аналізу екстремальних значень.

Приклад 4.1. Обчислення критерію для виявлення екстремального значення. Є ряд значень: $X = \{23,2 \ 23,4 \ 23,5 \ 24,1 \ 25,5\}$. Встановити, чи можна вважати значення 25,5 екстремальним і чи необхідно його видалити з вибірки?

Розв’язок. Обчислимо: $\bar{x} = 23,9$; $\tilde{x}_5 = 25,5 - 23,9 = 1,6$; $S_x = 0,77$. Для $\alpha = 0,05$; $v = 4$ і $N = 5$ за виразом (2.2.3) маємо:

$$\left(\frac{15c^2}{16 - 5c^2} \right)^{1/2} = 2,776^3.$$

Звідси, за методом проб та помилок, $c = 1,49$. Згідно з критерієм

$$|1,6| > 1,49 \cdot 0,77 = 1,05,$$

тобто спостереження x_5 видаляється.

4.3 Аналіз наявності нелінійностей

Однією з проблем при визначені структури моделі є встановлення факту наявності нелінійностей в досліджуваному процесі та їх типу. Для розв'язання цієї проблеми обов'язково використовують *візуальний аналіз* даних та формальні тести на наявність нелінійностей. Досвідченому фахівцю з моделювання візуальний аналіз дозволяє оперативно виявити наявність участків з лінійним або нелінійним трендом, в якісь мірі наявність гетероскедастичності та значних викидів (імпульсів), які можуть суттєво впливати на якість моделі. Необхідно зазначити, що існують окремі навчальні курси з візуального аналізу даних. Це свідчить про те, що не варто нехтувати такою доступною, але ефективною можливістю дослідження даних. Психологи встановили, що один інформативний рисунок може замінити до двох тисяч слів.

Існує також ряд формальних тестів на наявність нелінійності. Розглянемо простий тест для визначення наявності нелінійності. Цей тест можна застосувати у випадку, коли можна набрати кілька груп (вибірок) спостережень для одного і того ж процесу:

$$\hat{F} = \frac{\frac{1}{m-2} \sum_{i=1}^m n_i (\bar{y}_i - \hat{y}_i)^2}{\frac{1}{n-m} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2},$$

де \bar{y}_i – середнє значення для i -ї групи (вибірки або групи) даних; \hat{y}_i – середнє для лінійної апроксимації даних; m – число груп даних; n_i – число вимірювань в i -й групі; n – загальне число вимірювань. Фактично, наведена статистика представляє собою таке відношення:

$$\hat{F} = \frac{\text{Відхилення середніх значень від прямої регресії}}{\text{Відхилення значень } y(k) \text{ від групових середніх}}.$$

Якщо статистика \hat{F} з $v_1 = m - 2$ та $v_2 = n - m$ степенями свободи досягає або перевищує рівень значущості, то гіпотезу щодо лінійності необхідно відхилити. Недоліком даного підходу є те, що для його

застосування необхідно мати декілька (не менше трьох) груп даних для одного і того ж процесу, які можна отримати в результаті виконання повторних експериментів. Очевидно, що це не завжди можливо.

Наявність нелінійності можна встановити також за допомогою вибіркових *нелінійних кореляційних функцій* (НКФ), тобто, кореляційних функцій, розрахованих за вибірками експериментальних (статистичних) даних. Наприклад, якщо дискретна НКФ

$$r_{yx^2}(s) = r_{y(k)x^2(k-s)} = \frac{1}{N} \frac{\sum_{k=s+1}^N \{ [y(k) - \bar{y}] [x(k-s) - \bar{x}]^2 \}}{\sigma_y \sigma_x^2}, \quad s = 0, 1, 2, 3, \dots,$$

містить значення, які суттєво відрізняються від нуля в статистичному смислі, то процес містить квадратичну нелінійність відносно регресора x .

Наявність нелінійного детермінованого тренду в процесі можна визначити шляхом оцінювання рівняння:

$$y(k) = a_0 + c_1 k + c_2 k^2 + \dots + c_m k^m,$$

яке представляє собою поліном порядку m відносно часу. Якщо хоча б один із коефіцієнтів c_i , $i = 1, \dots, m$ є статистично значущим, то гіпотеза щодо відсутності тренду відхиляється. Якщо тренд відносно швидко змінює свій напрям розвитку і для нього трудно знайти адекватне функціональне описання, то застосовують моделі випадкових трендів, які ґрунтуються на комбінаціях випадкових величин. Наявність нелінійності стосовно регресора $x(k)$ можна встановити за допомогою відповідного полінома:

$$y(k) = a_0 + c_1 x(k) + c_2 x^2(k) + \dots + c_m x^m(k) + \varepsilon(k).$$

Автоматично оцінює структуру математичної моделі *метод групового врахування аргументів* (МГВА), запропонований академіком Івахненком О.Г. (Київський інститут кібернетики) [...]. Цей метод вже багаторазово застосовували до опису широкого класу процесів з метою оцінювання прогнозів та керування. Його успішно застосовують і сьогодні до моделювання процесів різної природи з нелінійностями та нестаціонарностями. Подальшим розвитком даного методу є нечіткий МГВА, який ґрунтується на нечіткому представленні параметрів оцінюваної моделі. Загалом, задача встановлення наявності та визначення типу нелінійності залишається предметом дослідження.

4.4 Формування інших елементів структури моделі

На третьому етапі необхідно вибрати структури моделей кандидатів. Поняття структури моделі було розглянуто вище, але нагадаємо, що воно включає *порядок моделі* (найвищий порядок рівнянь, що його утворюють); *вимірність* (число рівнянь моделі); час *запізнення* по входу (лаг) та його оцінка; *можливі нелінійності* та їх тип; *зовнішні збурення* та їх тип (детерміновані або випадкові; адитивні та мультиплікативні; імпульсні та неперервні).

Для того щоб визначити які незалежні змінні (регресори) необхідно включити в праву частину рівняння, обчислюють коефіцієнт кореляції між залежною та відповідною незалежною змінною.

Коефіцієнт кореляції, а в загальному випадку кореляційна функція, дозволяє встановити факт існування зв'язку між змінними. Кореляція може бути лінійною або нелінійною в залежності від типу функціональної залежності, яка фактично має місце між змінними. В більшості практичних випадків розглядають лінійну кореляцію (взаємозв'язок) між змінними, але більш глибокий аналіз потребує використання нелінійних залежностей. Складну нелінійну залежність можна спростити, але знати про її існування необхідно для того щоб побудувати, при необхідності, (складнішу за структурою) модель процесу з вищим ступенем адекватності.

Кореляційна матриця дає можливість встановити існування зв'язку між залежною (ендогенною) змінною та незалежними (екзогенними) змінними в правій частині. Розглянемо кореляційну матрицю \mathbf{R} вимірності 3×3 , яка будеться для трьох змінних x, y, z :

$$\mathbf{R} = \begin{bmatrix} r_{yy} & r_{xy} & r_{zy} \\ r_{yx} & r_{xx} & r_{zx} \\ r_{yz} & r_{xz} & r_{zz} \end{bmatrix}, \quad \text{де } r_{yx} = r_{xy}, \quad r_{yz} = r_{zy}, \quad r_{xz} = r_{zx}.$$

Нехай y – показник якості технологічного процесу; x, z – технологічні параметри, які, за припущенням, впливають на показник якості. Тобто, ставиться задача встановлення існування залежності вигляду:

$$y = f(x, z),$$

яка може бути представлена у формі регресії змінної y на незалежні змінні x, z :

$$y(k) = a_0 + a_1 x(k) + a_2 z(k) + \varepsilon(k),$$

де k – дискретний час (наприклад, в долях секунди, секундах, хвилинах, годинах, днях, тижнях, місяцях і т.і.); $\varepsilon(k)$ – випадкова змінна, введення якої в модель пояснюється наступними причинами:

- часто буває неможливо встановити всі незалежні змінні, які впливають на залежну змінну, а тому наведене рівняння описує процес з похибкою;
- можуть існувати такі незалежні змінні, які неможливо виміряти і включити в модель, а тому їх розглядають як збурення і вважають, що їх спільний вплив на залежну змінну описується випадковою змінною $\varepsilon(k)$;
- в наведене вище регресійне рівняння можуть бути введені пояснюючі змінні, які є формально корельованими із залежною змінною, але фактично вони не впливають на неї;
- для будь-якого методу оцінювання параметрів рівнянь властиві методичні похибки, які повинні бути, по можливості, враховані в моделі.

Вважається, що сукупний вплив всіх вказаних факторів можна описати в деякій мірі за допомогою випадкової змінної $\varepsilon(k)$. Оскільки вона не вимірюється, то оцінити її значення (*похибку* моделі або *залишок*) можна тільки після оцінювання коефіцієнтів моделі, тобто

$$\hat{\varepsilon}(k) = e(k) = y(k) - \hat{y}(k),$$

де $\hat{y}(k)$ – оцінка змінної $y(k)$, отримана за допомогою моделі; $y(k)$ – фактичний вимір.

Для обчислення елементів матриці \mathbf{R} необхідно мати синхронізовані в часі вибірки значень всіх трьох змінних y, x, z . Формула для розрахунку коефіцієнтів кореляції має вигляд:

$$r_{yx} = \frac{\frac{1}{N-1} \sum_{k=1}^N \{[x(k) - \bar{x}][y(k) - \bar{y}]\}}{\sigma_x \sigma_y},$$

де \bar{x}, \bar{y} – вибіркові середні значення змінних x, y ; σ_x, σ_y – стандартні відхилення цих змінних, тобто корені квадратні з їх дисперсії. Наприклад,

$$\sigma_y = \sqrt{\sigma_y^2} = \left[\frac{1}{N-1} \sum_{k=1}^N [y(k) - \bar{y}]^2 \right]^{1/2},$$

де N – число вимірів змінної y ; \bar{y} – вибіркове середнє значення ряду $\{y(k)\}$, яке обчислюється за відомою формулою:

$$\bar{y} = \frac{1}{N} \sum_{k=1}^N y(k).$$

Очевидно, що перед формальним обчисленням коефіцієнтів кореляції, необхідно зробити аналіз процесу і визначити наявність (або відсутність) логічних зв'язків між змінними. Це дозволяє ввести до розгляду тільки ті змінні, які дійсно впливають на залежну змінну, наприклад, на показник якості. Очевидно, що для правильного вибору незалежних (екзогенних) змінних необхідно глибоко знати технологічний або інший процес, що моделюється.

На основі значень коефіцієнтів кореляції приймається рішення про включення їх у рівняння регресії:

$$y(k) = a_0 + b_1 x(k) + b_2 z(k) + \varepsilon(k),$$

яке може бути представлено в загальному вигляді як

$$y(k) = a_0 + a_1 x_1(k) + a_2 x_2(k) + a_3 x_3(k) + \dots + a_{p-1} x_{p-1}(k) + \varepsilon(k).$$

Можна показати, що між коефіцієнтами регресії b_1, b_2 і коефіцієнтами кореляції r_{yx}, r_{yz} існує однозначний аналітичний взаємозв'язок.

Останнє рівняння представляє собою *лінійну регресію*, яка містить p параметрів (коефіцієнтів), але досить часто необхідно застосовувати складніші нелінійні моделі. Характерним представником нелінійної стосовно змінних регресії є поліноміальна регресія порядку $p - 1$:

$$y(k) = a_0 + a_1 x(k) + a_2 x^2(k) + a_3 x^3(k) + \dots + a_{p-1} x^{p-1}(k) + \varepsilon(k).$$

Хоча в це рівняння включено тільки одну незалежну змінну $x(k)$, очевидно, що воно може бути розширене будь-якими іншими змінними.

Для визначення необхідності введення в рівняння регресії авторегресійної складової необхідно обчислити і дослідити вибіркову *автокореляційну функцію* змінної $y(k)$. Рівняння з авторегресійною складовою має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + b_1 x(k) + b_2 z(k) + \varepsilon(k),$$

тобто, в рівняння регресії введено авторегресійну (AP) складову другого порядку. Порядок авторегресії визначається за допомогою автокореляційної функції (АКФ) і часткової АКФ (ЧАКФ). Кількість значень автокореляційної функції, які відмінні від нуля в статистичному смислі, і буде складати оцінку порядку авторегресії.

Природа авторегресії пояснюється існуванням так званої “пам’яті” процесу, яка проявляється в тому, що його поточний стан в значній мірі визначається попередніми станами. Наприклад, стан людини вранці залежить від того, яким було самопочуття ввечері та в попередні дні. На формування ринкових цін суттєво впливають значення цін в попередні періоди часу. Очевидно, що поточний рівень валового внутрішнього продукту (ВВП) залежить від попередніх значень, а поточний стан технічної системи або технологічного процесу також залежить від їх стану в попередні моменти часу.

АКФ та ЧАКФ використовують для визначення попередньої оцінки порядку авторегресійної частини моделі, тобто скільки затриманих в часі значень необхідно брати для опису процесу. При цьому необхідно врахувати, що АКФ дає менш «чітку» оцінку порядку процесу ніж ЧАКФ. Наприклад, для процесу AP(1) значення основної змінної $y(k)$ та $y(k-2)$ будуть корельованими, незважаючи на те, що $y(k-2)$ не присутнє в моделі. Кореляція між $y(k)$ і $y(k-2)$, тобто ρ_2 , дорівнює коефіцієнту кореляції між значеннями $y(k)$ і $y(k-1)$, помноженому на коефіцієнт кореляції між $y(k-1)$ і $y(k-2)$ або $\rho_2 = \rho_1 \rho_1 = \rho_1^2$. Подібні «непрямі» кореляції наявні в АКФ будь-якого процесу авторегресії.

Вибіркова АКФ обчислюється за виразом:

$$r_y(s) = r_{y(k)y(k-s)} = \frac{1}{N-1} \frac{\sum_{k=s+1}^N \{ [y(k) - \bar{y}] [y(k-s) - \bar{y}] \}}{\sigma_y^2}, \quad s=1,2,3,\dots,$$

де σ_y^2 – вибіркова дисперсія змінної $y(k)$; \bar{y} – середнє значення вибірки даних.

Кількість значень АКФ, відмінних від нуля в статистичному смислі, вказує на порядок авторегресійної частини моделі. Для стаціонарного процесу (це процес із постійними середнім, дисперсією та коваріацією) коефіцієнти $r_y(s)$ мають нормальній розподіл та нульове середнє.

На відміну від АКФ, часткова АКФ між значеннями $y(k)$ та $y(k-s)$ виключає вплив величин $y(k-1) \dots y(k-s+1)$, а це означає, що коефіцієнти ЧАКФ чіткіше відображають зв'язок між окремими значеннями основної змінної. Так, для процесу AR(1) ЧАКФ між $y(k)$ та $y(k-2)$ дорівнює нулю за визначенням, що підтверджується обчисленими значеннями ЧАКФ. Для того, щоб знайти попередню оцінку порядку моделі, вибіркові коефіцієнти ЧАКФ (тобто, коефіцієнти, знайдені за вибіркою даних) можна обчислити також за допомогою простого методу, який полягає в наступному.

а) Формують додатковий часовий ряд з відхилень основної змінної:

$$\{y'(k)\} = \{y(k)\} - \mu,$$

де μ – середнє значення ряду.

б) Формують рівняння першого порядку

$$y'(k) = \Phi_{11} y'(k-1) + e(k),$$

де $e(k)$ – похибка моделі. В такому рівнянні Φ_{11} відіграє роль коефіцієнта АКФ та ЧАКФ між $y(k)$ та $y(k-1)$. Для оцінювання двох коефіцієнтів можна сформувати рівняння другого порядку:

$$y'(k) = \Phi_{11} y'(k-1) + \Phi_{22} y'(k-2) + e(k),$$

де Φ_{22} – коефіцієнт ЧАКФ між $y(k)$ та $y(k-2)$.

Коефіцієнти ЧАКФ можна обчислити також за допомогою коефіцієнтів АКФ, використовуючи наступні вирази [...]:

$$\Phi_{11} = r(1), \quad \Phi_{22} = \frac{r_2 - r_1^2}{1 - r_1^2};$$

$$\Phi_{ss} = \frac{r_s - \sum_{j=1}^{s-1} \Phi_{s-1,j} r_{s-j}}{1 - \sum_{j=1}^{s-1} \Phi_{s-1,j} r_j}.$$

В загальному випадку коефіцієнти ЧАКФ стаціонарного процесу AR(p,q) повинні збігатися до нуля, починаючи із p -го значення. АКФ процесу AR(p,q) починається збігатися до нуля при значеннях зміщення $s \geq q$.

Коли ми говоримо, що значення автокореляційної функції повинні бути відмінними від нуля в статистичному смыслі, це означає, що існує вираз (формула), який дозволяє підтверджити або спростовувати цей факт. Одним із загальноприйнятих підходів до встановлення того факту, що значення АКФ суттєво відмінні від нуля в статистичному смыслі, є обчислення та аналіз значущості статистичного параметра (або просто статистики) Льюнга-Бокса $Q(r_k)$ за формулою [7]:

$$Q(r_k) = N(N+2) \sum_{k=1}^s r_k^2 / (N-k),$$

де N – довжина вибірки даних змінної, для якої обчислено значення автокореляційної функції r_k ; s – число відліків АКФ, які досліджуються на суттєву відмінність від нуля. Якщо дані згенеровані процесом АР чи АРКС, то значення $Q(r_k)$ асимптотично мають розподіл χ^2 з s ступенями свободи, а тому для перевірки їх значущості необхідно користуватись відповідними статистичними таблицями. Очевидно, що більші значення вибіркової автокореляційної функції приводять до більших значень $Q(r_k)$.

Так, для ідеального процесу білого шуму $Q(r_k)=0$. Якщо значення $Q(r_k)$, обчислене за наведеним виразом, перевищує критичне значення з розподілу χ^2 з s ступенями свободи, то існує щонайменше одне значення $r(k)$, яке є відмінним від нуля в статистичному смыслі. Статистику Льюнга-Бокса можна застосовувати також для встановлення близькості залишків моделі до білого шуму. Однак, необхідно пам'ятати, що при обчисленні s значень кореляційної функції число ступенів вободи зменшується на кількість коефіцієнтів моделі. Таким чином, при аналізі залишків моделі АРКС(p,q) статистика $Q(r_k)$ має розподіл χ^2 з $s-p-q$ ступенями свободи, а із врахуванням константи $s-p-q-1$.

Цей етап закінчується формуванням структур декількох моделей-кандидатів з векторами параметрів $\theta_1, \dots, \theta_m$, де m – число кандидатів. Кандидатів може бути декілька, оскільки встановити структуру точно за один раз, як правило, неможливо. Загалом побудова моделі високого ступеня адекватності – це трудомісткий ітераційний процес, який вимагає значних зусиль. На наступному етапі оцінюють параметри моделей-кандидатів.

4.5 Оцінювання коефіцієнтів моделей-кандидатів

На четвертому етапі оцінюють коефіцієнти (параметри) рівняння, використовуючи принцип економії або збереження. Цей принцип означає,

що кількість коефіцієнтів, що оцінюються, не повинна перевищувати їх необхідне число (“необхідність” можна визначити, наприклад, як необхідність збереження в моделі основних статистичних характеристик процесу).

При моделюванні процесів будь-якої природи необхідно пам'ятати, що поведінку процесу необхідно *апроксимувати* за допомогою рівнянь, а не старатися описати її до найменших дрібниць. Необхідно враховувати також, що різні за структурою моделі можуть мати одинакові властивості. Наприклад, рівняння авторегресії першого порядку

$$y(k) = 0,5y(k-1) + \varepsilon(k)$$

еквівалентне ковзному середньому у вигляді

$$y(k) = \varepsilon(k) + 0,5\varepsilon(k-1) + 0,25\varepsilon(k-2) + 0,125\varepsilon(k-3) + \dots .$$

Модель, що оцінюється, повинна задовольняти принципу інверсії, тобто щоб за допомогою отриманого рівняння можна було б згенерувати початковий ряд, на основі якого оцінювались коефіцієнти. Це означає, що хоча модель і спрощена, вона повинна співпадати з досліджуваним процесом за такими основними характеристиками як середнє, дисперсія та коваріація.

В процедурі оцінювання часто використовують не абсолютні значення змінних, а їх відхилення від середнього, тобто

$$y(k) = Y(k) - \mu_y,$$

де $Y(k)$ – значення виміру, μ_y – середнє значення ряду. Якщо для оцінювання параметрів використовується рекурсивна процедура, то поточне середнє можна обчислювати за формулою:

$$\mu_y(k) = \mu_y(k-1) + \frac{1}{k} [y(k) - \mu_y(k-1)].$$

Найбільш поширеними методами оцінювання параметрів моделі є такі:

- *метод найменших квадратів* (МНК);
- *метод максимальної правдоподібності* (ММП),
- *метод допоміжної (інструментальної) змінної* (МДП),
- *нелінійний метод найменших квадратів* (НМНК)

та їх рекурсивні версії (РМНК, РММП, РМДП). Деякі методи будуть розглянуті в розділі, присвяченому методам оцінювання. Оцінки (звичайного) МНК обчислюють за допомогою такого виразу:

$$\hat{\theta} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y},$$

де $\theta [p]$ – вектор оцінок параметрів вимірності p ; $\mathbf{X}[N \times p]$ – матриця вимірів; $\mathbf{y}[N]$ – вектор вимірів залежної змінної. В квадратних дужках вказана вимірність векторів і матриці. Елементи матриці вимірів обчислюються по-своєму для кожної конкретної моделі. Так, для моделі

$$y(k) = a_0 + a_1 x_1(k) + a_2 x_2(k) + a_3 x_3(k) + \varepsilon(k)$$

матриця вимірів має вигляд:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1(1) & x_2(1) & x_3(1) \\ 1 & x_1(2) & x_2(2) & x_3(2) \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_1(N) & x_2(N) & x_3(N) \end{bmatrix}.$$

Одници в першому стовпчику матриці \mathbf{X} означають, що вимір при коефіцієнті a_0 завжди дорівнює одиниці.

Елементи матриці вимірів дещо ускладнюються у випадку використання поліноміальної моделі, але її також можна оцінювати за допомогою лінійних методів. Безпосереднє застосування методу мінімізації суми квадратів похибок до поліноміальної моделі порядку p приводить до формування такої матриці вимірів:

$$\mathbf{X}' = \begin{bmatrix} N & \sum_{k=1}^N x(k) & \sum_{k=1}^N x^2(k) & \dots & \sum_{k=1}^N x^p(k) \\ \sum_{k=1}^N x(k) & \sum_{k=1}^N x^2(k) & \sum_{k=1}^N x^3(k) & \dots & \sum_{k=1}^N x^{p+1}(k) \\ \cdot & \cdot & \cdot & \dots & \dots \\ \sum_{k=1}^N x^p(k) & \sum_{k=1}^N x^{p+1}(k) & \sum_{k=1}^N x^{p+2}(k) & \dots & \sum_{k=1}^N x^{2p}(k) \end{bmatrix}.$$

При такому представленні векторно-матричне рівняння для N вимірів залежної та незалежної змінних можна записати так: $\mathbf{y}' = \mathbf{X}' \theta$; звідси

$$\hat{\theta} = [\mathbf{X}']^{-1} \mathbf{y}', \text{ де } \mathbf{y}' = \left[\sum_1^N y(k) \quad \sum_1^N x(k) y(k) \quad \dots \quad \sum_1^N x^p(k) y(k) \right]^T;$$

$\hat{\theta}$ – вектор оцінок параметрів моделі. Тобто оцінку вектора параметрів можна знайти шляхом знаходження розв'язку системи лінійних (нормальних) рівнянь.

Для отримання незміщених, консистентних та ефективних оцінок вектора параметрів θ лінійної регресійної математичної моделі, наприклад, моделі змішаної регресії:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + b_1 x(k) + b_2 z(k) + \varepsilon(k)$$

за допомогою методу найменших квадратів необхідно задовольнити наступні умови:

- а) $\varepsilon(k)$ – некорельована послідовність випадкових чисел з нульовим середнім, тобто, $E[\varepsilon(k)] = 0$, $\text{cov}[\varepsilon(k)] = E[\varepsilon(k)\varepsilon(j)] = \begin{cases} \sigma_\varepsilon^2, & k = j; \\ 0, & k \neq j. \end{cases}$
- б) послідовності $\varepsilon(k)$ і $y(k)$ не повинні бути корельовані між собою.

Зазначимо, що перевірити виконання наведених умов ми можемо тільки після оцінювання коефіцієнтів моделі, а до оцінювання можна тільки постулювати їх виконання. Тобто, після оцінювання моделі оцінка значень випадкового процесу визначається похибками моделі:

$$\hat{\varepsilon}(k) = e(k) = y(k) - \hat{y}(k),$$

що дає можливість виконати аналіз характеристик випадкового процесу $\{\varepsilon(k)\}$.

4.6 Діагностика моделей – вибір кращої з множини оцінених кандидатів

На п'ятому етапі аналізується якість моделі, тобто виконується перевірка оцінених кандидатів на адекватність процесу. Діагностика складається з кроків, наведених нижче.

- а) *Візуальне дослідження графіка похибок* моделі $e(k) = y(k) - \hat{y}(k)$, де $\hat{y}(k)$ – оцінка змінної, отримана за допомогою побудованої моделі. На графіку не повинно бути значних викидів та довгих інтервалів, на яких похибка приймає великі значення (тобто довгих інтервалів суттєвої неадекватності). У випадку застосування рекурсивних методів оцінювання найбільші похибки будуть в переходному процесі, коли інформаційна матриця ще не містить достатньо інформації про процес.

б) *Похибки моделі не повинні бути корельовані між собою.* Для аналізу наявності кореляції між значеннями похибок необхідно обчислити АКФ та ЧАКФ для ряду $\{e(k)\}$ і за допомогою Q -статистики визначити ступінь корельованості (наприклад, Q -статистика вважається несуттєвою до рівня 10%).

Крім того, корельованість похибок визначають за допомогою статистики Дарбіна-Уотсона (DW), яка розраховується за формулою:

$$DW = 2 - 2\rho,$$

де $\rho = E[e(k)e(k-1)]/\sigma_e^2$ – коефіцієнт кореляції між сусідніми значеннями похибки; σ_e^2 – дисперсія послідовності похибок $\{e(k)\}$. Таким чином, при повній відсутності кореляції між похибками $DW = 2$ – це ідеальне значення. Границями значеннями для DW є 0 (при $\rho = 1$) та +4 (при $\rho = -1$).

Отримати формулу $DW = 2 - 2\rho$ можна досить просто. Автори цієї статистики (Durbin i Watson) запропонували скористатись для перевірки корельованості похибок моделі таким виразом:

$$DW = \frac{\sum_{k=2}^N [e(k) - e(k-1)]^2}{\sum_{k=1}^N e^2(k)} = \frac{\sum_{k=2}^N [e(k) - e(k-1)][e(k) - e(k-1)]}{\sum_{k=1}^N e^2(k)},$$

тобто, DW можна, в деякій мірі, трактувати як коефіцієнт автокореляції для (перших різниць) приростів похибок.

Розкриваючи квадрат різниці в чисельнику, отримаємо:

$$DW = \frac{\sum_{k=2}^N e^2(k)}{\sum_{k=1}^N e^2(k)} + \frac{\sum_{k=2}^N e^2(k-1)}{\sum_{k=1}^N e^2(k)} - 2 \frac{\sum_{k=2}^N e(k)e(k-1)}{\sum_{k=1}^N e^2(k)},$$

$$\text{де } \frac{\sum_{k=2}^N e^2(k)}{\sum_{k=1}^N e^2(k)} \approx 1; \quad \frac{\sum_{k=2}^N e^2(k-1)}{\sum_{k=1}^N e^2(k-1)} \approx 1; \quad \text{а } \frac{\sum_{k=2}^N e(k)e(k-1)}{\sum_{k=1}^N e^2(k-1)} = \rho.$$

Тому можна записати, що $DW = 2 - 2\rho$.

в) Для лінійної моделі 2-3 порядку *оцінки параметрів повинні збігатися до усталених значень* після 30-40 (не більше) ітерацій алгоритму оцінювання. Якщо кількість ітерацій набагато перевищує вказані числа, то це свідчить про те, що процес може бути нестационарним.

г) Перевірка значущості оцінок параметрів моделі. *Статистика Стьюдента або t-статистика* (випадкова величина, що має t-розподіл), яка використовується для визначення значущості оцінки кожного коефіцієнта в статистичному смыслі, визначається за виразом:

$$t = \frac{\hat{a} - a^0}{SE_{\hat{a}}},$$

де \hat{a} – оцінка коефіцієнта моделі; a^0 – нуль-гіпотеза (початкова гіпотеза) щодо цієї оцінки; $SE_{\hat{a}}$ – стандартна похибка оцінки. За нуль-гіпотезу щодо значущості оцінки можна висувати будь-яку: що коефіцієнт значущий, тобто, $(H_0 : a^0 \neq 0)$ або незначущий $(H_0 : a^0 = 0)$. Статистична теорія перевірки гіпотез пропонує висувати нуль-гіпотезу, яка є протилежною бажаному результату. В даному випадку бажаним результатом є значущість коефіцієнтів математичної моделі. Таким чином, необхідно висувати нульову гіпотезу, що коефіцієнт незначущий. Це дає можливість коректно підійти до визначення значущості оцінок коефіцієнтів та дещо спростити розрахунки.

Для того щоб встановити, чи є оцінка коефіцієнта значущою, необхідно знати довжину вибірки даних N (потужність вибірки); число ступенів свободи $f = N - n$, де n – число коефіцієнтів моделі, які оцінюються на основі ряду даних, і вибрati рівень значущості $\alpha = 1\%$ або $\alpha = 5\%$ або $\alpha = 10\%$ (для цих значень існують розраховані таблиці для критичних значень t-статистики). Фактично, рівень значущості означає ймовірність припуститись *помилки першого роду* при перевірці гіпотези. Згадаємо, що

$$\alpha = p\{X \in G/\omega | H_0\} = \int_{n-m(G/\omega)} L_{H_0}(X) dx,$$

де $X = [x_1, \dots, x_n] \in R^n$ – вся вибірка, яка розбивається на дві множини, що перетинаються: ω і G/ω (ω – область прийняття нуль-гіпотези); G/ω – критична область: якщо $X \in G/\omega$, то H_0 відхиляється; $L_{H_0}(X)$ – закон розподілу X . Помилка першого роду означає відхилення вірної гіпотези.

Користуючись значеннями N, f і α , з таблиць для t – розподілу знаходять критичне значення t – статистики, тобто t_{kp} . Для перевірки правильності висунутої гіпотези розраховане значення t порівнюють з критичним t_{kp} . Якщо

$$-t_{kp} < t < t_{kp} \quad \text{або} \quad |t| < |t_{kp}|,$$

то нуль-гіпотеза щодо незначущості коефіцієнта приймається (його можна не враховувати в регресії). Звідси випливає, що чим більшим є значення t – статистики для оцінки коефіцієнта, тим імовірніше, що цей коефіцієнт є значущим.

Загалом послідовність дій при перевірці значущості оцінок коефіцієнтів побудованої моделі можна сформулювати так:

- сформулювати нуль-гіпотезу щодо значущості коефіцієнта;
- обчислити значення t – статистики для кожного коефіцієнта регресії (це робить кожний пакет для математичного моделювання);
- за допомогою значень N, f і α знайти із таблиць для t – статистики її критичне значення;
- перевірити нуль-гіпотезу за наведеним вище простим правилом (аналіз виконання нерівності $-t_{kp} < t < t_{kp}$).

д) Коефіцієнт множинної детермінації R^2 , який обчислюється так:

$$R^2 = \frac{\text{var}(\hat{y})}{\text{var}(y)} = 1 - \frac{SSE}{SST},$$

де $\text{var}(\hat{y})$ – дисперсія залежності змінної, оціненої за допомогою побудованої моделі; $\text{var}(y)$ – дисперсія вимірюваних залежності змінної;

$SSE = \sum_{k=1}^N [y(k) - \hat{y}(k)]^2$ – сума квадратів похибок (залишків) моделі (*sum of squared errors*); $SST = \sum_{k=1}^N [y(k) - \bar{y}]^2$ – загальна сума квадратів (*total sum of squares*); \bar{y} – середнє значення; $SST = SSE + SSR$, де $SSR = \sum_{k=1}^N [\hat{y}(k) - \bar{y}]^2$ – загальна сума квадратів для регресії (*sum of squares for regression*).

Очевидно, що найкращим значенням є $R^2 = 1$, тобто, коли дисперсії вимірюваних змінної, та цієї ж змінної, оціненої за рівнянням, збігаються. Цей параметр можна трактувати, також, як міру інформативності моделі, якщо за

міру інформативності вибрати дисперсію. Таким чином, R^2 показує рівень інформативності моделі по відношенню до інформативності вибірки даних, за допомогою якої вона була оцінена.

е) Сума квадратів похибок для вибраної моделі повинна бути мінімальною, тобто

$$\sum_{k=1}^N e^2(k) = \sum_{k=1}^N [\hat{y}(k) - y(k)]^2 \rightarrow \min_{\hat{\theta}}$$

у порівнянні з усіма іншими моделями.

е) Для оцінки адекватності моделі також використовують інформаційний критерій Акайке:

$$AIC = N \ln \left(\sum_{k=1}^N e^2(k) \right) + 2n,$$

та критерій Байєса-Шварца:

$$BSC = N \ln \left(\sum_{k=1}^N e^2(k) \right) + n \ln(N),$$

де $n = p + q + 1$ – кількість параметрів моделі, які оцінюються за допомогою статистичних даних (p - число параметрів авторегресійної частини моделі; q - число параметрів ковзного середнього; 1 з'являється тоді, коли оцінюється зміщення (або *перетин*, тобто a_0).

Критерії Акайке і Байєса-Шварца містять в правій частині суму квадратів похибок, а тому за цими критеріями вибирають ту модель, для якої критерії приймають найменші значення. Введення нового регресора приводить до збільшення критерію (при цьому збільшується n), але одночасно зменшується сума квадратів похибок і критерій в цілому зменшується. Якщо регресор не покращує модель, то критерій збільшується. Необхідно також зазначити, що асимптотичні властивості для довгих виборок кращі у критерія Байєса-Шварца, тобто, його рекомендують застосовувати при відносно великих значеннях N ($N > 100$).

ж) Окрім згаданих параметрів, для визначення адекватності моделі в цілому використовують *F*-статистику Фишера, яка пропорціональна відношенню:

$$F \sim \frac{R^2}{1-R^2},$$

а для множинної (багатофакторної) регресії вона визначається за виразом

$$F = \frac{R^2}{1-R^2} \cdot \frac{(N-p-1)}{p},$$

де, як і раніше, N – число значень ряду; p – число параметрів моделі без врахування перетину (константи).

Таким чином, якщо $R^2 \rightarrow 1$, то $F \rightarrow \infty$. Порядок застосування F -статистики такий же, як і t -статистики. Нуль-гіпотезою є в даному випадку припущення про те, що модель неадекватна в цілому, тобто,

$$H_0 : a_1 = a_2 = \dots = a_p = 0$$

проти альтернативної гіпотези

H_1 : хоча б одне значення a_i відмінне від нуля в статистичному смислі.

Значення $F_{\text{крит}}$ знаходять із таблиць для F -розподілу. Послідовність застосування цієї статистики можна представити наступним чином:

1. Сформулювати нуль-гіпотезу стосовно адекватності моделі в цілому.

Наприклад, H_0 : модель неадекватна в цілому (або

$$H_0 : a_1 = a_2 = \dots = a_p = 0).$$

2. Розрахувати значення F для оціненої моделі (як правило, воно розраховується всіма пакетами статистичної обробки даних).
3. Задати рівень значущості $\alpha = 1\%$ або $\alpha = 5\%$ або $\alpha = 10\%$.
4. Користуючись значеннями N , f і α , знайти критичне значення $F_{\text{крит}}$ знаходять із таблиць для F -розподілу при $(p, N-p-1)$ степенях свободи.
5. Перевірити нуль-гіпотезу:
якщо $F > F_{\text{крит}}$, то нуль-гіпотеза щодо неадекватності моделі в цілому відкидається на вираному рівні значущості.

Коректне застосування запропонованої методики забезпечує побудову адекватної математичної моделі процесу, якщо експериментальні дані

відповідають *вимогам представництва та інформативності*. Перша вимога означає, що вибірка даних повинна охоплювати досить довгий проміжок часу, щоб повністю відображати поведінку того режиму функціонування процесу, для яких буде створюватися модель. Вимога *інформативності* означає, що вибірка повинна містити в собі об'єм інформації, достатній для оцінювання коефіцієнтів моделі. Наприклад, якщо моделюється процес другого порядку, то вибірка повинна забезпечувати коректне обчислення першої та другої похідної. Іноді інформативність формально оцінюють за допомогою величини дисперсії процесу, а також за кількістю гармонічних складових, які містяться в процесі. Чим більше гармонік містить вибірка, тим вищою є її інформативність.

Умову інформативності даних пов'язують з умовою *достатнього збудження* процесу. Достатнє збудження означає, що вхідний сигнал повинен охоплювати всю смугу частот, які може пропускати на вихід процес (об'єкт). Тобто вхідний вплив (сигнал) повинен охоплювати всю амплітудно-частотну характеристику процесу. Ця вимога залишається справедливою для процесів будь-якої природи.

4.7 Приклади побудови математичних моделей за допомогою експериментальних даних

Приклад 4.2. Побудуємо математичні моделі динаміки для наступних макроекономічних процесів України: формування внутрішнього валового продукту (ВВП), індекс споживчих цін (ІСЦ) і грошовий агрегат М3. Для побудови використано фактичні щомісячні дані з січня 1996 по січень 2005 року, всього 109 значень. Кореляційна матриця для цих змінних:

ІСЦ	ВВП	М3
1	-0.3043610	-0.2491123
-0.3043610	1	0.9317529
-0.2491123	0.9317529	1

Корельованість ІСЦ з ВВП та агрегатом М3 незначна; в подальшому ця інформація буде використана при побудові альтернативних варіантів математичних моделей процесів, що розглядаються. Корельованість між ВВП і агрегатом М3 складає 0,931 – велике значення, яке може негативно вплинути на якість оцінок моделі при використанні цих змінних у правій частині рівняння.

4.7.1 Модель індексу споживчих цін

Авторегресійна модель індексу споживчих цін

Спочатку розглянемо можливість описання індексу споживчих цін за допомогою моделі авторегресії з ковзним середнім. Авторегресійні моделі самі прості за своєю структурою, але досить часто вони мають високу ступінь адекватності досліджуваному процесу, прийнятну для подальшого використання. Автокореляційна функція процесу наведена в таблиці 4.1. При побудові моделі індекс споживчих цін позначимо як isc .

Таблиця 4.1. Автокореляційна функція процесу формування споживчих цін (ІСЦ)

Часові дані: 1996:01 2005:01

Всього спостережень: 109

АКФ	Часткова АКФ	АКФ	ЧАК Ф	Q- стат.	Ймо в
. ****	. ****	1 0.57	0.57	36.37	0.00
		0 0	0	3 0	
. **	. *.	2 0.23	-	42.37	0.00
		0 0.14		8 0	
			0		
. *	. *	3 0.12	0.08	44.28	0.00
		9 9	9	4 0	
. *	. .	4 0.07	-	44.95	0.00
		6 0.02		7 0	
			2		
. .	. .	5 0.06	0.04	45.43	0.00
		4 3	3	3 0	
. *	. *	6 0.12	0.10	47.13	0.00
		0 1	7	0 0	
. *	. *	7 0.18	0.08	50.96	0.00
		0 8	0	0 0	
. *	. *.	8 0.06	-	51.51	0.00
		8 0.13		5 0	
			6		
. .	. *	9 0.05	0.10	51.84	0.00
		2 5	0	0 0	
. *	. .	1 0.07	-	52.46	0.00
		0 2 0.00		8 0	
			1		

В таблиці 4.1 скорочення АКФ – автокореляційна функція, а ЧАКФ – часткова АКФ.

Як видно з АКФ, при побудові моделей необхідно починати з моделей нижчих порядків, які часто мають прийнятну адекватність процесу і

забезпечують високу якість прогнозу. Результати оцінювання моделі авторегресії першого порядку за методом найменших квадратів наведені в таблиці 4.2.

Таблиця 4.2. Результати оцінювання моделі AP(1) для ІСЦ

Метод оцінювання: метод найменших квадратів
 Скорегована часова вибірка даних: 1996:02 2005:01
 Використано спостережень: 108 після корегування
 Модель AP(1): $I = C(1) + C(2)*I(-1)$

	Оцінки коef-в	Станд. похибка	t-статистика	Ймов.
C(1)	43.29386	6.624742	6.535178	0.0000
C(2)	0.569826	0.065703	8.672790	0.0000
R-квадрат	0.415067	Середнє залежної змінної		100.7407
Скорегований R- квадрат	0.409549	Станд. відхил. зал. змінної		1.506189
Станд. похибка регресії	1.157368	Інформ. критерій Акайке		3.148519
Сума квадратів похибок	141.9871	Критерій Шварца		3.198188
		Стат. Дарбіна-Уотсона		1.931805

Отримана модель AP(1):

$$\begin{aligned} isc(k) &= a_0 + a_1 isc(k-1) + \varepsilon(k) = \\ &= 43,294 + 0,57 isc(k-1) + e(k), \end{aligned}$$

де $e(k)$ – залишки (похибки) моделі, значення яких можна знайти у відповідному файлі пакету програм, що застосовується для побудови моделі. Статистичні характеристики моделі:

$$R^2 = 0,415; \quad J = СКП = 141,99; \quad DW = 1,931.$$

Коефіцієнт детермінації має низьке значення (0,415), сума квадратів похибок досить високе (141,99), а статистика Дарбіна-Уотсона (1,931) наближається до найкращого значення. Таким чином, загалом адекватність моделі AP(1) досить низька, а тому структура моделі потребує уточнення. Характеристики якості (історичного – на навчальній вибірці) однокрокового (історичного) прогнозу:

$$CeK\pi = 1,36; \quad CA\pi = 1,02; \quad CA\pi\pi = 1,008; \quad U = 0,0067,$$

тобто середньоквадратична похибка ($CeK\pi$), середня абсолютна похибка ($CA\pi$), середня абсолютна похибка в процентах ($CA\pi\pi$) і коефіцієнт Тейла, який свідчить про загальну придатність моделі для прогнозування (ідеальне значення – нуль).

В таблиці 4.3 наведені результати оцінювання авторегресійної моделі AR(3). Всі три коефіцієнти моделі значимі.

Таблиця 4.3. Результати оцінювання моделі AR(3)

Метод оцінювання: метод найменших квадратів

Скорегована часова вибірка даних: 1996:04 2005:01

Всього спостережень після корегування крайніх значень: 106

Модель: $I=C(1)+C(2)*I(-1)+C(3)*I(-2)+C(4)*I(-3)$

	Оцінки коеф-в	Станд. похиб.	t-статистика	Ймов.
C(1)	46.96556	9.330171	5.033729	0.0000
C(2)	0.613458	0.098563	6.223995	0.0000
C(3)	-0.157670	0.113613	-1.387787	0.1682
C(4)	0.077517	0.086493	0.896229	0.3722
R-квадрат	0.316655	Середнє залежності змінної		100.6604
Скорегований R- квадрат	0.296556	Станд. відхилення змінної		1.372428
Станд. похибка регресії	1.151076	Інформ. критерій Акайке		3.156277
Сума квадратів похибок	135.1476	Критерій Шварца		3.256785
		Стат. Дарбіна-Уотсона		1.992389

Таким чином, можна записати наступну модель:

$$\begin{aligned} isc(k) &= a_0 + a_1 isc(k-1) + a_2 isc(k-2) + a_3 isc(k-3) + \varepsilon(k) = \\ &= 46,96 + 0,61 isc(k-1) - 0,15 isc(k-2) + 0,08 isc(k-3) + e(k), \end{aligned}$$

Для цієї моделі спостерігається зменшення коефіцієнта детермінації від 0,415 до 0,317 (деяке погіршення); зменшення суми квадратів похибок від 141,99 до 135,148 і деяке покращення статистики Дарбіна-Уотсона: від 1,931 до 1,992 (похибки моделі можна вважати практично некорельзованими). Тобто отримані значення характеристик моделі:

$$R^2 = 0,317; \quad J = CKP = 135,148; \quad DW = 1,992.$$

Характеристики однокрокового прогнозу для даної моделі:

$$CeKP = 1,36; \quad CAP = 1,02; \quad CAPP = 1,01; \quad U = 0,0068,$$

тобто модель загалом придатна для прогнозування (за коефіцієнтом Тейла, який наближається до ідеального значення), а три інших показники свідчать про високу точність прогнозу. Необхідно зазначити, що показники якості прогнозу для моделей AP(1) і AP(3) є практично одинаковими.

Розглянемо характеристики моделі вищого порядку. В таблиці 4.4 наведені результати оцінювання моделі AP(7).

Таблиця 4.4. Результати оцінювання моделі AP(7) для ІСЦ

Метод оцінювання: метод найменших квадратів

Скорегована часова вибірка даних: 1996:08 2005:01

Всього використано спостережень після корегування крайніх значень:

102

Модель: $I = C(1) + C(2)*I(-1) + C(3)*I(-2) + C(4)*I(-3) + C(5)*I(-5) + C(6)*I(-6) + C(7)*I(-7)$

	Оцінки коеф-в	Станд. похибка	t-статистика	Ймов.
C(1)	33.13446	12.79240	2.590168	0.0111
C(2)	0.603211	0.101625	5.935625	0.0000
C(3)	-0.184907	0.117804	-1.569623	0.1198
C(4)	0.125878	0.101891	1.235415	0.2197
C(5)	-0.031530	0.102954	-0.306254	0.7601
C(6)	0.093246	0.117601	0.792904	0.4298
C(7)	0.064850	0.089527	0.724365	0.4706
R-квадрат	0.346348	Середнє залежної змінної		100.6667
Скорегований R- квадрат	0.305064	Станд. відхилен. зал. змінної		1.388306
Станд. похибка регресії	1.157330	Інформ. критерій Акайке		3.196268
Сума квадратів похибок	127.2443	Критерій Шварца		3.376413
		Стат. Дарбіна-Уотсона		1.811821

Порівнюючи моделі AP(3) і AP(7), можна сказати, що коефіцієнт детермінації збільшився (від 0,317 до 0,346; сума квадратів похибок моделі

зменшилась від 135,15 до 127,24, а статистика Дарбіна-Уотсона зменшилась від 1,992 до 1,811). Характеристики однокрокового прогнозу для AP(7):

$$CeKP = 1,36; \quad CAP = 1,012; \quad CAPP = 1,002; \quad U = 0,0067,$$

тобто на 0,008 зменшились середня абсолютна похибка і середня абсолютна похибка в процентах .

Результати оцінювання моделі 12-го порядку наведені в таблиці 4.5.

Таблиця 4.5. Результати оцінювання моделі AP(12) для ІСЦ

Метод оцінювання: метод найменших квадратів

Скорегована часова вибірка даних: 1997:01 2005:01

Всього використано спостережень після корегування крайніх значень: 97

Модель: $I = C(1) + C(2)*I(-1) + C(3)*I(-2) + C(4)*I(-3) + C(5)*I(-6) + C(6)*I(-7) + C(7)*I(-8) + C(8)*I(-9) + C(9)*I(-11) + C(10)*I(-12)$

	Оцінки коеф-в	Станд. похибка	t-статистика	Ймов.
C(1)	26.23923	14.28763	1.836500	0.0697
C(2)	0.666410	0.102198	6.520741	0.0000
C(3)	-0.247996	0.122488	-2.024654	0.0460
C(4)	0.179378	0.103248	1.737345	0.0859
C(5)	-0.038240	0.096715	-0.395392	0.6935
C(6)	0.016650	0.112406	0.148120	0.8826
C(7)	-0.037042	0.111025	-0.333639	0.7395
C(8)	0.030055	0.095744	0.313915	0.7543
C(9)	0.020445	0.093392	0.218911	0.8272
C(10)	0.149194	0.081977	1.819964	0.0722
R-квадрат	0.435002	Середнє залежної змінної		100.6082
Скорегований R- квадрат	0.376553	Станд. відхилен. зал. змінної		1.342857
Станд. похибка регресії	1.060301	Інформ. критерій Акайке		3.052366
Сума квадратів похибок	97.80880	Критерій Шварца		3.317800
		Статистика Дарбіна- Уотсона		1.940130

Модель AP(12) має кращі характеристики ніж попередні моделі:

$$R^2 = 0,435; \quad J = CKP = 97,80; \quad DW = 1,94.$$

Значно зменшилась сума квадратів похибок, підвищилося значення R^2 , а значення статистики Дарбіна-Уотсона майже таке ж, як для моделі AP(3).

Характеристики однокрокового прогнозу для цієї моделі:

$$CeKP = 1,337; \quad CAP = 1,02; \quad CAPP = 1,013; \quad U = 0,0066.$$

Таким чином, характеристики однокрокового прогнозу також найкращі для моделі AP(7). Можна зробити висновок, що процес формування індексу оптових цін може бути описаний моделлю авторегресії AP(7) з високим ступенем адекватності. Ця модель забезпечує також отримання кращого однокрокового прогнозу.

Модель авторегресії для відхилень ІСЦ від середнього

Якщо із вихідних (фактичних) значень ряду ІСЦ відняти середнє ($isc(k) - 100,66$), $k = 1, \dots, 109$, то отримаємо ряд відхилень від середнього. Автокореляційна функція залишається фактично незмінною. Результати оцінювання моделі AP(1) наведені в таблиці 4.6.

Таблиця 4.6. Результати оцінювання моделі AP(1) для відхилень ІСЦ

	Оцінки коef-в	Станд. похибка	t-статистика	Ймов.
C(1)	-0.007477	0.111831	-0.066857	0.9468
C(2)	0.569826	0.065703	8.672790	0.0000
R-квадрат	0.415067	Середнє залежної змінної		0.080741
Скорегований R- квадрат	0.409549	Станд. відхил. зал. змінної		1.506189
Станд. похибка регресії	1.157368	Інформ. критерій Акайке		3.148519
Сума квадратів похибок	141.9871	Критерій Шварца		3.198188
		Стат. Дарбіна-Уотсона		1.931805

Три вибрані статистичні характеристики адекватності цієї моделі:

$$R^2 = 0,415; \quad J = СКП = 141,99; \quad DW = 1,931.$$

Характеристики якості однокрокового прогнозу:

$$СeКП = 1,36; \quad САП = 1,02; \quad САПП = 100,40; \quad U = 0,662.$$

У порівнянні з моделлю AP(1) для повних значень (без віднімання середнього) середньоквадратична похибка і середня абсолютна похибка не змінились, але в 100 разів збільшилась САПП і коефіцієнт Тейла: від 0,0067 до 0,662. Таким чином, модель AP(1) для відхилень не придатна для прогнозування. Це можна пояснити тим, що відхилення мають різні знаки, що утруднює побудову моделі.

Авторегресія 13-го порядку для відхилень ІСЦ від середнього наведена в табл.. 4.7.

Таблиця 4.7. Результати оцінювання моделі AP(13) для відхилень ІСЦ.

Метод оцінювання: метод найменших квадратів

Скорегована часова вибірка даних: 1997:02 2005:01

Всього використано спостережень після корегування крайніх значень: 96

Модель: $IV = C(1) + C(2)*IV(-1) + C(3)*IV(-2) + C(4)*IV(-3) + C(5)*IV(-5) + C(6)*IV(-6)$

$$+ C(7)*IV(-7) + C(8)*IV(-8) + C(9)*IV(-9) + C(10)*IV(-11) + C(11)*IV(-12) + C(12)*IV(-13)$$

	Оцінки коef-в	Станд. похибка	t-статистика	Ймов.
C(1)	-0.045818	0.110205	-0.415756	0.6786
C(2)	0.706976	0.107583	6.571429	0.0000
C(3)	-0.261301	0.123875	-2.109395	0.0379
C(4)	0.194879	0.105046	1.855172	0.0671
C(5)	-0.110121	0.107924	-1.020352	0.3105
C(6)	0.043907	0.120182	0.365334	0.7158
C(7)	0.002351	0.115561	0.020347	0.9838
C(8)	-0.023657	0.114078	-0.207375	0.8362
C(9)	0.033030	0.096947	0.340705	0.7342
C(10)	0.002612	0.096331	0.027112	0.9784
C(11)	0.187169	0.109248	1.713242	0.0904
C(12)	-0.074932	0.085883	-0.872494	0.3854
R-квадрат	0.440922	Середнє зал. Змінної	-0.066250	
Скорегований R-квадрат	0.367709	Станд. відхил. зал. змінної	1.342254	
Станд. похибка регресії	1.067316	Інформ. крит. Акайке	3.084639	
Сума квадратів похибок	95.68969	Критерій Шварца	3.405183	

Логарифм правдоподібн.	-136.0627	Статистика Дарбіна- Уотсона	1.965193
---------------------------	-----------	--------------------------------	----------

Три вибрані статистичні характеристики адекватності цієї моделі:

$$R^2 = 0,440; \quad J = СКП = 95,69; \quad DW = 1,965.$$

Характеристики якості однокрокового прогнозу:

$$СeКП = 1,346; \quad САП = 1,02; \quad САПП = 109,09; \quad U = 0,848.$$

Спостерігається погіршення характеристик прогнозу, особливо погіршилось значення САПП і коефіцієнта Тейла: від 0,662 для моделі АР(1) до 0,848 для моделі АР(13). За цим параметром модель є непридатною для прогнозування. Коефіцієнти С(6) – С(10) є незначущими, а тому їх можна видалити з моделі.

Після вилучення цих коефіцієнтів із моделі та відповідних їм складових процесу отримаємо модель, параметри якої наведені в таблиці 4.8.

Таблиця 4.8. Результати оцінювання моделі АР(13) для відхилень

Метод оцінювання: метод найменших квадратів

Скорегована часова вибірка даних: 1997:02 2005:01

Всього використано спостережень після корегування крайніх значень: 96

Модель: IV=C(1)+C(2)*IV(-1)+C(3)*IV(-2)+C(4)*IV(-3)+C(5)*IV(-5)+C(6)*IV(-12)+C(7)*IV(-13)

	Оцінки коеф-в	Станд. похибка	t-статистика	Ймов.
C(1)	-0.046129	0.107257	-0.430077	0.6682
C(2)	0.700062	0.103640	6.754728	0.0000
C(3)	-0.256660	0.120078	-2.137440	0.0353
C(4)	0.195599	0.101703	1.923247	0.0576
C(5)	-0.081877	0.083177	-0.984368	0.3276
C(6)	0.190417	0.089745	2.121760	0.0366
C(7)	-0.067455	0.081058	-0.832180	0.4075
R-квадрат	0.438765	Середнє залеж. змінної	-0.066250	
Скорегований R- квадрат	0.400929	Станд. відхил. залеж. змінної	1.342254	
Станд. похибка регресії	1.038900	Інформ. критерій Акайке	2.984323	
Сума квадратів похибок	96.05881	Критерій Шварца	3.171307	
Логарифм правдоподібн.	-136.2475	Статистика Дарбіна- Уотсона	1.959533	

Вибрані статистичні характеристики адекватності цієї моделі:

$$R^2 = 0,439; \quad J = CKP = 96,05; \quad DW = 1,96.$$

Характеристики якості однокрокового прогнозу:

$$CeKP = 1,346; \quad CAP = 1,017; \quad CAPP = 108,2; \quad U = 0,854.$$

Таким чином, після видалення з моделі незначимих коефіцієнтів її характеристики залишились практично незмінними. В цілому можна зробити висновок, що побудовані авторегресійні моделі для відхилень ІСЦ від середнього не придатні для прогнозування. Це свідчить про те, що побудувати модель для процесу, який має різні знаки вимірів в різні моменти часу (тобто розвиток відбувається в двох квадрантах), складніше ніж для процесу, який змінюється в межах одного квадранту.

Авторегресія з ковзним середнім для ІСЦ

Розглянемо можливість описання ІСЦ за допомогою моделі АРКС. Характеристики моделі АРКС(1,1) наведені в таблиці 4.9.

Таблиця 4.9. Результати оцінювання моделі АРКС(1, 1) для ІСЦ

Метод оцінювання: метод найменших квадратів
 Скорегована часова вибірка даних: 1996:02 2005:01
 Всього використано спостережень після корегування крайніх значень:
 108

Змінна	Оцінки коеф-в	Станд. похибка	t-статистика	Ймов.
C	100.6529	0.263645	381.7742	0.0000
AR(1)	0.546826	0.093191	5.867793	0.0000
KC(1)	0.060696	0.133153	0.455838	0.6494
R-квадрат	0.415932	Середнє залеж. змінної	100.7407	
Скорегований R-квадрат	0.404807	Станд. відхилен. залеж. змінної	1.506189	
Станд. похибка регресії	1.162006	Інформ. критерій Акайке	3.165558	
Сума квадратів похибок	141.7772	Критерій Шварца	3.240062	
Логарифм	-167.9401	F-статистика	37.38676	

правдоподібн.				
Статист. Дарбіна-	1.996202	Ймовірність (F-стат.)	0.000000	
Уотсона				
Інвертовані AP корені	.55			
Інвертовані KC корені	-.06			

Вибрані статистичні характеристики адекватності цієї моделі:

$$R^2 = 0,416; \quad J = CKP = 141,78; \quad DW = 1,996.$$

Характеристики якості однокрокового прогнозу:

$$CeKP = 1,362; \quad CAP = 1,016; \quad CAPP = 1,005; \quad U = 0,0067.$$

Моделі AP(1) і APC(1,1) мають практично одинакові характеристики адекватності та якості однокрокового прогнозу, а тому перевагу (при виборі з цих двох моделей) можна надати моделі AP(1), яка є простішою. Нижче наведена порівняльна таблиця для всіх побудованих моделей.

Врахування впливу на ICЦ агрегату M3

Коефіцієнт кореляції між ICЦ та агрегатом M3 від'ємний: -0,249, тобто формальний взаємозв'язок між цими змінними незначний, але цікаво розглянути вплив M3 на ICЦ за допомогою моделі. Врахування регресора може покращити деякі характеристики моделі, а також врахувати причинний зв'язок між вибраними змінними. Характеристики змішаної моделі: авторегресія AP(1) + парна регресія наведені в таблиці 4.10.

Таблиця 4.10. Результати оцінювання моделі ICЦ: AP(1) + регресор M3

Метод оцінювання: метод найменших квадратів

Скорегована часова вибірка даних: 1996:02 2005:01

Всього використано спостережень після корегування крайніх значень:

108

Модель: $I = C(1) + C(2)*I(-1) + C(3)*M$

	Оцінки коef-в	Станд. похибка	t-статистика	Ймов.
C(1)	44.97869	6.920057	6.499757	0.0000
C(2)	0.554264	0.068261	8.119764	0.0000
C(3)	-2.82E-06	3.30E-06	-0.854463	0.3948
R-квадрат	0.419106	Середнє залеж. змінної		100.7407
Скорегований R-	0.408041	Станд. відхилен. залеж.		1.506189

квадрат Станд. похибка регресії	1.158844	змінної Інформ. критерій Акайке	3.160108
Сума квадратів похибок	141.0066	Критерій Шварца	3.234612
Логарифм правдоподібн.	-167.6458	Стат. Дарбіна-Уотсона	1.919823

Спостерігається незначне покращення характеристик моделі та якості прогнозу у порівнянні з АР(1), але коефіцієнт при МЗ дуже малий (-2,82E-06). Однак, формально він є значимим. Таким чином, на розглянутому часовому інтервалі вплив агрегату МЗ на індекс споживчих цін незначний і ним можна знехтувати.

Характеристики прогнозу:

$$CeK\bar{P} = 1,340; \quad CA\bar{P} = 1,004; \quad CA\bar{PP} = 0,994; \quad U = 0,0066.$$

Можна припустити, що в обороті було недостатньо грошової маси для того, щоб її вплив на ІСЦ був істотним. З іншого боку, недостатній об'єм грошової маси в національній валюті компенсувався (і продовжує компенсуватись) "твердою" іноземною валютою, зокрема доларами США і, в деякій мірі, євро. Таким чином, слабка українська економіка позитивно впливає на курс долара США завдяки фактичному введенню його в частковий оборот на відносно великій території України. Встановити фактичне співвідношення між об'ємами національної та іноземної валют в обороті можливо, але для цього необхідно отримати додаткові статистичні дані і виконати спеціальне дослідження. Зокрема, необхідно мати дані щодо об'ємів реалізації торговельних операцій у валюті підприємствами всіх форм власності. Очевидно, що отримані такі дані для тіньового обороту непросто і пов'язані з деяким ризиком.

Визначення впливу ВВП на ІСЦ

Як було показано на початку цього параграфа, коефіцієнт кореляції між ІСЦ та ВВП складає -0,304, тобто формально це невелике значення. Також логічно припустити, що зростання ВВП має приводити до зменшення ІСЦ (про це свідчить також знак коефіцієнта кореляції між цими змінними).

Характеристики змішаної моделі: авторегресія АР(1) + парна регресія наведені в таблиці 4.11.

Таблиця 4.11. Результати оцінювання змішаної регресії для ІСЦ і ВВП

Метод оцінювання: метод найменших квадратів
 Скорегована часова вибірка даних: 1996:02 2005:01

Всього використано спостережень після корегування крайніх значень:
 108

Модель: $I = C(1) + C(2)*I(-1) + C(3)*V$

	Оцінки коef-в	Станд. похибка	t-статистика	Ймов.
C(1)	45.53125	7.156626	6.362112	0.0000
C(2)	0.549611	0.070124	7.837741	0.0000
C(3)	-1.37E-05	1.64E-05	-0.833627	0.4064
R-квадрат	0.418913	Середнє залеж. змінної	100.7407	
Скорегований R-квадрат	0.407844	Станд. відхилення залеж. змінної	1.506189	
Станд. похибка регресії	1.159037	Інформ. критерій Акайке	3.160441	
Сума квадратів похибок	141.0536	Критерій Шварца	3.234945	
Логарифм правдоподібності	-167.6638	Стат. Дарбіна-Уотсона	1.916289	

Отримано рівняння:

$$isc(k) = 45,53 + 0,55 \cdot isc(k-1) - (1,37E-05) \cdot vvp(k) + e(k)$$

з характеристиками:

$$R^2 = 0,419; \quad J = CKP = 141,05; \quad DW = 1,92.$$

Характеристики якості однокрокового прогнозу:

$$CeKP = 1,335; \quad CAP = 1,004; \quad CAPP = 0,993; \quad U = 0,0066.$$

Коефіцієнт при змінній VVP невеликий і знаходиться на межі статистичної значимості, тобто вплив VVP на ICP незначний. Причиною такого незначного впливу може бути некоректний розподіл VVP , який призводить до того, що більшість населення проживає на межі або навіть нижче межі бідності. Одночасно інша (менша) частина привласнює більшість благ і користується ними, але, очевидно, що це користування не приводить до позитивного впливу на ICP . Тому і маємо, що VVP зростає начебто значними темпами, а добробут населення сильно від нього відстає. Таким чином, можна зробити висновок, що на часовому періоді, який розглядається в даному прикладі, відбувався несправедливий розподіл суспільних благ, що проявилось, в даному випадку, в незначному впливі VVP на споживчі ціни для більшості населення України. Побудована

математична модель є формальним (об'єктивним) підтвердженням відомого факту нерівномірного розподілу матеріальних благ, вироблених в Україні.

В таблиці 4.12 зведені характеристики математичних моделей, побудованих для індексу споживчих цін, і характеристики однокрокових прогнозів, обчислені на основі цих моделей. Ця таблиця дає можливість оперативно порівняти результати моделювання та прогнозування, отримані за допомогою методики Бокса-Дженкінса, а також встановити можливість практичного використання результатів.

Таблиця 4.12. Результати моделювання і однокрокового прогнозування індексу оптових цін

Тип моделі	Характеристики моделі			Характеристики однокрокового прогнозу			
	R^2	$\sum e^2(k)$	DW	CeKP	САП	САПП	Коеф-т Тейла
AP(1)	0,415	141,99	1,931	1,360	1,020	1,008	0,0067
AP(3)	0,317	135,148	1,992	1,360	1,020	1,011	0,0068
AP(7)	0,346	127,244	1,811	1,360	1,012	1,002	0,0067
AP(12)	0,435	97,80	1,941	1,337	1,020	1,013	0,0066
APKC(1,1)	0,416	141,78	1,996	1,362	1,016	1,005	0,0067
AP(1)+МЗ	0,419	141,007	1,919	1,340	1,004	0,994	0,0066
AP(1)+ВВП	0,419	141,054	1,916	1,335	1,004	0,993	0,0066
Моделі для відхилень ІСЦ від середнього							
AP(1)	0,415	141,99	1,931	1,360	1,020	100,40	0,662
AP(13)	0,440	95,69	1,965	1,346	1,020	109,09	0,848

Результати моделювання, наведені в таблиці 4.12, свідчать про те, що практично всі моделі, побудовані для ІСЦ (перші сім моделей), є придатними для прогнозування, оскільки коефіцієнт Тейла вимірюється тисячними долями. Характеристики однокрокового прогнозу для моделей різного порядку відрізняються несуттєво. Найкращі характеристики щодо прогнозування має модель AP(1)+ВВП. Для моделі AP(12) отримано найменше значення суми квадратів похибок моделі, але характеристики прогнозів, отриманих за цією моделлю, не кращі від інших.

Дві моделі, які побудовані для відхилень ІСЦ від середнього, мало придатні для прогнозування. Про це свідчать високі значення коефіцієнта Тейла та середньої абсолютної похибки в процентах. Цей факт можна пояснити тим, що описати за допомогою АРКС процес, динаміка якого спостерігається у двох квадрантах, трудніше, ніж процес, який спостерігається в одному квадранті. Тому для прогнозування відхилень від середнього необхідно знайти іншу структуру математичної моделі, придатну для описання різнознакових величин часового ряду.

Запитання і вправи до четвертого розділу

1. Назвіть етапи побудови математичних моделей за методикою Бокса-Дженкінса. Що забезпечує коректне використання цієї методики на практиці?
2. Яка мета аналізу функціонування процесу? Які елементи моделі ми встановлюємо за допомогою аналізу процесу?
3. Яка мета попередньої обробки даних? Назвіть основні операції, які виконують в процесі попередньої обробки даних? До чого призводить визначення значень змінних у великому діапазоні?
4. Які два основних типи нелінійностей зустрічаються в аналізі часових рядів? Поясніть на прикладах.
5. Яким чином можна встановити наявність нелінійностей у процесі?
6. Який метод дає можливість автоматизувати процес визначення та врахування нелінійностей процесу?
7. Яку інформацію можна отримати на основі візуального аналізу даних? Як можна нею скористатись?
8. Яким чином можна знайти оцінку порядку авторегресійної частини моделі?
9. В чому полягає відмінність між автокореляційною та частковою автокореляційною функціями процесу? Чи існує необхідність розрахунку обох функцій в процесі аналізу даних?
10. На чому ґрунтуються відбір незалежних змінних (регресорів, екзогенних змінних) для включення в праву частину математичної моделі?
11. Назвіть три умови коректного застосування методу найменших квадратів до оцінювання параметрів математичної моделі?
12. Для чого призначена статистика Льюнга-Бокса і як вона обчислюється? Скористайтеся наявним пакетом програм для статистичного аналізу даних для обчислення цього статистичного параметру і поясніть його значущість (чи незначущість) на прикладі.
13. В чому полягає принципова різниця між методом найменших квадратів (МНК), призначеним для оцінювання лінійних моделей, та МНК для оцінювання моделей, нелінійних відносно параметрів? Які критерії якості мінімізують ці методи?
14. Виведіть формулу МНК для поліноміальної моделі?
15. Сформуйте матрицю вимірів для математичної моделі вигляду:

$$y(k) = a_0 + a_1 y(k-1) + b_1 x_1(k) + b_2 x_2(k) + \varepsilon(k).$$
16. Про що свідчить корельованість похибок моделі між собою? За допомогою якого статистичного параметра якості моделі можна визначити ступінь корельованості похибок? Яке значення приймає ця статистика в ідеальному випадку?

17. Що означає значущість оцінки параметра (коефіцієнта) моделі в статистичному смыслі? За допомогою якої статистики можна встановити значущість оцінки параметра моделі?
18. Що означають помилки першого і другого роду при перевірці статистичних гіпотез?
19. Поясніть фізичну суть коефіцієнта (множинної) детермінації, яке його ідеальне значення? Для чого корегують R^2 із врахуванням числа ступенів свободи?
20. В чому полягає різниця між критеріями Акайке та Байєса-Шварца? Чи можна обмежитись використанням тільки одного з цих критеріїв?
21. Яка мета розрахунку статистики Фішера? Наведіть послідовність розрахунку та аналізу цієї статистики.
22. Сформулюйте правила перевірки гіпотез для t – статистики Стьюдента і F – статистики Фішера. Яким чином можна визначити критичні (порогові) значення цих статистик?
23. Поясніть фізичну суть умови достатнього збудження процесу?
24. Яким чином досягають достатнього збудження процесів (об'єктів) на практиці? Які сигнали використовують для досягнення цієї мети? Які властивості цих сигналів використовують в даному випадку?
25. Нарисуйте наближений графік перетворення Фур'є для білого шуму.
26. Які інші ортогональні перетворення даних ви знаєте крім перетворення Фур'є? Чому на практиці використовують декілька різних ортогональних перетворень, а не одне?

Розділ 5

МЕТОДИКА ПРОГНОЗУВАННЯ ДИНАМІКИ ПРОЦЕСІВ ЗА ДОПОМОГОЮ РІЗНИЦЕВИХ РІВНЯНЬ

5.1 Для чого необхідні прогнози?

Прогнозування подій в процесі приватної та ділової активності – це суттєва складова нашого повсякденного життя, яка стосується кожного. Навряд чи знайдеться такий напрям людської діяльності, де не потрібно робити прогнози. Для того щоб продемонструвати важливість коректного розв'язання задачі прогнозування розглянемо ієрархію процесу прийняття рішень на рівні виробничої, торговельної або іншої компанії. В управлінському процесі, що стосується діяльності цих компаній, можна виділити такі *рівні ієрархії* планування:

1. *Стратегічне бізнес-планування.* На цьому рівні визначаються стратегічні цілі функціонування організації (фірми, підприємства) на довгому часовому інтервалі: від 3-х до 20-и років. План складається по *місяцях, кварталах* або *роках*. Основними цілями, як правило, є наступні: отримання визначеного рівня прибутку; досягнення необхідного рівня якості продукції; визначення своєї участі (частини) на ринку конкретних товарів; задоволення вимог покупців; досягнення визначеного рівня автоматизації виробництва; досягнення необхідного рівня відносин між роботодавцем і найманими робітниками.
2. *Середньострокове планування управлінської діяльності.* На даному рівні визначаються управлінські дії, необхідні для досягнення стратегічних цілей на часовому інтервалі від 3-х місяців до 3-х років. План складається по *тижнях* або по *місяцях*. Планування стосується необхідних фінансів; маркетингу; об'ємів продаж, перевезення (розповсюдження) продукції; числа, місцезнаходження та характеру необхідного устаткування; агрегування (концентрації) виробництва – яка конкретна продукція буде вироблятись на

конкретних відділеннях підприємства; формування основних календарних планів.

3. *Оперативне планування та управління*. Головна мета цього рівня управління полягає в тому, щоб ефективно розподілити ресурси для досягнення наведених вище головних цілей. Оперативне планування та управління реалізується на часових інтервалах від 1 місяця до 2-х років; план складається по *тижнях* або по *місяцях*. Воно стосується управління та контролю розподілу матеріальних ресурсів, виробництва та розподілу продукції, необхідних виробничих потужностей (приміщення, устаткування, робоча сила).
4. *Обробка транзакцій (щоденне управління)*. На цьому рівні реалізуються управлінські операції на часовому інтервалі від 1-го дня до 1-го місяця. При цьому контролюються такі операції: закупівля необхідних матеріалів; формування замовлень; виставлення та прийняття рахунків; оперативний контроль якості продукції; управління транспортним господарством підприємства.

Докладний аналіз кожного рівня прийняття рішень на підприємстві показує, що управлінський персонал не може ефективно вирішувати задачі планування та контролю без належного високоякісного *прогнозування* показників розвитку підприємства та інших процесів, пов'язаних з діловою активністю. При цьому горизонт прогнозу (проміжок часу, на який робиться прогноз) залежить від рівня ієархії і може бути дуже довгим (до 20-и років) або коротким (один день). Таким чином, ієархія процесу прийняття рішень потребує створення та використання ієархічної системи прогнозування.

Необхідно зазначити, що рух вниз по управлінській ієархії супроводжується суттєвими змінами характеру інформації та її докладності. Так, на вершині піраміди необхідно приймати *неструктуровані* та *погано структуровані* рішення при неповній інформації (або її відсутності). Чим нижчим є рівень управління, тим більше програмованими (структуркованими, упорядкованими, прозорими) є рішення, для яких існує інформація у все повнішому об'ємі.

При наявності на підприємстві комп'ютеризованої системи підтримки прийняття рішень з'являється можливість автоматизації процесів управління та контролю. Наприклад, автоматизуються операції обліку на складах, контролю якості, аналіз номенклатури обладнання, формування технологічних карт для реалізації технологічних процесів, накопичення та підтримки необхідних запасів сировини і т. ін.. В цілому, рухаючись вниз по ієархії прийняття рішень можна сформулювати такі висновки:

- рішення та прогнози стають менш стратегічними;
- зменшується участь в прийнятті рішень управлінського персоналу верхнього рівня;

- скорочується горизонт прогнозування від років до місяців, тижнів та днів;
- рішення та прогнози стають більш програмованими (стандартизованими, структурованими, чітко визначеними);
- значно зростає докладність описання ситуацій та задач, щодо яких приймаються рішення;
- зростає ступінь автоматизації процесу прийняття рішень;
- для прийняття рішень, а особливо тих, що повторюються, доцільно застосовувати відповідні комп’ютерні системи підтримки прийняття рішень при прогнозуванні та плануванні діяльності.

Із сказаного випливає, що прогнозуюча система також повинна бути побудована за ієрархічною архітектурою (структурою). Такий підхід дає можливість структуризувати (упорядкувати) задачі прогнозування, створити множину необхідних для прогнозування спеціалізованих моделей, ситуацій та методів прийняття рішень на кожному рівні, а також наблизити комп’ютеризований процес прийняття рішень до того, що є звичним для нас.

Крім бізнес-процесів досить часто ставиться задача оцінювання прогнозу при прийнятті приватних рішень. Наприклад, необхідно спрогнозувати сімейний бюджет на 4-5 років із врахуванням оплати навчання у навчальному закладі та виплат за житловими кредитами, або спрогнозувати витрати часу на виконання різних робіт в залежності від їх складності.

5.2 Які складові процесу можна прогнозувати

В загальному випадку прогноз може бути представлений одним (точковим) значенням змінної, інтервалом, в який попадає випадкова змінна, а також ймовірністю прийняття змінною (чи подією) деякого значення у вибраному інтервалі. Якщо для опису процесу застосовують лінгвістичні змінні, то прогнозом буде нечітке значення, але його також можна перетворити в чітке число.

Можна по-різному ставити задачу прогнозування в залежності від рівня прийняття рішення та конкретної поставленої задачі управління чи контролю. Прогнозування може стосуватись наступних складових процесу:

- детермінованого тренду, як індикатора довгострокових змін процесу;
- випадкового (нерегулярного) тренду, як показника коротко- та середньострокових змін;
- короткострокових змін, тобто, прогнозування коливань (відхилень), що накладаються на тренд;
- сезонних ефектів;

- приростів (швидкості) зміни процесу, які визначаються першими різницями;
- дисперсії або стандартного відхилення, як міри розсіювання процесу (наприклад, волатильність, яку часто використовують за міру ризику у інвестуванні або міру якості на виробництві);
- якісних змінних (за допомогою нечітких множин, мереж Байєса і т. ін.);
- комбінацій вказаних елементів процесів.

Відповідно до того, які складові процесу необхідно прогнозувати, ставиться задача побудови математичної, ймовірнісної (статистичної) або логічної моделі, що має меті забезпечити високу якість прогнозу на заданий часовий горизонт. Розглянемо деякі можливості створення математичного опису складових процесів різної природи.

Детермінований тренд

Якщо описати детермінований тренд за допомогою полінома від часу довільного порядку

$$y(k) = a_0 + a_1 \cdot k + a_2 \cdot k^2 + \dots + a_p \cdot k^p + \varepsilon(k), \quad E[\varepsilon(k)] = 0,$$

то визначення прогнозу тренду зводиться до підстановки в це рівняння бажаного значення часу k і застосування безумовного математичного сподівання. Прогнозування значень тренду вважається довгостроковим прогнозом, оскільки детермінований тренд вказує на довгострокові зміни процесів. Обмеження на випадковий процес $E[\varepsilon(k)] = 0$ необхідне для коректного застосування методів оцінювання моделей, а також для виконання подальшого аналізу якості оцінок прогнозів.

Стохастичний тренд

Для опису і прогнозування стохастичного тренду можна скористатись, наприклад, відомим рівнянням випадкового кроку із перетином (константою):

$$y(k) = a_0 + y(k-1) + \varepsilon(k),$$

розв'язок якого має вигляд:

$$y(k) = y_0 + k a_0 + \sum_{i=1}^k \varepsilon(i).$$

Сума $\sum_{i=1}^k \varepsilon(i)$ в правій частині останнього рівняння описує випадкову складову тренду. Цю складову називають ще *нерегулярною*.

Прогнозування коливань, що накладаються на тренд

Коливання, що накладаються на тренд, або короткострокові зміни можна описати рівняннями авторегресії з ковзним середнім

$$y(k) = \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j \varepsilon(k-j) + \varepsilon(k).$$

Техніку обчислення такого прогнозу за допомогою умовного та безумовного математичного сподівання розглянемо нижче в цьому розділі.

При переході до перших та різниць вищих порядків з процесу вилучається тренд відповідного порядку. Наприклад, якщо процес містить лінійний тренд, то перші різниці вилучають його і після переходу до різниць ми маємо справу з коливаннями, що накладаються на тренд.

Можливість прогнозування *сезонних ефектів* досягається за рахунок введення в модель процесу відповідних значень основної змінної із затримками (лагами), що відповідають періодичності ефекту. Як буде показано нижче, сезонний ефект може бути врахований як за допомогою основної змінної, так і за допомогою процесу ковзного середнього.

Прогнозування дисперсії

Якщо дисперсія процесу змінюється в часі, то для її описання можна вибрати рівняння для формування відповідної функції прогнозування та обчислення в подальшому оцінки прогнозованого значення. Процеси зі змінною дисперсією отримали назву гетероскедастичних, вони будуть розглянуті нижче, в іншому розділі. Методика побудови моделей гетероскедастичних процесів передбачає наступні кроки:

- математичний опис самої змінної рівнянням авторегресії невисокого порядку (наприклад, першого);
- математичний опис умовної дисперсії, як динамічної змінної, за допомогою рівняння прийнятної (за якістю прогнозу) структури.

5.3 Умовні та безумовні статистичні характеристики

При виконанні статистичного аналізу випадкових процесів використовують *умовні* та *безумовні статистичні характеристики*. Зокрема, для знаходження короткострокових та довгострокових прогнозів

розвитку процесів застосовують умовне, E_k , та безумовне, E , математичне сподівання, відповідно. Різниця у визначенні цих характеристик полягає у наступному.

Безумовні статистичні характеристики – це константи, які розглядають і оцінюють на довільних часових інтервалах, не накладаючи умов на змінні, функції та об'єми інформації, необхідної для визначення цих характеристик. Тобто інформація для обчислення дисперсії вважається відомою на всьому інтервалі, що розглядається в процесі аналізу.

Так, безумовне математичне сподівання використовують для знаходження давгострокових прогнозів або умов економічної рівноваги. Наприклад, безумовні вибіркове середнє і дисперсія обчислюються за відомими формулами:

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N y(k), \quad Var(x) = \frac{1}{N-1} \sum_{k=1}^N [x(k) - \bar{x}]^2,$$

а безумовне математичне сподівання позначається символом E . У виразі для дисперсії немає необхідності зменшувати значення N на одиницю, якщо потужність ряду досить велика, наприклад, декілька десятків.

Умовні статистичні характеристики в аналізі динаміки процесів, представлених часовими рядами, – це функції часу, які визначають на кожний конкретний момент часу k . При цьому для їх обчислення необхідно, щоб на вибраний (заданий) момент була в наявності інформація щодо значень змінних та функцій, яка необхідна для виконання відповідних обчислень.

Умовне математичне сподівання застосовують для визначення короткострокових та середньострокових прогнозів. Умовну дисперсію процесу і стандартне відхилення також часто використовують як міру ризику. Наприклад, в аналізі фінансових процесів, при формуванні правил торгівлі на біржі, аналізі банківських та економічних ризиків. В технічних системах дисперсія також відіграє значну при визначені ступеня відхилення вузла (наприклад, підшипника, корпусного елемента, крила) від нормального (заданого) стану. Тому вміння правильно аналітично описати дисперсію дає можливість з високою точністю описати і спрогнозувати значення відхилень від норми.

Так, умовне математичне сподівання стохастичного процесу AP(1) визначається за виразом:

$$\begin{aligned} E_k[y(k+1) | y(k), y(k-1), \dots, y(0), \varepsilon(k), \varepsilon(k-1), \dots, \varepsilon(0)] &= \\ &= E_k[a_0 + a_1 y(k) + \varepsilon(k+1)] = a_0 + a_1 y(k) \end{aligned}$$

за умови, що $E_k[\varepsilon(k+l)] = 0$, $l \geq 1$. Умовне вибіркове середнє та умовну вибіркову дисперсію можна наблизено обчислити за виразами:

$$\bar{x}(k) = \frac{1}{k} \sum_{i=1}^k x(i); \quad \text{var}_k(x) = \frac{1}{k-1} \sum_{i=1}^k [x(i) - \bar{x}(k)]^2, \quad k = 2, \dots, N.$$

В результаті застосування цих виразів отримаємо ряд значень умовного середнього та умовної дисперсії, тобто ще дві характеристики процесу, якими можна скористатись при побудові математичних і статистичних моделей.

Наблизені значення умовних вибіркових статистичних характеристик можна визначити, також, по аналогії з обчисленням ковзного середнього. При такому підході необхідно вибрати ширину ковзного вікна і обчислити значення статистичного параметра рухаючись крок за кроком від початку до кінця часового ряду. Наприклад, якщо вибрати вікно шириною p ять значень ряду, то умовна дисперсія буде обчислюватись за виразом:

$$\text{var}_k(x) = \frac{1}{4} \sum_{i=k-2}^{k+2} [y(i) - \bar{y}]^2, \quad k = 3, \dots, N-2,$$

або в загальному вигляді:

$$\text{var}_k(x) = \frac{1}{w-1} \sum_{i=k-(w-1)/2}^{k+(w-1)/2} [y(i) - \bar{y}]^2, \quad k = \frac{d-1}{2}, \dots, N - \frac{d-1}{2},$$

де w – ширина ковзного вікна. Ширина вікна залежить від того, наскільки швидко змінюється дисперсія. Якщо вона має високу динаміку, то для розміру вікна вибирають значення 5-9. Незважаючи на наблизеність таких розрахунків, практика моделювання свідчить, що обчислені значення виявляються, як правило, досить корисними при побудові моделей.

У складніших випадках обчислюють функцію умовної дисперсії, яка точніше відображає характер її зміни у часі. Наприклад, можна розглянути наступну модель процесу:

$$y(k) = \sqrt{f[x(k)]} \varepsilon(k), \quad (5.3.1)$$

$$\varepsilon(k) = \beta \varepsilon(k-1) + v(k), \quad (5.3.2)$$

де $y(k) \in \Re$, $k = 1, 2, \dots, N$; $\{v(k)\}$ – множина незалежних однаково розподілених (НОР) величин, що мають нормальній розподіл з параметрами $N(0,1)$; $\beta \in \Theta = (-1; 1)$; $f[x(k)] \in C^p[0,1]$. Припущення про те, що $v(k)$ – гаусів

процес зроблено для зручності викладок. Змінна $x(k) \in [0,1]$ є упорядкованою за значеннями, тобто:

$$x(1) \leq x(2) \leq \dots \leq x(N),$$

де $x(k) = \frac{k}{N}$, $k = 1, 2, \dots, N$. Функцію $f[x(k)]$ називають функцією дисперсії, хоча вона не повністю описує дисперсійно-коваріаційну структуру процесу (5.3.1)-(5.3.2). Припустимо, що $f[x(k)]$ має p неперервних похідних. Безумовну дисперсія процесу $y(k)$ можна визначити наступним чином:

$$\begin{aligned} \text{var}[y(k)] &= f[x(k)] E[\varepsilon^2(k)] = f[x(k)] E\{\beta \varepsilon(k-1) + v(k)\}^2 = \\ &= \{\beta^2 E[\varepsilon^2(k-1)] + E[v^2(k)]\} f[x(k)] = \\ &= \beta^2 f[x(k)] E[\varepsilon^2(k-1)] + f[x(k)] = \\ &= \beta^2 \text{var}[y(k)] + f[x(k)]. \end{aligned}$$

Звідси отримаємо:

$$\text{var}[y(k)] = \frac{f[x(k)]}{1 - \beta^2}.$$

Таким чином, процес (5.3.1)-(5.3.2) є умовно і безумовно гетероскедастичним. Задача полягає у тому, щоб оцінити функцію $f[x(k)]$. Така задача буде розглянута нижче.

5.4 Прогнозування без знаходження розв'язку рівнянь

Обчислення оцінки прогнозу без розв'язку рівнянь

Структура різницевого рівняння така, що воно дозволяє виконувати прогнозування на один крок (один період дискретизації вимірів) без додаткових перетворень. Тобто в праву частину необхідно підставити минулі значення змінних і обчислити оцінку прогнозу головної змінної в лівій частині. Але для того щоб знайти оцінку прогнозу на більше число кроків, необхідно застосувати деякі попередні перетворення різницевих рівнянь. Розглянемо деякі можливі підходи до формування функцій прогнозування та обчислення оцінок прогнозованих значень.

Як приклад, розглянемо рівняння AP(1):

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k), \quad E[\varepsilon(k)] = 0. \quad (5.4.1)$$

Збільшимо незалежну змінну час на одну одиницю і запишемо рівняння знову:

$$y(k+1) = a_0 + a_1 y(k) + \varepsilon(k+1). \quad (5.4.2)$$

Якщо коефіцієнти a_0, a_1 відомі, то можна знайти умовне математичне сподівання на основі відомої інформації до моменту k включно:

$$\begin{aligned} E_k[y(k+1)] &= E_k[y(k+1) | y(k), y(k-1), \dots, \varepsilon(k), \varepsilon(k-1), \dots] = \\ &= a_0 + a_1 E_k[y(k)] = a_0 + a_1 y(k), \end{aligned} \quad (5.4.3)$$

оскільки $y(k)$ в момент k є відомою константою.

По аналогії запишемо рівняння (5.4.1) для моменту $k+2$:

$$y(k+2) = a_0 + a_1 y(k+1) + \varepsilon(k+2) \quad (5.4.4)$$

і знайдемо умовне математичне сподівання

$$\begin{aligned} E_k[y(k+2)] &= a_0 + a_1 E_k[y(k+1)] = a_0 + a_1 E_k[a_0 + a_1 y(k)] = \\ &= a_0 + a_0 a_1 + a_1^2 y(k). \end{aligned}$$

Для наступного моменту часу маємо

$$E_k[y(k+3)] = a_0 + a_0 a_1 + a_0 a_1^2 + a_1^3 y(k).$$

Таким чином, для загального випадку прогнозування на s кроків можна записати

$$E_s[y(k+s)] = a_0 \left(\sum_{i=0}^{s-1} a_1^i \right) + a_1^s y(k) = a_0 \sum_{i=0}^{s-1} a_1^i + a_1^s y(k). \quad (5.4.5)$$

Отримане рівняння називають функцією прогнозування на довільне число кроків. Прогноз представляє собою збіжний процес, якщо $|a_1| < 1$, тобто

$$\lim_{s \rightarrow \infty} E_k[y(k+s)] = \frac{a_0}{1 - a_1}, \quad |a_1| < 1, \quad (5.4.6)$$

де a_1 – знаменник геометричної прогресії. Вираз (5.4.6) свідчить про те, що для будь-якого стаціонарного процесу АР чи АРКС оцінка умовного прогнозу асимптотично ($s \rightarrow \infty$) збігається до безумовного середнього.

Похибка прогнозу на довільне число кроків

Знайдемо похибку прогнозування при умові, що $E[\varepsilon(k)] = 0$:

$$f_k(s) = y(k+s) - E_k[y(k+s)]. \quad (5.4.7)$$

Похибка прогнозу на один крок:

$$f_k(1) = y(k+1) - E_k[y(k+1)] = a_0 + a_1 y(k) + \varepsilon(k+1) - a_0 - a_1 y(k) = \varepsilon(k+1).$$

Похибка прогнозу на два кроки

$$\begin{aligned} f_k(2) &= y(k+2) - E_k[y(k+2)] = \\ &= a_0 + a_1[a_0 + a_1 y(k) + \varepsilon(k+1)] + \varepsilon(k+2) - E_k[y(k+2)] = \\ &= a_0 + a_0 a_1 + a_1^2 y(k) + a_1 \varepsilon(k+1) + \varepsilon(k+2) - a_0 - a_0 a_1 - a_1^2 y(k) = \\ &= \varepsilon(k+2) + a_1 \varepsilon(k+1). \end{aligned}$$

Таким чином, можемо записати вираз для похибки для довільного числа кроків прогнозування як

$$f_k(s) = \varepsilon(k+s) + a_1 \varepsilon(k+s-1) + a_1^2 \varepsilon(k+s-2) + \dots + a_1^{s-1} \varepsilon(k+1). \quad (5.4.8)$$

Враховуючи те, що $E[f_k(s)] = 0$, оцінка прогнозу, яка обчислюється за виразом (5.4.5), є незміщеною. Дисперсія похибки прогнозування

$$Var[f_k(s)] = \sigma^2 [1 + a_1^2 + a_1^4 + a_1^6 + \dots + a_1^{2(s-1)}],$$

тобто дисперсія є функцією s . Асимптотичне значення дисперсії похибки прогнозу для стаціонарного процесу

$$\lim_{s \rightarrow \infty} Var[f_k(s)] = \frac{\sigma^2}{1 - a_1^2}, \quad (5.4.9)$$

де a_1^2 – знаменник геометричної прогресії.

Узагальнення функції прогнозування (на процес APKC(p,q))

Знайдемо спочатку функцію прогнозування для процесу APKC(2,1), який описується наступним рівнянням:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + \varepsilon(k) + \beta_1 \varepsilon(k-1). \quad (5.4.10)$$

Для наступного моменту часу можна записати

$$y(k+1) = a_0 + a_1 y(k) + a_2 y(k-1) + \varepsilon(k+1) + \beta_1 \varepsilon(k).$$

Умовне математичне сподівання для $y(k+1)$ має вигляд

$$E_k[y(k+1)] = a_0 + a_1 y(k) + a_2 y(k-1) + \beta_1 \varepsilon(k),$$

де $\varepsilon(k)$ розглядається як відома величина на момент k включно. При цьому

$$E_k[\varepsilon(k+j)] = 0, \quad \forall j > 0.$$

Для моменту часу $k+2$ маємо

$$y(k+2) = a_0 + a_1 y(k+1) + a_2 y(k) + \varepsilon(k+2) + \beta_1 \varepsilon(k+1)$$

і умовне математичне сподівання

$$\begin{aligned} E_k[y(k+2)] &= a_0 + a_1 E_k[y(k+1)] + a_2 E_k[y(k)] = \\ &= a_0 + a_1 [a_0 + a_1 y(k) + a_2 y(k-1) + \beta_1 \varepsilon(k)] + a_2 y(k) = \\ &= a_0 + a_0 a_1 + a_1^2 y(k) + a_1 a_2 y(k-1) + a_1 \beta_1 \varepsilon(k) + a_2 y(k) = \\ &= a_0 (1 + a_1) + (a_1^2 + a_2) y(k) + a_1 a_2 y(k-1) + a_1 \beta_1 \varepsilon(k). \end{aligned}$$

Можна знайти також наступне умовне математичне сподівання для оцінки прогнозу на три кроки:

$$E_k[y(k+3)] = a_0 + a_1 E_k[y(k+2)] + a_2 E_k[y(k+1)] =$$

$$= a_0(1 + a_1 + a_1^2 + a_2) + (a_1^3 + 2a_1a_2)y(k) + (a_1^2a_2 + a_2^2)y(k-1) + \\ + \beta_1(a_1^2 + a_2)\varepsilon(k).$$

З отриманих виразів для умовного математичного сподівання видно, що рекурсивна формула для прогнозу на довільне число кроків s може бути записана тільки у наступному вигляді:

$$E_k[y(k+s)] = a_0 + a_1 E_k[y(k+s-1)] + a_2 E_k[y(k+s-2)]. \quad (5.4.11)$$

Якщо корені характеристичного рівняння, записаного для (5.4.11), знаходяться всередині одиничного кола, то оцінка прогнозу асимптотично збігається до безумовного середнього значення

$$\lim_{s \rightarrow \infty} E[y(k+s)] = \frac{a_0}{1 - a_1 - a_2},$$

а для процесу ARKC(p,q) довільного порядку оцінку умовного прогнозу можна записати так:

$$E_k[y(k+s)] = a_0 + \sum_{i=1}^p a_i E_k[y(k+s-i)].$$

5.5 Побудова функції прогнозування на основі розв'язку різницевого рівняння

Розглянемо як приклад рівняння ARKC(1,1)

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k) + \beta_1 \varepsilon(k-1), \quad |a_1| < 1, \quad (5.5.1)$$

де $\varepsilon(k)$ – білий шум з нульовим середнім; $y(0) = y_0$ – відома початкова умова. Для однорідного рівняння $y(k) - a_1 y(k-1) = 0$ розв'язком є $A a_1^k$, де A – довільна константа.

Частковий розв'язок можна знайти за допомогою лагового оператора L в наступному вигляді:

$$y(k) = \frac{a_0}{1 - a_1} + \frac{\varepsilon(k)}{1 - a_1 L} + \frac{\beta_1 \varepsilon(k-1)}{1 - a_1 L}. \quad (5.5.2)$$

Використовуючи властивості лагового оператора, запишемо загальний розв'язок як

$$y(k) = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) + \beta_1 \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i-1) + A a_1^k. \quad (5.5.3)$$

Для того щоб знайти значення довільної константи скористаємось початковою умовою:

$$k=0: \quad y_0 = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(-i) + \beta_1 \sum_{i=0}^{\infty} a_1^i \varepsilon(-i-1) + A.$$

Запишемо розв'язок із врахуванням отриманого значення довільної константи:

$$\begin{aligned} y(k) = & \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) + \beta_1 \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i-1) + \\ & + \left[y_0 - \frac{a_0}{1-a_1} - \sum_{i=0}^{\infty} a_1^i \varepsilon(-i) + \beta_1 \sum_{i=0}^{\infty} a_1^i \varepsilon(-i-1) \right] a_1^k. \end{aligned}$$

Оскільки

$$\begin{aligned} \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) &= \varepsilon(k) + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + \dots + a_1^k \varepsilon(0) + a_1^{k+1} \varepsilon(-1) + \dots, \\ -a_1^k \sum_{i=0}^{\infty} a_1^i \varepsilon(-i) &= -a_1^k \varepsilon(0) - a_1^{k+1} \varepsilon(-1) - a_1^{k+2} \varepsilon(-2) - \dots, \end{aligned}$$

то

$$y(k) = \frac{a_0}{1-a_1} + \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i) + \beta_1 \sum_{i=0}^{k-1} a_1^i \varepsilon(k-i-1) + \left[y_0 - \frac{a_0}{1-a_1} \right] a_1^k. \quad (5.5.4)$$

Знайдемо рівняння для оцінки прогнозу із врахуванням того, що на момент $k=0$ відоме значення збурення $E_0[\varepsilon(0)] = \varepsilon_0$, тобто необхідно знайти $E_o[y(k)]$. Таким чином, функція прогнозу приймає наступний вигляд:

$$\begin{aligned} E_0[y(0+k) | y(0), \varepsilon(0)] = & \frac{a_0}{1-a_1} + E_0 \left[\sum_{i=0}^{k-1} a_1^i \varepsilon(k-i) + \beta_1 \sum_{i=0}^{k-1} a_1^i \varepsilon(k-1-i) \right] + \\ & + E_0 \left(\left[y_0 - \frac{a_0}{1-a_1} \right] a_1^k \right). \end{aligned}$$

Враховуючи, що

$$\sum_{i=0}^{k-1} a_1^i \varepsilon(k-i) = \varepsilon(k) + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + \dots + a_1^{k-1} \varepsilon(-1)$$

і

$$\beta_1 \sum_{i=0}^{k-1} a_1^i \varepsilon(k-1-i) = \beta_1 \varepsilon(k-1) + \beta_1 a_1 \varepsilon(k-2) + \beta_1 a_1^2 \varepsilon(k-3) + \dots + \beta_1 a_1^{k-1} \varepsilon(0),$$

отримаємо:

$$\hat{y}(0+k) = E_0[y(k)] = \frac{a_0}{1-a_1} + \beta_1 a_1^{k-1} \varepsilon_0 + \left[y_0 - \frac{a_0}{1-a_1} \right] a_1^k. \quad (5.5.5)$$

Рівняння (5.5.5) можна розглядати як функцію прогнозування на k кроків наперед на основі інформації, наявної на момент $k=0$.

Знайдемо функцію прогнозування на s кроків вперед на основі інформації, наявної на момент k . Спочатку зробимо заміну індексів в рівнянні (5.5.5):

$$\begin{aligned} \hat{y}(0+s) &= E_0[y(s)] = \frac{a_0}{1-a_1} + \beta_1 a_1^{s-1} \varepsilon_0 + \left[y_0 - \frac{a_0}{1-a_1} \right] a_1^s = \\ &= \left(\frac{a_0}{1-a_1} \right) (1 - a_1^s) + \beta_1 a_1^{s-1} \varepsilon_0 + y_0 a_1^s. \end{aligned}$$

Тепер виконаємо оновлення часового індексу для змінних y , i , ε на k одиниць вперед:

$$\hat{y}(k+s) = E_k[y(k+s)] = \left(\frac{a_0}{1-a_1} \right) (1 - a_1^s) + \beta_1 a_1^{s-1} \varepsilon(k) + y(k) a_1^s. \quad (5.5.6)$$

Отримане рівняння представляє собою функцію прогнозування на основі відомої інформації про процес на момент k , включно. Використовуючи наведені вище викладки, можна записати функції прогнозування для довільного числа кроків наступним чином:

$$\begin{aligned} s=1: \quad E_k[y(k+1)] &= a_0 + \beta_1 \varepsilon(k) + a_1 y(k); \\ s=2: \quad E_k[y(k+2)] &= \left(\frac{a_0}{1-a_1} \right) (1 - a_1^2) + \beta_1 a_1 \varepsilon(k) + a_1^2 y(k); \end{aligned}$$

$$s=3: \quad E_k[y(k+3)] = \left(\frac{a_0}{1-a_1} \right) (1 - a_1^3) + \beta_1 a_1^2 \varepsilon(k) + a_1^3 y(k);$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$\text{При цьому} \quad \lim_{s \rightarrow \infty} E_k[y(k+s)] = \frac{a_0}{1-a_1}.$$

Можна досить просто перейти від моделі ARKC(1,1) до моделі AP(1), якщо покласти $\beta_1 = 0$. Для AP(1) отримаємо функцію прогнозування у вигляді:

$$E_k[y(k+s)] = \left(\frac{a_0}{1-a_1} \right) (1 - a_1^s) + a_1^s y(k). \quad (5.5.7)$$

По аналогії можна знайти функції прогнозування для моделей іншої структури, табулювати отримані функції і використовувати їх при прогнозуванні процесів різної природи.

Примітка: Експериментально встановлено, що модель, яка може бути використана для прогнозування, повинна бути отримана (оцінена) на основі не менше ніж 50 спостережень (вимірювань) основної змінної.

Якщо модель побудована на основі вибірки даних довжиною N , то для рівняння ARKC(2,1)

$$y(k) = \hat{a}_0 + \hat{a}_1 y(k-1) + \hat{a}_2 y(k-2) + \hat{\varepsilon}(k) + \hat{\beta}_1 \hat{\varepsilon}(k-1),$$

функцію прогнозу можна записати наступним чином:

$$E_N[y(N+1)] = \hat{a}_0 + \hat{a}_1 E_N[y(N)] + \hat{a}_2 E_N[y(N-1)] + \hat{\beta}_1 \hat{\varepsilon}(k);$$

$$E_N[y(N+2)] = \hat{a}_0 + \hat{a}_1 E_N[y(N+1)] + \hat{a}_2 E_N[y(N)];$$

$$\cdot$$

$$\cdot$$

$$E_N[y(N+s)] = \hat{a}_0 + \hat{a}_1 E_N[y(N+s-1)] + \hat{a}_2 E_N[y(N+s-2)],$$

$$s \geq 2.$$

Приклад 5.1. Побудувати функцію прогнозування для рівняння AP(2)

$$y(k) = 3,0 + 0,9 y(k-1) - 0,2 y(k-2) + \varepsilon(k)$$

З початковими умовами: $y(0) = y_0$; $y(1) = y_1$.

а) Знайдемо однорідний розв'язок:

$$\alpha^2 - 0,9\alpha + 0,2 = 0;$$

$$\alpha_1 = 0,5; \alpha_2 = 0,4.$$

$$y^h(k) = A_1 0,5^k + A_2 0,4^k.$$

б) Частковий розв'язок для детермінованої частини:

$$d = 3,0 + 0,9d - 0,2d, \quad \text{або} \quad d = \frac{3,0}{0,3} = 10.$$

в) Частковий розв'язок для стохастичної частини:

$$y^{ps}(k) = \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i),$$

$$\text{де } \alpha_0 = 1; \alpha_1 = 0,9; \alpha_i = 0,9\alpha_{i-1} - 0,2\alpha_{i-2}, \quad \forall i > 1.$$

г) Повний розв'язок:

$$y(k) = 10,0 + A_1 0,5^k + A_2 0,4^k + \sum_{i=0}^{\infty} \alpha_i \varepsilon(k-i),$$

$$\alpha_i = 0,9\alpha_{i-1} - 0,2\alpha_{i-2}.$$

Для того щоб знайти значення невідомих довільних констант, використаємо початкові умови y_0 , y_1 . В результаті отримаємо два рівняння:

$$y_0 = 10,0 + A_1 + A_2 + \sum_{i=0}^{\infty} \alpha_i \varepsilon(-i),$$

$$y_1 = 10,0 + 0,5A_1 + 0,4A_2 + \sum_{i=0}^{\infty} \alpha_i \varepsilon(1-i),$$

з яких можна визначити невідомі константи. Із врахуванням початкових умов розв'язок прийме наступний вигляд:

$$\begin{aligned} y(k) = & 10,0 + 0,4^k [5,0(y_0 - 10) - 10(y_1 - 10)] + \\ & + 0,5^k [10(y_1 - 10) - 4(y_0 - 10)] + \sum_{i=0}^{k-1} \alpha_i \varepsilon(k-i). \end{aligned}$$

Значення залежної змінної в момент $k+s$ при початкових умовах $y(k)$ і $y(k-1)$ можна записати як

$$\begin{aligned} y(k+s) = & 10,0 + 0,4^s \{5,0[y(k-1) - 10] - 10[(y(k) - 10)]\} + \quad (5.5.8) \\ & + 0,5^s \{[10y(k) - 10] - 4[y(k-1) - 10]\} + \sum_{i=0}^{s-1} \alpha_i \varepsilon(k+s-i), \end{aligned}$$

де $y(k-1), y(k)$ – початкові умови відносно k – го моменту часу. В рівнянні (5.5.8) виконано заміну часової змінної k на $k+s$, але із врахуванням того, що прогнозоване значення визначається відносно k – го моменту часу. Тому верхнє граничне значення над знаком суми визначається як $k-2=k+s-2$, а із врахуванням того, що початковим моментом часу (на який відома необхідна інформація) є $k=1$, отримаємо $1+s-2=s-1$. Перевірка нижнього індекса свідчить, що $i=0 \Rightarrow \varepsilon(k+s)$ і $i=s-1 \Rightarrow \varepsilon(k+1)$, тобто границі зміни часового індекса для збурення визначені правильно.

Запишемо умовне математичне сподівання або прогноз на короткий проміжок часу

$$\begin{aligned} E_k[y(k+s)] &= 10,0 + 0,4^s \{ 5,0[y(k-1)-10] - 10[(y(k)-10)] + \\ &+ 0,5^s \{ [10y(k)-10] - 4[y(k-1)-10] \}. \end{aligned} \quad (5.5.9)$$

Безумовне середнє або асимптотичний прогноз приймає значення

$$\lim_{s \rightarrow \infty} E_k[y(k+s)] = 10,$$

тобто прогноз на нескінченності дорівнює безумовному середньому.

5.6 Прогнозування з мінімальною дисперсією

Функцію прогнозування можна будувати різними способами, які відрізняються припущеннями щодо процесу (який описується математично і прогнозується), критерію оптимальності оцінок (якщо він використовується) та вигляду самої функції. Побудова функції прогнозування, яка забезпечує мінімум дисперсії, ґрунтуються на таких припущеннях:

1. Процес, що прогнозується, є стаціонарним, гаусовим і має дробово раціональну спектральну густину.
2. Оптимальною вважається функція прогнозування, яка генерує оцінку прогнозу шляхом мінімізації дисперсії похибки прогнозу, тобто

$$E[y(k+s) - \hat{y}(k+s)]^2 \Rightarrow \min.$$

3. Значення $y(k+s)$, для якого оцінюється прогноз, є функцією всіх попередніх спостережень до моменту k включно.

Результат буде таким же, якщо замінити припущення про нормальність процесу припущенням про те, що він є двічі диференційованим, а функцією прогнозування є лінійна функція спостережень [7].

Розглянемо стохастичне різницеве рівняння

$$y(k+1) = -a y(k) + \varepsilon(k+1) + \beta \varepsilon(k), \quad (5.6.1)$$

де $\{\varepsilon(k)\} \sim N(0,1)$, тобто нормальну розподілений процес із нульовим середнім та одиничною дисперсією. Розглянемо спочатку функцію прогнозування на один крок за допомогою рівняння (5.6.1). Для того щоб знайти значення $y(k+1)$, необхідно знати $y(k), \varepsilon(k)$. Значення $y(k)$ – це значення останнього виміру, а $\varepsilon(k)$ необхідно обчислити.

Нехай відомі початкові значення $\varepsilon(0) = \varepsilon_0, y(0) = y_0$. Запишемо рівняння (5.6.1) у вигляді:

$$\varepsilon(\tau+1) - y(\tau+1) + \beta \varepsilon(\tau) = a y(\tau), \quad 0 \leq \tau \leq k$$

і знайдемо розв'язок для цього рівняння. Додамо до лівої і правої частин член $-\beta y(\tau)$ і отримаємо

$$[\varepsilon(\tau+1) - y(\tau+1)] + \beta [\varepsilon(\tau) - y(\tau)] = (a - \beta) y(\tau). \quad (5.6.2)$$

Тепер введемо нову змінну $x(\tau) = \varepsilon(\tau) - y(\tau)$ і запишемо різницеве рівняння для нової змінної

$$x(\tau+1) = -\beta x(\tau) + (a - b) y(\tau) \quad (5.6.3)$$

і знайдемо розв'язок цього рівняння методом прямої ітерації:

$$\begin{aligned} x(1) &= -\beta x(0) + (a - \beta) y(0), \\ x(2) &= -\beta x(1) + (a - \beta) y(1) = -\beta[-\beta x(0) + (a - b) y(0)] + (a - b) y(1) = \\ &\quad = -\beta^2 x(0) - \beta(a - b) y(0) + (a - \beta) y(1), \\ x(3) &= -\beta x(2) + (a - \beta) y(2) = \\ &\quad = -\beta^3 x(0) + \beta^2 (a - \beta) y(0) - \beta(a - \beta) y(1) + (a - \beta) y(2). \end{aligned}$$

Для довільного k розв'язок має наступний вигляд:

$$x(k) = (-\beta)^k x(0) + (a - \beta) \sum_{i=0}^{k-1} (-\beta)^{k-1-i} y(i). \quad (5.6.4)$$

Зробимо перевірку індексів:

- при $k = 1$ останній член в правій частині має вигляд: $(a - \beta) y(0)$;
- при $k = 2$ маємо наступне рівняння:

$$(a - \beta) \sum_{i=0}^1 (-\beta)^{1-i} y(i) = -\beta(a - \beta)y(0) + (a - \beta)y(1),$$

тобто, часові індекси в рівнянні (5.6.4) змінюються в належних інтервалах.

Використовуючи рівняння (5.6.4), можемо записати вираз для $\varepsilon(k)$:

$$\varepsilon(k) = [\varepsilon(0) - y(0)](-\beta)^k + y(k) + (a - \beta) \sum_{i=0}^{k-1} (-\beta)^{k-1-i} y(i), \quad (5.6.5)$$

а для довільного початкового моменту часу k_0 отримаємо

$$\varepsilon(k) = [\varepsilon(0) - y(0)](-\beta)^{k-k_0} + y(k) + (a - \beta) \sum_{i=0}^{k-1} (-\beta)^{k-1-i} y(i). \quad (5.6.6)$$

Оскільки для стаціонарного процесу $|\beta| < 1$, то перший член (5.6.5) і (5.6.6) наближається до нуля при $k_0 \rightarrow \infty$ для довільних початкових умов. Таким чином, $\varepsilon(k)$ може може бути обчислено безпосередньо за допомогою результатів вимірювань із рівняння

$$\varepsilon(k) = y(k) + (a - \beta) \sum_{i=-\infty}^{k-1} (-\beta)^{k-1-i} y(i).$$

Таким чином, можемо записати рівняння для прогнозування значення $y(k+1)$ на основі спостережень $y(k), y(k-1), y(k-2), \dots$ у вигляді:

$$\begin{aligned} \hat{y}(k+1|k) &= \hat{y}(k+1, k) = \hat{y}(k+1) = -a y(k) + \beta \varepsilon(k) = \\ &= -a y(k) + \beta y(k) + \beta(a - \beta) \sum_{i=-\infty}^{k-1} (-\beta)^{k-i-1} y(i) = \\ &= -(a - \beta) y(k) + \beta(a - \beta) \sum_{i=-\infty}^{k-1} (-\beta)^{k-i-1} y(i) = (\beta - a) \sum_{i=-\infty}^k (-\beta)^{k-i} y(i). \end{aligned} \quad (5.6.7)$$

Рівняння (5.6.7) можна переписати у зручнішій рекурсивній формі

$$\begin{aligned} \hat{y}(k+1, k) &= (\beta - a)[y(k) + \sum_{i=-\infty}^{k-1} (-\beta)^{k-i} y(i)] = \\ &= (\beta - a)[y(k) + (-\beta) \sum_{i=-\infty}^{k-1} (-\beta)^{k-1-i} y(i)] = -\beta \hat{y}(k, k-1) + (\beta - a) y(k). \end{aligned} \quad (5.6.8)$$

Тобто функція прогнозування на один крок описується РР першого порядку. Динаміка прогнозу визначається коефіцієнтом β .

За визначенням, похибка прогнозу дорівнює

$$f(k+1, k) = y(k+1) - \hat{y}(k+1, k) = \varepsilon(k+1),$$

а її математичне сподівання $E[f(k+1, k)] = 0$.

Альтернативна форма функції прогнозування на один крок

Введемо оператор зсуву $z x(k) = x(k+1)$ або $z^{-1} x(k) = x(k-1)$. Тепер рівняння

$$y(k+1) = -a y(k) + \varepsilon(k+1) + \beta \varepsilon(k) \quad (5.6.9)$$

можна переписати у вигляді

$$\begin{aligned} y(k+1) &= \frac{1 + \beta z^{-1}}{1 + a z^{-1}} \varepsilon(k+1) = \varepsilon(k+1) - \varepsilon(k+1) + \frac{1 + \beta z^{-1}}{1 + a z^{-1}} \varepsilon(k+1) = \\ &= \varepsilon(k+1) + \frac{-\varepsilon(k+1) - a \varepsilon(k) + \varepsilon(k+1) + \beta \varepsilon(k)}{1 + a z^{-1}} = \\ &= \varepsilon(k+1) + \frac{\beta - a}{1 + a z^{-1}} \varepsilon(k). \end{aligned} \quad (5.6.10)$$

Із рівняння (5.6.9) знайдемо вираз для випадкової величини

$$\varepsilon(k) = \frac{1 + a z^{-1}}{1 + \beta z^{-1}} y(k) \quad (5.6.11)$$

і підставимо цей вираз в (5.6.10) щоб уникнути змінної $\varepsilon(k)$ з метою зменшення невизначеності прогнозу:

$$y(k+1) = \varepsilon(k+1) + \frac{\beta - a}{1 + \beta z^{-1}} y(k). \quad (5.6.12)$$

Віднімемо від обох частин цього рівняння прогнозоване значення $\hat{y}(k+1, k)$ і знайдемо математичне сподівання квадрата похибки прогнозу:

$$E_k[y(k+1) - \hat{y}(k+1, k)]^2 = E[\varepsilon^2(k+1)] + E\left[\frac{\beta - a}{1 + \beta z^{-1}} y(k) - \hat{y}(k+1, k)\right]^2$$

при умові, що $E[y(k)\varepsilon(k+1)] = 0$. Якщо дисперсію збурення нормувати до одиниці, тобто $E[\varepsilon^2(k)] = 1$, то можна записати наступне співвідношення:

$$E[y(k+1) - \hat{y}(k+1, k)] \geq E[\varepsilon^2(k+1)] = 1,$$

де рівність виконується тільки при умові, що

$$\hat{y}(k+1, k) = \frac{\beta - a}{1 + \beta z^{-1}} y(k).$$

Таким чином, оптимальне з точки зору мінімуму дисперсії похибки прогнозу значення для прогнозу на один крок можна знайти за допомогою рівняння

$$\hat{y}(k+1, k) + \beta \hat{y}(k, k-1) = (\beta - a) y(k),$$

або

$$\hat{y}(k+1, k) = -\beta \hat{y}(k, k-1) + (\beta - a) y(k), \quad (5.6.13)$$

що співпадає з отриманим вище рівнянням (5.6.8).

Функція прогнозування на два кроки

Побудуємо функцію прогнозування на два кроки $y(k+2, k)$ для випадкового процесу AR(1,1):

$$y(k) = -a y(k-1) + \varepsilon(k) + \beta \varepsilon(k-1),$$

де $\{\varepsilon(k) : N(0,1)\}$ - випадковий процес з одиничною дисперсією та нульовим середнім значенням. Для моменту $k+2$ можна записати попереднє рівняння за допомогою оператора зсуву в наступній формі:

$$y(k+2) = \frac{1 + \beta z^{-1}}{1 + a z^{-1}} \varepsilon(k+2). \quad (5.6.14)$$

Права частина цього рівняння є лінійною функцією від $\varepsilon(k+2), \varepsilon(k+1), \varepsilon(k), \varepsilon(k-1), \dots$. При цьому змінні $\varepsilon(k), \varepsilon(k-1), \dots$ можна визначити за допомогою спостережень $y(k), y(k-1), \dots$, а випадкові змінні $\varepsilon(k+1), \varepsilon(k+2)$ не залежать від отриманих спостережень на момент k . Виділимо в правій частині члени з $\varepsilon(k+1), \varepsilon(k+2)$. З цією метою додамо та віднімемо $\varepsilon(k+2)$ в правій частині:

$$\begin{aligned} y(k+2) &= \varepsilon(k+2) - \varepsilon(k+2) + \frac{1+\beta z^{-1}}{1+z^{-1}} \varepsilon(k+2) = \\ &= \varepsilon(k+2) + \frac{-\varepsilon(k+2) - a\varepsilon(k+1) + \varepsilon(k+2) + \beta\varepsilon(k+1)}{1+a z^{-1}} = \\ &= \varepsilon(k+2) + \frac{\beta - a}{1+a z^{-1}} \varepsilon(k+1). \end{aligned}$$

Тепер домножимо чисельник на многочлен $(1+a z^{-1}-a z^{-1})$:

$$\begin{aligned} y(k+2) &= \varepsilon(k+2) + \frac{(\beta-a)(1+a z^{-1}-a z^{-1})\varepsilon(k+1)}{1+a z^{-1}} = \\ &= \varepsilon(k+2) + \frac{(\beta-a)(1+a z^{-1})}{1+a z^{-1}} \varepsilon(k+1) - \frac{a(\beta-a)\varepsilon(k)}{1+a z^{-1}} = \\ &= \varepsilon(k+2) + (\beta-a)\varepsilon(k+1) - \frac{a(\beta-a)}{1+a z^{-1}} \varepsilon(k), \end{aligned} \tag{5.6.15}$$

де $\varepsilon(k)$ можна обчислити за спостереженнями:

$$\varepsilon(k) = \frac{1+a z^{-1}}{1+\beta z^{-1}} y(k).$$

Якщо підставити цей вираз в (5.6.15), то отримаємо наступне рівняння:

$$y(k+2) = \varepsilon(k+2) + (\beta-a)\varepsilon(k+1) - \frac{a(\beta-a)}{1+a z^{-1}} \varepsilon(k) =$$

$$= \varepsilon(k+2) + (\beta - a)\varepsilon(k+1) - \frac{a(\beta - a)}{1 + \beta z^{-1}} y(k). \quad (5.6.16)$$

Нехай $\hat{y}(k+2, k)$ – прогноз на два кроки на основі інформації, що є в наявності на момент k , включно. Запишемо дисперсію похибки прогнозу за допомогою рівняння (5.6.16):

$$\begin{aligned} E[y(k+2) - \hat{y}(k+2, k)]^2 &= E[\varepsilon^2(k+2)] + (\beta - a)^2 E[\varepsilon^2(k+1)] + \\ &+ E[\hat{y}(k+2, k) + \frac{a(\beta - a)}{1 + \beta z^{-1}} y(k)]^2. \end{aligned} \quad (5.6.17)$$

Таким чином, можна записати, що

$$E[y(k+2) - \hat{y}(k+2, k)]^2 \geq 1 + (\beta - a)^2.$$

Із (5.6.17) випливає, що рівність можлива при умові, що

$$\hat{y}(k+2, k) = -\frac{a(\beta - a)}{1 + \beta z^{-1}} y(k). \quad (5.6.18)$$

Тобто функція двокривого прогнозу задовільняє наступному різницевому рівнянню:

$$\hat{y}(k+2, k) = -\beta \hat{y}(k+1, k) - a(\beta - a) y(k). \quad (5.6.19)$$

Знайдемо похибку двокривого прогнозу:

$$\begin{aligned} f(k+2, k) &= y(k+2) - \hat{y}(k+2, k) = \\ &= \varepsilon(k+2) + (\beta - a)\varepsilon(k+1) - \frac{a(\beta - a)}{1 + \alpha z^{-1}} y(k) + \frac{a(\beta - a)}{1 + \beta z^{-1}} y(k) = \\ &= \varepsilon(k+2) + (\beta - a)\varepsilon(k+1), \end{aligned} \quad (5.6.20)$$

тобто, це процес КС(2).

Загальна постановка задачі прогнозування з мінімумом дисперсії

Модель процесу можна представити у вигляді:

$$A(z^{-1})y(k) = B(z^{-1})\varepsilon(k), \quad (5.6.21)$$

де $A(z^{-1}), B(z^{-1})$ – поліноми відносно z виду:

$$A(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_n z^{-n}, \quad (5.6.22)$$

$$B(z^{-1}) = 1 + b_1 z^{-1} + \dots + b_n z^{-n}, \quad (5.6.23)$$

$$\{\varepsilon(k)\} \sim N(0,1),$$

тобто, $\varepsilon(k)$ – нормальню розподілена випадкова величина з нульовим середнім та одиничною дисперсією.

Необхідно побудувати найкращу, в смыслі мінімуму дисперсії похибки, функцію прогнозування на s кроків при умові, що відомі значення $y(k), y(k-1), y(k-2), \dots$, тобто із умови

$$E_k [y(k+s) - \hat{y}(k+s)]^2 \rightarrow \min_{\hat{y}},$$

де \hat{y} – оцінка прогнозу.

На основі рівняння (5.6.21) можна записати:

$$y(k+s) = \frac{B(z^{-1})}{A(z^{-1})} \varepsilon(k+s). \quad (5.6.24)$$

Права частина цього рівняння є лінійною комбінацією змінних $\varepsilon(k+s), \varepsilon(k+s-1), \varepsilon(k+s-2), \dots, \varepsilon(k+1), \varepsilon(k), \varepsilon(k-1), \varepsilon(k-2), \dots$.

Значення змінних $\varepsilon(k), \varepsilon(k-1), \dots$ можна обчислити за результатами вимірювань $y(k), y(k-1), \dots$, а значення $\varepsilon(k+1), \dots, \varepsilon(k+s)$ не залежать від вимірювань, тому що ці виміри ще невідомі. Для того щоб виділити вказані групи змінних, перепишемо праву частину (5.6.24) у вигляді (це допустимо для лінійних систем):

$$\frac{B(z^{-1})}{A(z^{-1})} \varepsilon(k+s) = F(z^{-1}) \varepsilon(k+s) + z^{-s} \frac{G(z^{-1})}{A(z^{-1})} \varepsilon(k+s), \quad (5.6.25)$$

де $F(z^{-1}), G(z^{-1})$ – поліноми від z^{-1} степені $s-1$ та $n-1$, відповідно, тобто,

$$F(z^{-1}) = 1 + f_1 z^{-1} + \dots + f_{s-1} z^{-(s-1)},$$

$$G(z^{-1}) = g_0 + g_1 z^{-1} + \dots + g_{n-1} z^{-(n-1)}. \quad (5.6.26)$$

Таке розщеплення є допустимим для лінійних систем, якщо правильно вибрати степені z та коефіцієнти поліномів F, G . Для першої складової в правій частині рівняння (5.6.25) можна записати

$$F(z^{-1})\varepsilon(k+s) = \varepsilon(k+s) + f_1\varepsilon(k+s-1) + \dots + f_{s-1}\varepsilon(k+s-s+1)$$

або

$$F(z^{-1})\varepsilon(k+s) = \varepsilon(k+s) + f_1\varepsilon(k+s-1) + \dots + f_{s-1}\varepsilon(k+1).$$

Враховуючи, що $z^{-s}\varepsilon(k+s) = \varepsilon(k)$, для чисельника другого члена в правій частині (5.6.25) отримуємо ковзне середнє порядку $n-1$:

$$g_0\varepsilon(k) + g_1\varepsilon(k-1) + \dots + g_{n-1}\varepsilon(k-n+1).$$

Тепер рівняння (5.6.24) можна записати у вигляді:

$$y(k+s) = F(z^{-1})\varepsilon(k+s) + \frac{G(z^{-1})}{A(z^{-1})}\varepsilon(k). \quad (5.6.27)$$

Другий член в правій частині цього рівняння є функцією випадкових змінних $\varepsilon(k), \varepsilon(k-1), \dots$, а тому його можна обчислити за результатами спостережень $y(k), y(k-1), \dots$, тобто, скориставшись рівнянням (5.6.21), отримаємо:

$$\varepsilon(k) = \frac{A(z^{-1})}{B(z^{-1})}y(k).$$

Тепер рівняння (5.6.27) можна звести до рівняння

$$y(k+s) = F(z^{-1})\varepsilon(k+s) + \frac{G(z^{-1})}{B(z^{-1})}y(k). \quad (5.6.28)$$

Перший член в правій частині цього рівняння є лінійною комбінацією змінних $\varepsilon(k+1), \varepsilon(k+2), \dots, \varepsilon(k+s)$, які не залежать від спостережень, а другий представляє собою лінійну функцію результатів спостережень.

Нехай $\hat{y}(k+s, k)$ – оцінка прогнозу на s кроків. Запишемо похибку цього прогнозу як

$$f(s) = y(k+s) - \hat{y}(k+s, k),$$

де $y(k+s)$ – спостереження, яке точно описується рівнянням (5.6.28). Віднімемо $\hat{y}(k+s, k)$ від обох частин рівняння (5.6.28), що приводить до виразу

$$y(k+s) - \hat{y}(k+s, k) = F(z^{-1})\varepsilon(k+s) + \frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k)$$

і знайдемо квадрат похибки прогнозу:

$$\begin{aligned} [y(k+s) - \hat{y}(k+s, k)]^2 &= \left[F(z^{-1})\varepsilon(k+s) + \frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k) \right]^2 = \\ &= [F(z^{-1})\varepsilon(k+s)]^2 + 2[F(z^{-1})\varepsilon(k+s)] \left[\frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k) \right] + \\ &\quad + \left[\frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k) \right]^2. \end{aligned}$$

Тепер запишемо умовне математичне сподівання для правої частини:

$$\begin{aligned} E_k [F(z^{-1})\varepsilon(k+s)]^2 + 2E_k \left\{ [F(z^{-1})\varepsilon(k+s)] \left[\frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k) \right] \right\} + \\ + E_k \left[\frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k) \right]^2. \end{aligned}$$

Оскільки випадкові величини $\varepsilon(k+s), \varepsilon(k+2), \dots$ не залежать від $y(k), y(k-1), \dots$, то середній член в правій частині дорівнює нулю, а тому

$$\begin{aligned} E[y(k+s) - \hat{y}(k+s, k)]^2 &= E_k [F(z^{-1})\varepsilon(k+s)]^2 + \\ &\quad + E_k \left[\frac{G(z^{-1})}{B(z^{-1})}y(k) - \hat{y}(k+s, k) \right]^2. \quad (5.6.29) \end{aligned}$$

Якщо знехтувати другим членом в правій частині цієї рівності, то можна записати нерівність:

$$\begin{aligned} E[y(k+s) - \hat{y}(k+s, k)]^2 &\geq \\ &\geq [\varepsilon(k+s) + f_1 \varepsilon(k+s-1) + f_2 \varepsilon(k+s-2) + \dots + f_{s-1} \varepsilon(k+1)]^2. \end{aligned}$$

В цій нерівності рівність досягається при умові, що

$$E_k \left[\frac{G(z^{-1})}{B(z^{-1})} y(k) - \hat{y}(k+s, k) \right]^2 = 0,$$

або

$$\hat{y}(k+s, k) = \frac{G(z^{-1})}{B(z^{-1})} y(k). \quad (5.6.30)$$

Звідси випливає, що оптимальне (за мінімумом дисперсії) значення прогнозу на s кроків можна обчислити за виразом:

$$\begin{aligned} \hat{y}(k+s, k) + b_1 \hat{y}(k+s, k-1) + \dots + b_n \hat{y}(k+s-n, k-n) &= \\ &= g_0 y(k) + g_1 y(k-1) + \dots + g_{n-1} y(k-n-1). \end{aligned} \quad (5.6.31)$$

Похибка прогнозування:

$$\begin{aligned} f(k+s, k) &= y(k+s) - \hat{y}(k+s, k) = F(z^{-1}) \varepsilon(k+s) = \\ &= \varepsilon(k+s) + f_1 \varepsilon(k+s-1) + \dots + f_{s-1} \varepsilon(k+1). \end{aligned} \quad (5.6.32)$$

Мінімально можливе значення дисперсії похибки прогнозування при даному підході

$$\text{var}[f(k+s, k)] = 1 + f_1^2 + f_2^2 + \dots + f_{s-1}^2.$$

Для того щоб скористатись формулами (5.6.31) і (5.6.32), необхідно знати коефіцієнти поліномів $F(z^{-1})$ і $G(z^{-1})$. Знайти невідомі коефіцієнти можна за допомогою тотожності

$$B(z^{-1}) = A(z^{-1})F(z^{-1}) + z^{-s} G(z^{-1}). \quad (5.6.33)$$

Фактично, $F(z^{-1})$ – це частка від відношення $B(z^{-1})/A(z^{-1})$, а поліном $G(z^{-1})$ – залишок від ділення. Коефіцієнти поліномів $F(z^{-1})$ і $G(z^{-1})$ можна визначити тепер прирівнюванням коефіцієнтів при одинакових степенях z^{-1} .

Для заданої структури поліномів $A(z^{-1}), B(z^{-1}), F(z^{-1})$ і $G(z^{-1})$

$$\begin{aligned}
B(z^{-1}) &= b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}, \\
A(z^{-1}) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}, \\
F(z^{-1}) &= 1 + f_1 z^{-1} + f_2 z^{-2} + \dots + f_{s-1} z^{-(s-1)}, \\
G(z^{-1}) &= g_0 + g_1 z^{-1} + \dots + g_{n-1} z^{-(n-1)}, \\
z^{-s} G(z^{-1}) &= g_0 z^{-s} + g_1 z^{-(s+1)} + g_2 z^{-(s+2)} + \dots + g_{n-1} z^{-(n-1+s)}
\end{aligned} \tag{5.6.34}$$

отримаємо наступні рівності для визначення коефіцієнтів:

$$\begin{aligned}
\text{при } z^{-1}: \quad b_1 &= a_1 + f_1, \\
\text{при } z^{-2}: \quad b_2 &= a_2 + a_1 f_1 + f_2, \\
&\vdots \\
\text{при } z^{-(s-1)}: \quad b_{s-1} &= a_{s-1} + a_{s-2} f_1 + a_{s-3} f_2 + \dots + a_1 f_{s-2} + f_{s-1}, \\
\text{при } z^{-s}: \quad b_s &= a_k + a_{k-1} f_1 + a_{k-2} f_2 + \dots + a_1 f_{s-1} + g_0, \\
\text{при } z^{-(s+1)}: \quad b_{s+1} &= a_{k+1} + a_k f_1 + a_{k-1} f_2 + \dots + a_2 f_{s-1} + g_1, \\
&\vdots \\
\text{при } z^{-n}: \quad b_n &= a_n + a_{n-1} f_1 + a_{n-2} f_2 + \dots + a_{n-s-1} f_{s-1} + g_{n-s}, \\
\text{при } z^{-(n+1)}: \quad 0 &= a_n f_1 + a_{n-1} f_2 + \dots + a_{n-s+2} f_{s-1} + g_{n-s+1}, \\
&\vdots \\
\text{при } z^{-(n+s-1)}: \quad 0 &= a_n f_{s-1} + g_{n-1}.
\end{aligned} \tag{5.6.35}$$

Таким чином, коефіцієнти f_i, g_i можна обчислити за допомогою простих рекурсивних співвідношень при відомих $A(z^{-1})$ і $B(z^{-1})$, а також вибраному числі кроків s для прогнозування. Очевидно, що це число не може бути великим.

Примітка 1: Для даної постановки задачі кращою є лінійна оцінка прогнозу і результат прогнозування не залежить критично від вибору критерію мінімізації похибки прогнозу. Якщо послідовність $\{y(k)\}$ має нормальній розподіл, то функція прогнозування не зміниться, якщо вибрати за критерій мінімізації

$$E w[y(k+s) - \hat{y}(k+s, k)]^2,$$

де w – довільна симетрична функція.

Примітка 2. Вирішальну роль для справедливості наведеного вище виводу функції прогнозування відіграє той факт, що $\varepsilon(k)$ і $\varepsilon(l)$ є взаємно

незалежними, якщо $k \neq l$. Якщо випадкові змінні $\varepsilon(k)$ і $\varepsilon(l)$ є залежними одна від одної, то математичне сподівання добутку $\varepsilon(k+l)$ на довільну функцію від $y(k), y(k-1), \dots$ не обов'язково буде дорівнювати нулю при $l > 0$.

Приклад 5.2: Для стохастичного процесу $\{y(k)\}$, який визначається рівнянням

$y(k) = 1,5y(k-1) + 0,5y(k-2) = 2[\varepsilon(k) - 1,2\varepsilon(k-1) + 0,6\varepsilon(k-2)]$,
де $\{\varepsilon(k)\} \sim N(0,1)$, записати функцію прогнозування на три кроки, $s = 3$, яка мінімізує середньоквадратичну похибку прогнозу.

Для даного прикладу

$$A(z^{-1}) = 1 - 1,5z^{-1} + 0,5z^{-2},$$

$$B(z^{-1}) = 1 - 1,2z^{-1} + 0,6z^{-2}.$$

Використовуючи співвідношення (5.6.35), можна визначити коефіцієнти функції прогнозування:

$$b_1 = a_1 + f_1 \quad \text{або} \quad f_1 = b_1 - a_1 = -1,2 + 1,5 = 0,3;$$

$$b_2 = a_2 + a_1f_1 + f_2 \quad \text{або} \quad f_2 = b_2 - a_2 - a_1f_1 = 0,6 - 0,5 + 1,5 \cdot 0,3 = 0,55;$$

Таким чином, $F(z^{-1}) = 1 + 0,3z^{-1} + 0,55z^{-2}$.

Використовуючи тотожність (5.6.33), запишемо:

$$1 - 1,2z^{-1} + 0,6z^{-2} = (1 - 1,5z^{-1} + 0,5z^{-2}) \cdot (1 + 0,3z^{-1} + 0,55z^{-2}) + z^{-3}(g_0 + g_1z^{-1}).$$

Звідси визначимо, що

$$G(z^{-1}) = 0,675 + 0,275z^{-1},$$

тобто, всі необхідні коефіцієнти для функції прогнозування знайдено.

5.7 Функції прогнозування, отримані в даному розділі

Запишемо функції прогнозування динаміки процесу на довільне число s кроків, отримані в даному розділі:

- для рівняння першого порядку без розв'язку:

$$E_k[y(k+s)] = \hat{y}(k+s, k) = a_0 \sum_{i=0}^{s-1} a_1^i + a_1^s y(k);$$

- для рівняння AP(2) без розв'язку:

$$E_k[y(k+s)] = \hat{y}(k+s, k) = a_0 + a_1 E_k[y(k+s-1)] + a_2 E_k[y(k+s-2)];$$

- на основі повного розв'язку рівняння APKC(1,1):

$$E_k[y(k+s)] = \frac{a_0}{1-a_1} (1 - a_1^s) + \beta_1 a_1^{s-1} \varepsilon(k) + a_1^s y(k);$$

- на основі повного розв'язку рівняння AP(1):

$$E_k[y(k+s)] = \frac{a_0}{1-a_1} (1 - a_1^s) + a_1^s y(k);$$

- функція прогнозування з мінімальною дисперсією:

$$\hat{y}(k+s, k) = \frac{G(z^{-1})}{B(z^{-1})} y(k).$$

Інші види прогнозуючих моделей будуть розглянуті нижче в розділах, присвячених моделюванню нестационарних гетероскедастичних та коінтегрованих процесів, а також оптимальному оцінюванню станів за допомогою фільтра Калмана (ФК). ФК представляє собою рекурсивну процедуру для обчислення оптимальних оцінок стану процесу в умовах впливу випадкових збурень стану і шумів вимірювань. В процесі обчислення поточних оцінок стану він знаходить також однокроковий прогноз оцінок стану на наступний крок на основі попередньої оцінки.

Існують також інші методи прогнозування, які суттєво відрізняються від наведених в даному розділі. Зокрема, це підхід на основі теорії формування Байєсового висновку, який забезпечує отримання прогнозів в умовах наявності суттєвих невизначеностей, коли задача погано формалізується. Для визначення оцінки прогнозу необхідно розглядати декілька взаємозв'язаних процесів, коли статистичні та інші типи невизначеностей не можуть бути описані з прийнятною точністю за допомогою рівнянь авторегресії, APKC та у просторі станів. Відомий Байєсові метод використовують у різних формах постановки задачі та використовують різні підходи до їх розв'язання. Однією з популярних модифікацій даного методу є представлення проблеми прогнозування і

прийняття рішення у вигляді динамічної Байєсової мережі, яка є спрямованим ацикличним графом. Вершинами графа виступають змінні стану процесу.

5.8 Оцінювання якості прогнозу

Важливим моментом процесу прогнозування є об'єктивне визначення якості отриманого прогнозу. Оскільки прогнозовані значення – випадкові величини, то для оцінювання їх якості необхідно використовувати декілька статистичних критеріїв. Рис. 5.1 ілюструє часову вісь та відрізки часу, на яких виконується оцінювання моделі і перевірка якості прогнозу.



Рис. 5.1. Види прогнозування за часовим рядом

Наявну вибірку даних доцільно розділити на навчальну та перевірочну. На навчальній вибірці виконується оцінювання параметрів моделі процесу і реалізується так званий „історичний” прогноз, який дає змогу встановити якість однокрокового прогнозу на цьому участку ряду. Прогноз на перевірочній частині вибірки даних в науковій літературі називають ще прогнозом *ex post*. В різних емпіричних дослідженнях рекомендують залишати для перевірки (5 – 40) % значень ряду даних.Хоча при аналізі коротких рядів доцільно значно більшу частину ряду використовувати для оцінювання параметрів моделі. Прогнозування значень поза вибіркою даних називають прогнозом *ex ante* (рис. 5.1).

Як правило, для оцінювання якості прогнозів використовують множину взаємно доповнюючих статистичних критеріїв. Наприклад, значення середньоквадратичної похибки залежить від масштабу даних, а тому недостатньо використовувати тільки цей статистичний параметр для аналізу якості прогнозу. Поглиблене оцінювання якості прогнозів досягається за рахунок використання критеріїв, які дають відносні оцінки в

процентах. Розглянемо деякі статистичні критерії якості прогнозу та їх призначення.

Дисперсія і стандартне відхилення прогнозу

Досить часто проста сума похибок прогнозів дорівнює нулю, оскільки похибки приймають різні знаки, а тому необхідно використовувати інші міри похибки. Ступінь розсіювання значень змінної навколо її середнього можна виміряти за допомогою стандартного відхилення $\sigma(k)$, яке дорівнює квадратному кореню з дисперсії, тобто:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2}, \quad (5.8.1)$$

де $y(k)$ – фактичне значення змінної; $\hat{y}(k)$ – прогноз. (Ділення на $N-1$, а не на N має смисл при невеликій потужності (довжині) ряду.)

Стандартне відхилення залишків – один із основних показників якості прогнозу. Це досить широко застосовувана статистична характеристика, яка виявляється корисною при аналізі поведінки процесів різної природи. Вона використовується, наприклад, як міра ризику (волатильності) при аналізі фінансових процесів, як характеристика надійності в управлінні запасами та інших прикладних задачах управління, а також як міра розсіювання значень змінних стану при аналізі систем різної природи.

Оскільки дисперсія є квадратичною характеристикою процесу, то в процесі її обчислення не відбувається взаємна компенсація значень відхилень від середнього. Крім того, дисперсію $\sigma^2(k)$ і стандартне відхилення можна перевіряти на статистичну значимість, що в деякій мірі сприяє поглибленню аналізу похибки прогнозу. Вище було показано, що при застосуванні для прогнозування рівнянь АРКС прогноз незміщений (математичне сподівання похибок прогнозів дорівнює нулю), але дисперсія прогнозованих значень прямо пропорціональна числу кроків прогнозування. Однак вона збігається до скінченної величини для стаціонарних процесів.

При відносно невеликому горизонті прогнозування можна стверджувати, що майбутнє значення прогнозованого показника попадає в інтервал, який визначається як \pm два стандартних відхилення від обчисленого значення прогнозу.

Очевидно, що дисперсії та стандартного відхилення недостатньо для аналізу якості прогнозів. Наприклад, якщо прогноз попиту дорівнює 2500 одиниць товару із стандартним відхиленням похибки 100, то інтервал $2300 \div 2700$ буде достатньо інформативним. Однак, якщо при такому ж значенні прогнозу стандартне відхилення складе 400 одиниць, то

відповідний інтервал $1700 \div 3300$ навряд чи буде корисним для практичного використання.

Середнє абсолютне значення похибки

Обчислення середнього абсолютноого значення похибки (САП) ґрунтуються на виразі для експоненціального середнього і має вигляд:

$$\begin{aligned} САП(k) &= \alpha \cdot |y(k) - \hat{y}(k)| + (1 - \alpha) САП(k-1) = \\ &= \alpha |e(k)| + (1 - \alpha) САП(k-1), \end{aligned} \quad (5.8.2)$$

де $0 < \alpha < 1$; $e(k)$ – похибка прогнозу. При такому обчисленні середнє абсолютне значення завжди невід'ємне.

САП можна обчислити також за іншим виразом:

$$САП = \frac{1}{N} \sum_{k=1}^N |y(k) - \hat{y}(k, k)|.$$

Для досить широкого класу статистичних розподілів випадкових величин значення стандартного відхилення є дещо більшим САП і строго пропорціональним йому. Коефіцієнт пропорціональності коливається для різних розподілів між 1,2 і 1,3 (для нормального розподілу це значення дорівнює $\sqrt{\pi/2} = 1,2533$) [45]. Таким чином, можна записати, що

$$\sigma(k) \approx 1,25 САП. \quad (5.8.3)$$

Виходячи з цього, процедуру оцінювання *стандартної похибки* прогнозу можна зформулювати так:

1. Обчислити поточне значення похибки прогнозу як $e(k) = y(k) - \hat{y}(k)$.
2. Обчислити нове значення САП за допомогою рівняння (5.8.2).
3. Знайти поточне стандартне відхилення за виразом (5.8.3).

Середню абсолютну похибку і її стандартне відхилення використовують, як правило, для визначення якості оцінки прогнозу одночасно.

Середній квадрат похибки і сума квадратів похибок

Якщо середній квадрат похибок (СКП) визначається для розповсюдженого випадку однокрокового прогнозування при довжині часового ряду N , то СКП визначається за формулою:

$$СКП = \frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2. \quad (5.8.4)$$

У випадках, коли прогнозування виконується на s кроків відносно моменту часу k , то СКП обчислюється за виразом:

$$СКП = \frac{1}{s} \sum_{i=1}^s [y(k+i) - \hat{y}(k+i, k)]^2. \quad (5.8.5)$$

Очевидно, що формули (5.8.4) і (5.8.5) однакові, але кожна форма чітко відображає тип прогнозу.

Більшість відомих пакетів для математичного моделювання і виконання статистичних розрахунків розраховують суму квадратів похибок за виразом:

$$СмКП = \sum_{k=1}^N [y(k) - \hat{y}(k)]^2. \quad (5.8.6)$$

Це корисний інтегральний показник якості, який використовують для порівняльного аналізу різних методів прогнозування.

Саме СКП і СмКП є найбільш поширеними критеріями, які використовуються при порівняльному аналізі і виборі кращої моделі для обчислення оцінок прогнозів. Однак, їх також не достатньо для поглибленого аналізу результатів.

Середня абсолютна похибка в процентах

Середня абсолютна похибка в процентах (САПП) – це середнє абсолютних значень похибок оцінок прогнозу в процентах відносно фактичного значення показника:

$$САПП = \frac{1}{N} \sum_{k=1}^N \frac{|y(k) - \hat{y}(k)|}{|y(k)|} \times 100\% = \frac{1}{N} \sum_{k=1}^N \frac{|e(k)|}{|y(k)|} \times 100\%, \quad (5.8.7)$$

або у випадку прогнозування на s кроків відносно k -го моменту:

$$САПП = \frac{1}{s} \sum_{i=1}^s \frac{|y(k+i) - \hat{y}(k+i, k)|}{|y(k+i)|} \times 100\% = \frac{1}{s} \sum_{i=1}^s \frac{|e(k+i)|}{|y(k+i)|} \times 100\%. \quad (5.8.8)$$

Оскільки ця міра характеризує відносну якість прогнозу, то її використовують, в основному, для порівняння точності прогнозів

різноманітних об'єктів (процесів) прогнозування. Однак, вона є завжди корисною при виконанні порівняльного аналізу якості прогнозування одного й того ж процесу різними методами, оскільки відносна міра є чіткою і зрозумілою для дослідника і практичного користувача. Типові значення САПП та їх пропонована інтерпретація наведені в таблиці 5.1 [45].

Якщо в формулах (5.8.7) і (5.8.8) $y(k)$ або $y(k+i)$ прямують до нуля, то значення САПП прямуватиме до нескінченності. Про це необхідно пам'ятати при застосуванні даного критерію якості прогнозу. Для того щоб виконати обчислення даного критерію в таких випадках, нульові значення $y(k)$ або $y(k+i)$ необхідно пропускати з відповідним корегуванням значення N або s . Можливо, що такий підхід не відповідає деяким вимогам статистичного аналізу даних, але він дає можливість наблизено і більш повно виконати аналіз якості прогнозування.

Таблиця 5.1. Інтерпретація типових значень критерію САПП

САПП, %	Інтерпретація
< 10	Висока точність
10 – 20	Хороша точність
20 – 50	Задовільна точність
> 50	Незадовільна (неприйнятна) точність

Середня похибка (СП) і середня похибка в процентах (СПП)

Середня похибка – це не відносний показник, вона характеризує ступінь зміщення прогнозованих значень від фактичних і розраховується за формулою:

$$SP = \frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)] = \frac{1}{N} \sum_{k=1}^N e(k), \quad (5.8.9)$$

або

$$SP = \frac{1}{s} \sum_{i=1}^s [y(k+s) - \hat{y}(k+s, k)]. \quad (5.8.10)$$

Очевидно, що СП буде зменшуватись у випадках, коли похибки мають різні знаки.

Середню похибку в процентах (СПП) обчислюють за виразом:

$$SPP = \frac{1}{N} \sum_{k=1}^N \frac{|y(k) - \hat{y}(k)|}{y(k)} \times 100\%, \quad (5.8.11)$$

або

$$CPI = \frac{1}{s} \sum_{i=1}^s \frac{|y(k+s) - \hat{y}(k+s, k)|}{y(k+s)} \times 100\%. \quad (5.8.12)$$

СПП також характеризує зміщеність прогнозу. Якщо втрати при прогнозуванні, зв'язані із завищеннем фактичного майбутнього значення, врівноважуються заниженням, то ідеальний прогноз повинен бути незміщеним. В такому випадку СП і СПП повинні прямувати до нуля. Очевидно, що нуль – це ідеальні значення і забезпечити його на практиці неможливо. Емпірично встановлено, що прийнятними значеннями для СПП (так само як і для САПП) є $\leq 5\%$.

Максимальна і мінімальна абсолютна похибка

Очевидно, що максимальна абсолютна похибка (МАП) може бути визначена як

$$MAP = \max_k \{|y(k) - \hat{y}(k)|\}, \quad 1 \leq k \leq N, \quad (5.8.13)$$

або

$$MAP = \max_i \{|y(k+1) - \hat{y}(k+1, k)|, \dots, |y(k+s) - \hat{y}(k+s, k)|\}; \quad (5.8.14)$$

$$1 \leq i \leq s.$$

Мінімальна абсолютна похибка (MiAP) визначається як

$$MiAP = \min_k \{|y(k) - \hat{y}(k)|\}, \quad 1 \leq k \leq N, \quad (5.8.15)$$

або

$$MiAP = \min_i \{|y(k+1) - \hat{y}(k+1, k)|, \dots, |y(k+s) - \hat{y}(k+s, k)|\}, \quad (5.8.16)$$

$$1 \leq i \leq s.$$

Критерії МАП і MiAP також можуть бути корисними при виконанні порівняльного аналізу кількох методів прогнозування, особливо якщо нас цікавлять максимально або мінімально можливі відхилення прогнозів від фактичних значень на заданому інтервалі.

Коефіцієнт нерівності Тейла

Коефіцієнт нерівності Тейла U – це важливий індикатор якості моделі і прогнозу. За визначенням, $0 \leq U \leq 1$. Якщо $U = 1$, то модель має практично нульові (неприйнятні) прогнозуючі властивості, що випливає з формули для обчислення U :

$$U = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2}}{\sqrt{\frac{1}{N} \sum_{k=1}^N y^2(k)} + \sqrt{\frac{1}{N} \sum_{k=1}^N \hat{y}^2(k)}}. \quad (5.8.17)$$

При $U = 0$ прогнозовані значення співпадають з фактичними значеннями ряду – модель ідеальна. Тобто U дає можливість встановити придатність моделі (методу) в принципі для оцінювання прогнозу.

Коефіцієнт Тейла можна розкласти на три складові; складова, пропорційна зміщенню:

$$U^M = \frac{(\bar{y} - \bar{\hat{y}})^2}{\frac{1}{N} \sum_{i=1}^N [y(i) - \hat{y}(i)]^2}; \quad (5.8.18)$$

складова, пропорційна дисперсії:

$$U^S = \frac{(\sigma_{факт} - \sigma_{модель})^2}{\frac{1}{N} \sum_{i=1}^N [y(i) - \hat{y}(i)]^2}; \quad (5.8.19)$$

і складова, пропорційна коваріації:

$$U^C = \frac{2(1-\rho)(\sigma_{факт} \cdot \sigma_{модель})}{\frac{1}{N} \sum_{i=1}^N [y(i) - \hat{y}(i)]^2}, \quad (5.8.20)$$

де ρ – коефіцієнт кореляції між залишками моделі.

Величина U^M використовується для перевірки факту, чи є систематичне відхилення середніх фактичного і прогнозованого ряду. Іншими словами, чи існує систематичне зміщення на виході моделі в той чи інший бік. Чим меншим є обчислене значення U^M , тим кращою є модель. Якщо $U^M = 0$, то прогнозовані значення не містять зміщення і модель адекватна процесу за цим показником.

Величина U^S використовується для тестування динамічних властивостей моделі, тобто чи відповідає її дисперсія дисперсії фактичного ряду. Наприклад, модель може відтворювати систематично менші коливання

ніж ті, які має фактичний ряд. Чим меншим є значення U^S , тим меншим є відхилення дисперсії виходу моделі від дисперсії ряду.

Накінець, складова, пропорційна коваріації, є мірою корельованості фактичного і прогнозованого за моделлю ряду. Зазначимо, що за побудовою останнього критерію для нього виконується рівність:

$$U^M + U^S + U^C = 1.$$

Якість моделі також визначається тим, наскільки точно вона може прогнозувати зміни напряму розвитку процесу, тобто нахил або знак тренду. Моделі можуть мати високу точність відтворення ряду, але погано прогнозувати тренди або цикли. Інші моделі, навпаки, моджуть мати меншу точність (адекватність), але кращі динамічні властивості. Таким чином, завжди необхідно шукати компроміс між точністю моделі, тобто якістю апроксимації ряду, і її динамічними властивостями. Однак, для аналізу цієї властивості формальних тестів немає. Можна наблизено встановити якість моделі щодо відтворення динаміки ряду шляхом візуального аналізу фактичного та спрогнозованого ряду.

Ще однією характеристикою якості моделі є її чутливість до початкового (стартового) періоду імітаційного моделювання. Якщо модель генерує результати, які можна наблизено класифікувати як інваріантні до початкових умов, то вона вважається хорошиою. Інакше, якщо результати імітаційного моделювання залежать від початкових умов, можна припустити, що модель неякісна. Наприклад, вона може мати нестационарність деякого типу.

5.9 Довірчий інтервал для прогнозу

При використанні для обчислення прогнозу регресійних та рівнянь інших типів ми отримуємо точкову оцінку. Однак, така оцінка є далеко не завжди значущою. При прогнозуванні необхідно визначити інтервал, всередині якого з достатньою ступінню впевненості можна очікувати появу фактичного значення показника. В регресійному аналізі граници цього інтервалу задаються за допомогою довірчого інтервалу.

Довірчий інтервал – це інтервал, який містить сам прогноз і в якому з визначеною степінню впевненості можна очікувати появу фактичного значення прогнозованої змінної.

Так, значення прогнозу $\hat{y}(k+1) = 1500$ з довірчим інтервалом ± 150 і ступінню впевненості 95% означає, що з ймовірністю $\approx 0,95$ очікується, що наступне значення прогнозованої змінної буде знаходитись в межах $1350 \div 1650$. Якби довірчий інтервал складав ± 500 , то з ймовірністю $\approx 0,95$ очікується, що майбутнє значення буде знаходитись в інтервалі $1000 \div 2000$.

В регресійному аналізі *мінімальна ширина довірчого інтервалу* відповідає точці (\bar{y}, \bar{k}) – “середині спостережень”. По обидва боки від цієї середини довжина інтервалу збільшується. Для того щоб визначити довірчий інтервал, необхідно знайти стандартну похибку рівняння регресії S_r , яка обчислюється всіма пакетами статистичної обробки даних. Величину S_r ще називають середньоквадратичним відхиленням і обчислюють за формулою:

$$S_r = \sqrt{\frac{1}{N-p} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2}, \quad (5.9.1)$$

де p – число параметрів моделі, які оцінюються в процесі її побудови. Суму $\sum [y(k) - \hat{y}(k)]^2$ можна обчислити також як різницю між загальною сумою квадратів і сумою квадратів значень, знайденою за регресією.

Тепер можна визначити стандартну похибку прогнозу як

$$S_{\hat{y}(k)} = S_r \sqrt{1 + \frac{1}{N} + \frac{(k - \bar{k})^2}{\sum_{k=1}^N (k - \bar{k})^2}}, \quad (5.9.2)$$

тобто стандартна похибка прогнозу залежить від довжини ряду N і віддалі (в часі) від середини часового періоду, що розглядається, до моменту прогнозування [41]. Видно, що збільшення довжини ряду приводить до зменшення стандартної похибки прогнозу. Якщо число кроків прогнозування відносно N -го значення ряду дорівнює τ , то член $(k - \bar{k})$ в виразі (5.9.2) необхідно замінити на $(\tau + (N-1)/2)$. Таку процедуру введення нового позначення називають ще *стандартним прогностичним перепозначенням*.

Із врахуванням введеного позначення вираз для стандартної похибки приймає вигляд:

$$S_{\hat{y}(k+\tau)} = S_r \sqrt{1 + \frac{1}{N} + \frac{(\tau + (N-1)/2)^2}{\sum_{k=1}^N k^2 - (\sum k)^2 / N}}, \quad (5.9.3)$$

де S_r для рівняння регресії $y(k) = a_0 + b k$ можна обчислити за виразом:

$$S_r = \sqrt{\frac{\sum y^2(k) - a_0 \sum y(k) - b \sum k y(k)}{N-2}}. \quad (5.9.4)$$

Скориставшись стандартною похибкою прогнозу, можна визначити довірчі інтервали. Так, наближеними 99%, 95% і 68%-ми довірчими інтервалами будуть значення $\pm 3S_{\hat{y}(k+\tau)}$, $\pm 2S_{\hat{y}(k+\tau)}$ і $\pm S_{\hat{y}(k+\tau)}$, відповідно.

5.10 Комбінування прогнозів, отриманих за різними методами

5.10.1 Усереднення прогнозів (вагові коефіцієнти однакові)

Для двох методів прогнозування середнє визначається досить просто:

$$\hat{y}_c(k) = \frac{\hat{y}_1(k) + \hat{y}_2(k)}{2}, \quad (5.10.1)$$

де $\hat{y}_c(k)$ – комбінований прогноз; $\hat{y}_1(k)$, $\hat{y}_2(k)$ – прогнози, отримані за різними методами. Якщо окремі прогнози не зміщені (це повинен забезпечувати метод прогнозування), то комбінований прогноз також буде незміщеним. Похибка комбінованого прогнозу:

$$e_c(k) = y(k) - \hat{y}_c(k) = y(k) - \frac{\hat{y}_1(k) + \hat{y}_2(k)}{2} = \frac{e_1(k) + e_2(k)}{2}, \quad (5.10.2)$$

де $y(k)$ – фактичне значення прогнозованої змінної.

Дисперсія похибки комбінованого прогнозу:

$$\begin{aligned} \text{var}\left[\frac{e_1(k) + e_2(k)}{2}\right] &= E\left[\frac{e_1(k) + e_2(k)}{2}\right]^2 = \frac{1}{4}E[e_1^2(k) + 2e_1(k)e_2(k) + e_2^2(k)] = \\ &= \frac{1}{4}\{E[e_1^2(k)] + 2E[e_1(k)e_2(k)] + E[e_2^2(k)]\} = \\ &= \frac{1}{4}\left[\sigma_1^2 + 2\frac{E[e_1(k)e_2(k)]}{\sigma_1\sigma_2}\sigma_1\sigma_2 + \sigma_2^2\right] = \frac{\sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2}{4}. \end{aligned}$$

Таким чином, дисперсія комбінованого прогнозу обчислюється за виразом:

$$\sigma_c^2 = \frac{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}{4}, \quad (5.10.3)$$

де ρ – коефіцієнт кореляції між похибками прогнозу. Якщо похибки прогнозування за двома моделями незалежні, тобто, $\rho=0$, то формула (5.10.3) спрощується:

$$\sigma_c^2 = \frac{\sigma_1^2 + \sigma_2^2}{4}. \quad (5.10.4)$$

Таким чином, якщо дисперсії близькі за значеннями і похибки оцінок прогнозів незалежні, то дисперсія комбінованої похибки буде значно меншою будь-якої з двох дисперсій. Наприклад, нехай $\sigma_1^2 = \sigma_2^2 = 100$:

$$\sigma_c^2 = \frac{100+100}{4} = 50.$$

Але навіть при існуванні досить високої кореляції між похибками прогнозування дисперсія похибки комбінованого прогнозу буде меншою ніж дисперсія кожного методу окремо. Наприклад, нехай $\sigma_1^2 = \sigma_2^2 = 100$ і $\rho = 0.8$:

$$\sigma_c^2 = \frac{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}{4} = \frac{100+100+2 \cdot 0.8 \cdot 10 \cdot 10}{4} = 90.$$

Навіть в цій ситуації спостерігається зменшення дисперсії похибки прогнозу після усереднення оцінок, отриманих за двома методами.

Однак, ситуація змінюється у випадку, коли дисперсії індивідуальних похибок сильно відрізняються. Наприклад, нехай $\sigma_1^2 = 100$, $\sigma_2^2 = 16$ і $\rho = 0.8$:

$$\sigma_c^2 = \frac{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}{4} = \frac{100+16+2 \cdot 0.8 \cdot 10 \cdot 4}{4} = 45.$$

Таким чином, якщо дисперсії похибок сильно відрізняються, то просте усереднення результатів не потрібно робити,. Висновок такий: просте усереднення можна застосовувати у випадках, коли дисперсії індивідуальних похибок прогнозування приблизно рівні або не дуже сильно відрізняються за своїми значеннями.

5.10.2 Зважене усереднення прогнозів

Якщо інформація щодо характеристик індивідуальних прогнозів відсутня, то можна присвоїти різні вагові коефіцієнти окремим прогнозам на основі суб'єктивних або експертних суджень:

$$\hat{y}_c(k) = w_1 \hat{y}_1(k) + w_2 \hat{y}_2(k),$$

де w_1, w_2 – вагові коефіцієнти. Очевидно, що більші значення вагових коефіцієнтів необхідно присвоювати тим індивідуальним прогнозам, які мають меншу дисперсію похибок. При цьому для коректності обчислень необхідно, щоб виконувалась умова: $w_1 + w_2 = 1$.

5.10.3 Вибір вагових коефіцієнтів за допомогою похибок прогнозів

Як правило, похибки прогнозів для конкретних моделей і процесів відомі, або їх можна визначити. Це дає можливість об'єктивно підійти до розв'язку задачі вибору вагових коефіцієнтів. Оскільки моделі, які дають меші суми квадратів похибок прогнозів, генерують якініші прогнози, то логічно прийняти цю міру за основу для визначення вагових коефіцієнтів. Позначимо суму квадратів похибок прогнозування (для історичного прогнозу) через

$$sse = \sum_{k=1}^N e^2(k).$$

Тепер можна записати вирази для вагових коефіцієнтів окремих прогнозів:

$$w_1 = \frac{1/sse_1}{1/sse_1 + 1/sse_2}, \quad (5.10.5)$$

$$w_2 = \frac{1/sse_2}{1/sse_1 + 1/sse_2}, \quad (5.10.6)$$

де sse_1, sse_2 – суми квадратів похибок для кожного з методів, що використовуються в даному випадку.

Наприклад, нехай $sse_1 = 100, sse_2 = 16$:

$$w_1 = \frac{1/100}{1/100 + 1/16} = \frac{0.01}{0.01 + 0.0625} = 0.1379,$$

$$w_2 = \frac{1/16}{1/100 + 1/16} = \frac{0.0625}{0.01 + 0.0625} = 0.8621.$$

Таким чином, ми (об'єктивно) присвоїли більший ваговий коефіцієнт точнішому методу прогнозування. При цьому $\sum w_i = 1$, що необхідно для досягнення коректності застосування методу.

Запитання і вправи до п'ятого розділу

1. Наведіть приклади практичних ділових задач, при розв'язанні яких необхідно користуватись прогнозами? Яка різниця між прогнозом і передбаченням?
2. Наведіть приклади задач прогнозування, які виникають при прийнятті особистих рішень?
3. На які частини і в якій пропорції доцільно ділити вибірку даних при побудові моделі і оцінюванні прогнозів?
4. Що означає термін „історичний прогноз”? Чи є необгнозість у його виконанні, обґрунтуйте відповідь. Які існують інші види прогнозів?
5. Чи існує різниця між якістю прогнозу, оціненого без знаходження розв'язку і на основі розв'язку? У чому полягають переваги функції прогнозування, знайденої на основі розв'язку?
6. Побудуйте функції прогнозування на основі розв'язків таких рівнянь:

$$\begin{aligned}y(k) &= 0,35 - 0,75 y(k-1) + 0,25 y(k-2) + \varepsilon(k); \\y(k) &= a_0 + 1,5 y(k-1) - 0,5 y(k-2) + \varepsilon(k); \\y(k) &= a_0 + y(k-2) + \varepsilon(k); \\y(k) &= a_0 + 2 y(k-1) - y(k-2) + \varepsilon(k); \\y(k) &= a_0 + y(k-1) + 0,25 y(k-2) - 0,25 y(k-3) + \varepsilon(k).\end{aligned}$$

7. Запишіть модель (поліном) қубічного тренду з АР(2), поясніть які прогнози можна оцінити за допомогою цієї моделі?
8. Які асимптотичні статистичні характеристики має похибка прогнозу, оціненого за допомогою рівнянь АРКС? Поясніть на прикладі.
9. Поясніть суть методу прогнозування з мінімальною дисперсією? Наведіть приклад його доцільного застосування.
10. Порівняйте дисперсію похибки прогнозу, оціненого на основі розв'язку рівняння, з дисперсією прогнозу, отриманого за методом мінімальної дисперсії.

11. Чому для оцінювання якості прогнозу необхідно використовувати множину критеріїв? Поясніть суть середньої і середньої абсолютної похибки в процентах?
12. Поясніть суть коефіцієнта Тейла? Яке значення приймає цей коефіцієнт у випадку, коли модель має високі характеристики якості прогнозу?
13. Які складові має коефіцієнт Тейла? Як можна інтерпретувати їх значення?

Розділ 6

ЗАСТОСУВАННЯ ОПЕРАТОРІВ ЗСУВУ ПРИ ПОБУДОВІ МОДЕЛЕЙ

6.1 Властивості операторів зсуву

Оператор зсуву – це лінійний оператор, який виконує операцію зсуву (затримки або лагу) в часі значень часових рядів. Його часто використовують як елемент описання дискретних моделей об'єктів керування у техніці, на виробництві і т. ін. В спеціальній літературі використовують, в основному, два позначення цього оператора: z^{-1} та L , тобто

$$z^{-1}y(k) = L y(k) = y(k - 1),$$

або для зсуву на i періодів дискретизації вимірів

$$z^{-i} y(k) = L^i y(k) = y(k - i). \quad (6.1.1)$$

Оскільки є деяка різниця в написанні операторів z^{-1} і L , то інколи виникають непорозуміння щодо їх коректного застосування. Тому нижче буде показано як застосовувати обидва види оператора зсуву.

Вище ми вже використовували оператори зсуву при виведенні функції прогнозування методом мінімізації дисперсії, а в даному розділі покажемо як за допомогою цього оператора можна компактно записати рівняння авторегресії або авторегресії з ковзним середнім та знайти його частковий розв'язок.

Оператори зсуву мають такі властивості:

1. Оператор не діє на константу: $Lc = c$.
2. Дистрибутивний закон для додавання операторів:

$$(L^i + L^j)y(k) = L^i y(k) + L^j y(k) = y(k-i) + y(k-j). \quad (6.1.2)$$

3. Асоціативний закон для множення операторів:

$$L^i L^j y(k) = L^i (L^j y(k)) = L^i y(k-j) = y(k-i-j). \quad (6.1.3)$$

4. Від'ємна степінь оператора L та додатна степінь оператора z означають зсув у часі вперед:

$$L^{-i} y(k) = z^i y(k) = y(k+i). \quad (6.1.4)$$

Таку дію оператора можна пояснити наступним чином, нехай $j = -i$; тепер

$$L^j y(k) = y(k-j) = y(k+i).$$

5. Для деякої константи $|a| < 1$ справедлива наступна тотожність:

$$(1 + aL + a^2 L^2 + a^3 L^3 + \dots) y(k) = y(k)/(1 - aL). \quad (6.1.5)$$

Доведення: Помножимо обидві частини (6.1.5) на $(1 - aL)$ і отримаємо:

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

або

$$(1 - aL + aL - a^2 L^2 + a^2 L^2 - a^3 L^3 + a^3 L^3 - a^4 L^4 + \dots) y(k) = y(k).$$

Оскільки $|a| < 1$, то $a^n L^n y(k) \rightarrow 0$ при $n \rightarrow \infty$. Таким чином, приходимо до тотожності $y(k) = y(k)$.

6. Для деякої константи $|a| > 1$ справедлива наступна тотожність:

$$\left[1 + (aL)^{-1} + (aL)^{-2} + (aL)^{-3} + \dots \right] y(k) = \frac{-aL y(k)}{1 - aL} \quad (6.1.6)$$

або

$$\frac{y(k)}{1 - aL} = -(aL)^{-1} \sum_{i=0}^{\infty} (aL)^{-i} y(k).$$

Доведення. Помножимо обидві частини (6.1.6) на $(1 - aL)$:

$$(1 - aL) \left[1 + (aL)^{-1} + (aL)^{-2} + (aL)^{-3} + \dots \right] y(k) = -aL y(k)$$

або

$$[1 - aL + (aL)^{-1} - 1 + (aL)^{-2} - (aL)^{-1} + (aL)^{-3} - (aL)^{-2} \dots] y(k) = -aL y(k).$$

Оскільки при $|a| > 1$ границя $\lim_{n \rightarrow \infty} a^{-n} L^{-n} = 0$, то остаточно отримаємо:
 $-aL y(k) = -aL y(k)$.

6.2 Приклади використання операторів зсуву

Компактний запис різницевих рівнянь

Рівняння авторегресії порядку p

$$y(k) = a_0 + a_1 y(k-1) + \dots + a_p y(k-p) + \varepsilon(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \varepsilon(k)$$

запишемо за допомогою оператора зсуву

$$(1 - a_1 L - a_2 L^2 - \dots - a_p L^p) y(k) = a_0 + \varepsilon(k)$$

або компактніше

$$A(L) y(k) = a_0 + \varepsilon(k),$$

де $A(L) = 1 - a_1 L - a_2 L^2 - \dots - a_p L^p$. Іноді використовують поліном $A(L)$ з одиничним аргументом для того щоб компактно записати алгебраїчну суму коефіцієнтів:

$$A(1) = 1 - a_1 - a_2 - \dots - a_p = 1 - \sum_{i=1}^p a_i.$$

Як ще один приклад, розглянемо рівняння АРКС(p, q)

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j \varepsilon(k-j) + \varepsilon(k)$$

за допомогою оператора зсуву у вигляді

$$A(L)y(k) = B(L)\varepsilon(k),$$

де $A(L) = 1 - a_1L - a_2L^2 - \dots - a_pL^p$,
 $B(L) = 1 + b_1L + b_2L^2 + \dots + b_qL^q$.

$A(L)$ і $B(L)$ можна трактувати як поліноми відносно операторів зсуву, які використовують, зокрема, для визначення структури різницевих рівнянь.

Використання операторів зсуву для знаходження часткових розв'язків різницевих рівнянь

Приклад 6.1. Розглянемо як приклад стохастичний процес авторегресії першого порядку

$$y(k) = a_0 + a_1y(k-1) + \varepsilon(k), \quad |a_1| < 1$$

або $y(k) = a_0 + a_1Ly(k) + \varepsilon(k)$. Запишемо вираз для $y(k)$:

$$y(k) = \frac{a_0 + \varepsilon(k)}{1 - a_1L} = \frac{a_0}{1 - a_1} + \frac{\varepsilon(k)}{1 - a_1}. \quad (6.2.1)$$

Згідно із властивістю (6.1.5), останній член в рівнянні (6.2.1) можна записати у вигляді розкладу в нескінченний ряд:

$$\frac{\varepsilon(k)}{1 - a_1L} = \varepsilon(k) + a_1\varepsilon(k-1) + a_1^2\varepsilon(k-2) + \dots .$$

Таким чином, можна остаточно записати частковий розв'язок рівняння AP(1):

$$y^p(k) = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i).$$

А повний розв'язок із врахуванням однорідного, як і отриманий раніше іншими методами, має вигляд:

$$y(k) = \frac{a_0}{1 - a_1} + Aa_1^k + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i).$$

Приклад 6.2. Знайдемо за допомогою операторів зсуву частковий розв'язок рівняння АРКС(1,1), тобто,

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k) + b_1 \varepsilon(k-1), \quad |a_1| < 1. \quad (6.2.2)$$

Скориставшись дистрибутивним законом, отримаємо

$$(1 - a_1 L) y(k) = a_0 + (1 + b_1 L) \varepsilon(k),$$

а звідси

$$\begin{aligned} y(k) &= \frac{a_0 + (1 + b_1 L) \varepsilon(k)}{1 - a_1 L} = \\ &= \frac{a_0}{1 - a_1} + \frac{\varepsilon(k)}{1 - a_1 L} + b_1 \frac{\varepsilon(k-1)}{1 - a_1 L}. \end{aligned} \quad (6.2.3)$$

У відповідністю із властивістю (6.1.5) при $|a_1| < 1$ для другого члена в правій частині останнього рівняння можна записати

$$\begin{aligned} \frac{\varepsilon(k)}{1 - a_1 L} &= \varepsilon(k) + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + \dots = \\ &= \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) = \sum_{i=1}^{\infty} a_1^i \varepsilon(k-i) + \varepsilon(k). \end{aligned}$$

Відповідно, останній член (6.2.3) представимо у вигляді:

$$\begin{aligned} b_1 \frac{\varepsilon(k-1)}{1 - a_1 L} &= b_1 \frac{\varepsilon(k)L}{1 - a_1 L} = b_1 L \frac{\varepsilon(k)}{1 - a_1 L} = b_1 L \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i) = \\ &= b_1 \sum_{i=0}^{\infty} a_1^i \varepsilon(k-1-i). \end{aligned}$$

Частковий розв'язок рівняння (6.2.2) запишемо як суму отриманих складових

$$\begin{aligned} y^p(k) &= \frac{a_0}{1 - a_1} + \sum_{i=1}^{\infty} a_1^i \varepsilon(k-i) + b_1 \sum_{i=0}^{\infty} a_1^i \varepsilon(k-1-i) + \varepsilon(k) = \\ &= \frac{a_0}{1 - a_1} + a_1 \varepsilon(k-1) + a_1^2 \varepsilon(k-2) + a_1^2 \varepsilon(k-3) + \dots + b_1 \varepsilon(k-1) + b_1 a_1 \varepsilon(k-2) + \\ &\quad + b_1 a_1^2 \varepsilon(k-3) + \dots + \varepsilon(k) = \end{aligned}$$

$$= \frac{a_0}{1-a_1} + (a_1 + b_1) \sum_{i=0}^{\infty} a_1^i \varepsilon(k-1-i) + \varepsilon(k).$$

і повний розв'язок

$$y(k) = \frac{a_0}{1-a_1} + A a_1^k + (a_1 + b_1) \sum_{i=0}^{\infty} a_1^i \varepsilon(k-1-i) + \varepsilon(k).$$

Приклад 6.3. Розглянемо випадок, коли в стохастичному рівнянні авторегресії AP(1) коефіцієнт $|a_1| > 1$, тобто,

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k), \quad |a_1| > 1. \quad (6.2.4)$$

Використовуючи оператор зсуву, отримаємо

$$y(k)(1-a_1L) = a_0 + \varepsilon(k),$$

а звідси $y(k) = \frac{a_0}{1-a_1} + \frac{\varepsilon(k)}{1-a_1L}. \quad (6.2.5)$

У відповідностю із властивістю (6.1.6)

$$\frac{y(k)}{1-a_1L} = -(aL)^{-1} \sum_{i=0}^{\infty} (aL)^{-i} y(k), \quad |a_1| > 1.$$

Тепер (6.2.5) можна переписати так:

$$\begin{aligned} y(k) &= \frac{a_0}{1-a_1} - (a_1 L)^{-1} \sum_{i=0}^{\infty} (a_1 L)^{-i} \varepsilon(k) = \\ &= \frac{a_0}{1-a_1} - \frac{1}{a_1} \sum_{i=0}^{\infty} (a_1 L)^{-i} \varepsilon(k+1). \end{aligned}$$

Таким чином, повний розв'язок рівняння (6.2.4) має вигляд:

$$y(k) = \frac{a_0}{1-a_1} + A a_1^k - \frac{1}{a_1} \sum_{i=0}^{\infty} a_1^{-i} \varepsilon(k+1+i).$$

На відміну від попередніх випадків, у правій частині отриманого розв'язку ми отримали зважену суму випадкових величин з аргументом, який зростає по відношенню до моменту k . Розв'язок такого типу буде корисним для виконання аналізу впливу майбутніх значень збурень на досліджувані процеси. Наприклад, можна спрогнозувати яким буде вплив випадкових збурень (які виникнуть в моменти $k+1, k+2, \dots$) на біржові ціни в момент k . Хоча в цілому це якісний наближений аналіз, але він обґрунтovаний коректно знайденим рішенням і може бути корисним при дослідженні поведінки випадкових процесів. Це особливо важливо в тих випадках, коли збурення мають суттєвий вплив на протікання процесу.

Знаходження часткових розв'язків рівнянь вищих порядків

Для стохастичного рівняння AP(2)

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + \varepsilon(k)$$

можна записати

$$y(k)(1 - a_1 L - a_2 L^2) = a_0 + \varepsilon(k),$$

а звідси

$$y(k) = \frac{a_0 + \varepsilon(k)}{1 - a_1 L - a_2 L^2}.$$

Якщо розкласти знаменник на множники, то отримаємо таку форму рівняння:

$$y(k) = \frac{a_0 + \varepsilon(k)}{(1 - b_1 L)(1 - b_2 L)}.$$

При $|b_1| < 1$ і $|b_2| < 1$ можна застосувати властивість (6.1.5), що приводить до наступної форми запису часткового розв'язку:

$$\begin{aligned} y(k) &= \frac{a_0 + \varepsilon(k)}{(1 - b_1 L)(1 - b_2 L)} \cdot \frac{1}{(1 - b_2 L)} = \left[\frac{a_0}{1 - b_1} + \frac{\varepsilon(k)}{1 - b_1 L} \right] \cdot \frac{1}{(1 - b_2 L)} = \\ &= \frac{a_0}{(1 - b_1)(1 - b_2)} + \frac{1}{(1 - b_2 L)} \sum_{i=0}^{\infty} b_1^i \varepsilon(k-i) = \end{aligned}$$

$$= \frac{a_0}{(1-b_1)(1-b_2)} + \frac{\varepsilon(k)}{1-b_2L} + \frac{b_1\varepsilon(k-1)}{1-b_2L} + \frac{b_1^2\varepsilon(k-2)}{1-b_2L} + \dots$$

Очевидно, що можливе подальше спрощення отриманого результату, але вже на даному етапі видно, що розв'язування рівнянь вищих порядків за допомогою операторів зсуву потребує значного об'єму алгебраїчних перетворень.

Для моделі загального вигляду

$$A(L)y(k) = a_0 + B(L)\varepsilon(k)$$

формально можна записати частковий розв'язок так:

$$y^p(k) = \frac{a_0}{A(L)} + \frac{B(L)}{A(L)}\varepsilon(k).$$

Така форма запису сприяє спрощенню викладок в одно- та багатовимірному випадках.

6.3 Альтернативна форма характеристичного рівняння

Запишемо стохастичне рівняння авторегресії n -го порядку

$$y(k) = a_0 + a_1y(k-1) + a_2y(k-2) + \dots + a_ny(k-n) + \varepsilon(k)$$

за допомогою оператора зсуву:

$$(1 - a_1L - a_2L^2 - \dots - a_nL^n)y(k) = a_0 + \varepsilon(k),$$

а звідси

$$y(k) = \frac{a_0 + \varepsilon(k)}{1 - a_1L - a_2L^2 - \dots - a_nL^n}.$$

З попередніх розділів відомо, що умовою стійкості однорідного розв'язку є те, що корені характеристичного рівняння

$$\alpha^n - a_1\alpha^{n-1} - a_2\alpha^{n-2} - \dots - a_n = 0 \quad (6.3.1)$$

повинні знаходитись всередині кола одиничного радіуса на комплексній площині.

З іншого боку, характеристичне рівняння можна записати в іншій формі:

$$1 - a_1 L - a_2 L^2 - \dots - a_n L^n = 0. \quad (6.3.2)$$

Очевидно, що корені цього рівняння будуть оберненими по відношенню до коренів рівняння (6.3.1). Тому рівняння (6.3.2) називають ще *зворотним характеристичним рівнянням*.

При використанні зворотного характеристичного рівняння умовою стійкості розв'язку є те, що корені цього рівняння повинні знаходитись за межами одиничного кола на комплексній площині.

Тобто, ця умова є протилежною до умови, отриманої раніше для "прямого" характеристичного рівняння.

Знайдемо корені "прямого" і "зворотного" характеристичного рівняння для процесу АР(2) з відомими коефіцієнтами:

$$y(k) = 0,9 y(k-1) - 0,2 y(k-2).$$

A. Пряме характеристичне рівняння:

$$\alpha^2 - 0,9\alpha + 0,2 = 0$$

має корені $\alpha_1 = 0,5$ і $\alpha_2 = 0,4$.

B. Зворотне рівняння

$$0,2\alpha^2 - 0,9\alpha + 0,2 = 0$$

має корені $\alpha_1' = 2,5$ і $\alpha_2' = 2$, тобто, $\alpha_1' = 1/\alpha_2$, а $\alpha_2' = 1/\alpha_1$.

6.4 Знаходження розв'язку різницевого рівняння в „прямому” напрямі

Як вже було показано в попередньому параграфі, розв'язок різницевого рівняння може бути знайдений відносно майбутніх значень

збурень. Розглянемо ітераційний спосіб знаходження розв'язку в “прямому” напрямі на прикладі стохастичного рівняння AP(1):

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k). \quad (6.4.1)$$

Звідси

$$y(k-1) = \frac{1}{a_1} y(k) - \frac{1}{a_1} [a_0 + \varepsilon(k)]. \quad (6.4.2)$$

Будемо збільшувати час покроково на 1:

$$k := k+1: \quad y(k) = \frac{1}{a_1} y(k+1) - \frac{1}{a_1} [a_0 + \varepsilon(k+1)]; \quad (6.4.3)$$

$$k := k+1: \quad y(k+1) = \frac{1}{a_1} y(k+2) - \frac{1}{a_1} [a_0 + \varepsilon(k+2)]. \quad (6.4.4)$$

Тепер підставимо (6.4.4) в (6.4.3):

$$\begin{aligned} y(k) &= -\frac{1}{a_1} [a_0 + \varepsilon(k+1)] + \frac{1}{a_1^2} [y(k+2) - a_0 - \varepsilon(k+2)] = \\ &= -\frac{1}{a_1} [a_0 + \varepsilon(k+1)] - \frac{1}{a_1^2} [a_0 + \varepsilon(k+2)] + \frac{1}{a_1^2} y(k+2) = \\ &= -\frac{1}{a_1} [a_0 + \varepsilon(k+1)] - \frac{1}{a_1^2} [a_0 + \varepsilon(k+2)] + \frac{1}{a_1^3} [y(k+3) - a_0 - \varepsilon(k+3)]. \end{aligned}$$

Після n ітерацій отримаємо розв'язок у формі:

$$y(k) = -a_0 \sum_{i=1}^n a_1^{-i} - \sum_{i=1}^n a_1^{-i} \varepsilon(k+i) + \frac{1}{a_1^n} y(k+n). \quad (6.4.5)$$

Якщо $|a_1| < 1$, то розв'язок розбіжний при $n \rightarrow \infty$. Але якщо $|a_1| > 1$, то $a_1^{-n} \rightarrow 0$, $n \rightarrow \infty$, а вираз

$$-a_0(a_1^{-1} + a_1^{-2} + a_1^{-3} + \dots) \rightarrow -\frac{a_0 a_1^{-1}}{1 - a_1^{-1}} = \frac{a_0}{1 - a_1}.$$

Таким чином, розв'язок в прямому напрямі при $k \rightarrow \infty$ має вигляд:

$$y(k) = \frac{a_0}{1-a_1} - \sum_{i=1}^{\infty} a_1^{-i} \varepsilon(k+i). \quad (6.4.8)$$

Даний розв'язок відображає вплив майбутніх значень збурень на поточне значення $y(k)$. Він збігається при $|a_1| > 1$. Таким розв'язком можна скористатись, наприклад, для визначення впливу майбутніх значень грошової маси на інфляцію, а також для визначення впливу майбутніх збурень на ціну акцій на біржі.

Для порівняння наведемо отриманий раніше розв'язок, який враховує минулі значення збурень:

$$y(k) = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon(k-i). \quad (6.4.9)$$

Використання обох видів розв'язків (6.4.8) і (6.4.9) дає можливість дослідити вплив минулих і майбутніх збурень на поточне значення залежної змінної.

Запитання і вправи до шостого розділу

1. Як діє оператора зсуву на константу і на змінну?
2. Які позначення використовують для оператора зсуву при аналізі фінансово-економічних процесів і технічних систем? Поясніть чому вони відрізняються?
3. Яка мета застосування операторів зсуву?
4. Наведіть приклади застосування властивостей 5 і 6, відмінні від тих, що наведені в тексті.
5. Поясніть термін „зворотне” характеристичне рівняння? Наскільки зручним є його використання?
6. Запишіть рівняння авторегресії другого порядку з відомими коефіцієнтами – виберіть значення коефіцієнтів такими, щоб однорідний розв'язок був збіжним. Знайдіть корені «прямого» і «зворотного» характеристичних рівнянь, порівняйте їх значення.
7. Сформулюйте умови збіжності однорідного розв'язку рівняння авторегресії за допомогою коренів прямого і зворотного характеристичного рівняння?
8. Чи можна знайти неоднорідні розв'язки рівнянь високих порядків за допомогою оператора зсуву?

9. У яких випадках корисно застосовувати розв'язок різницевих рівнянь у „прямому” напрямі? Наведіть приклад конкретного процесу, при моделюванні якого необхідно скористатись таким розв'язком.
10. Наведіть приклад використання зворотного характеристичного рівняння для аналізу збіжності однорідного розв'язку?
11. Чи можна користуватись розв'язком у „прямому” напрямі для оцінювання прогнозів?
12. Чим відрізняється розв'язок різницевого рівняння в прямому від розв'язку у зворотному напрямі? Наведіть приклад розв'язку РР в прямому напрямі?
13. Знайдіть розв'язок у „прямому” напрямі для моделі процесу АРКС(1,1):

$$y(k) = a_0 + a_1 y(k-1) + \varepsilon(k) + b_1 \varepsilon(k-1).$$

14. Запишіть за допомогою оператора зсуву двовимірну модель АРКС(2,2) із врахуванням у ковзному середньому вхідних випадкових величин: $x_1(k)$ і $x_2(k)$.
15. Чи можна застосовувати різницеві рівняння різних порядків для моделювання динаміки процесів довільної природи, не користуючись операторами зсуву?

Розділ 7

ОПТИМАЛЬНЕ ОЦІНЮВАННЯ І ПРОГНОЗУВАННЯ СТАНУ ЗА ДОПОМОГОЮ ФІЛЬТРА КАЛМАНА

7.1 Принцип рекурсивного оцінювання

Алгоритми оцінювання стану динамічних систем ґрунтуються на рекурсивних обчислювальних схемах, які дуже зручні для комп’ютерної реалізації та практичного застосування. Без них неможливо обійтись при створенні систем керування та прогнозування, які працюють в реальному часі. Нижче будуть розглянуті недоліки та переваги, які мають рекурсивні алгоритми у порівнянні із звичними нерекурсивними схемами обчислень. Загалом, мета цього розділу полягає в тому, щоб надати необхідну початкову інформацію про рекурсивне оцінювання вимірів на фоні шумових складових, зокрема стосовно можливостей застосування популярного фільтра Калмана, його алгоритмізацію, переваги та недоліки.

Для ілюстрації спочатку розглянемо загальний принцип рекурсивного оцінювання даних, який в подальшому буде використано при побудові алгоритмів оптимального оцінювання (фільтрації). Принцип рекурсивного оцінювання можна легко пояснити на прикладі обчислення поточного вибіркового середнього значення часового ряду статистичних або експериментальних даних:

$$\bar{x}(k) = \frac{1}{k} \sum_{i=1}^k x(i), \quad (7.1.1)$$

де $\bar{x}(k)$ – оцінка середнього значення послідовності $x(k)$. Формулу для середнього значення представимо у вигляді:

$$\begin{aligned}\bar{x}(k) &= \frac{k-1}{k-1} \cdot \frac{1}{k} \left[\sum_{i=1}^{k-1} \bar{x}(i) + x(k) \right] = \frac{k-1}{k} \cdot \frac{1}{k-1} \left[\sum_{i=1}^{k-1} x(i) + x(k) \right] = \\ &= \left(1 - \frac{1}{k}\right) \cdot \left[\frac{1}{k-1} \sum_{i=1}^{k-1} x(i) + \frac{1}{k-1} x(k) \right] = \left(1 - \frac{1}{k}\right) \left[\bar{x}(k-1) + \frac{1}{k-1} x(k) \right] = \\ &= \bar{x}(k-1) - \frac{1}{k} \bar{x}(k-1) + \frac{1}{k} x(k).\end{aligned}$$

Тобто

$$\bar{x}(k) = \bar{x}(k-1) + \frac{1}{k} \cdot [x(k) - \bar{x}(k-1)]. \quad (7.1.2)$$

Це рівняння має рекурсивну форму, тому що поточна оцінка $\bar{x}(k)$ ґрунтуються на значенні попередньої оцінки $\bar{x}(k-1)$, до якої додається корекція (зважена нев'язка). Другий член в правій частині останньої формулі призначений для корекції оцінки середнього в момент k . Таким чином, ми отримали рівняння у формі, яка подібна до основного рівняння фільтра Калмана. Таке рівняння можна вважати рівнянням оцінювання поточного середнього у формі фільтра Калмана з нестационарним коефіцієнтом, який дорівнює $1/k$. Такою ж формою ми скористаємося у наступному параграфі з метою отримання рівнянь оптимального фільтра.

Перевага рекурсивного рівняння у порівнянні з нерекурсивною формою (7.1.1) полягає в тому, що рекурсивне рівняння не потребує запам'ятовування всієї вибірки значень $x(k)$, а середнє значення оцінки може знаходитись в будь-якому інтервалі. Воно ґрунтуються на оцінці, отриманій в попередній момент часу, та на поточних вимірах. Очевидно, що із збільшенням значення k поточне середнє прямує до константи, а тому вираз (7.1.2) можна використовувати для оцінювання середнього тільки на коротких часових інтервалах. Оцінювання і прогнозування середнього на довгих часових інтервалах потребує застосування моделей поточного середнього, наприклад, поліномів від часу або інших детермінованих функцій.

7.2 Дискретний фільтр Калмана для вільної динамічної системи

Дискретний фільтр Калмана описується множиною рекурсивних матричних рівнянь, які в лінійному випадку відносно легко розв'язуються на комп'ютері. В основі фільтра лежить розв'язок рівняння Ріккаті, який дає можливість визначити оптимальне значення матричного коефіцієнта фільтра. В даній главі ми отримаємо рівняння алгоритму оптимальної фільтрації для лінійної динамічної системи, які можна використовувати для обчислювання оптимальних оцінок стану процесу, а також для короткострокового прогнозування.

Постановка задачі оптимальної фільтрації. Нестаціонарна лінійна система описується в дискретному часі рівняннями з непостійними в часі коефіцієнтами (непостійність означає залежність від часу k):

$$\mathbf{x}(k) = \mathbf{A}(k, k-1)\mathbf{x}(k-1) + \mathbf{B}(k, k-1)\mathbf{u}(k-1) + \mathbf{w}(k-1), \quad (7.2.1)$$

де $\mathbf{x}(k)$ – n -вимірний вектор станів системи; $\mathbf{u}(k-1)$ – m -вимірний вектор детермінованих вхідних величин (сигнали керування); $\mathbf{w}(k-1)$ – n -вимірний вектор випадкових зовнішніх збурень; $\mathbf{A}(k, k-1)$ – $(n \times n)$ матриця динаміки системи (вона містить коефіцієнти, що характеризують динаміку, тобто швидкість зміни станів у часі); $\mathbf{B}(k, k-1)$ – $(n \times m)$ матриця коефіцієнтів керування. Подвійний часовий аргумент у вигляді $(k, k-1)$ означає, що величина з цим аргументом використовується в момент k , але її значення ґрунтуються на попередніх даних, які відомі на момент $k-1$ включно. Далі будемо записувати для простоти матриці \mathbf{A} і \mathbf{B} з одним аргументом, тобто $\mathbf{A}(k)$ та $\mathbf{B}(k)$. Очевидно, що стаціонарна система описується матрицями з постійними коефіцієнтами, які записують просто \mathbf{A} і \mathbf{B} . Оскільки матриця \mathbf{A} зв'язує поточний стан із попереднім, то її називають ще переходною матрицею станів. Нагадаємо, що дискретний час k зв'язаний з неперервним часом t періодом дискретизації вимірів T_s : $t = kT_s$.

В класичній постановці задачі оптимальної фільтрації послідовність зовнішніх збурень $\mathbf{w}(k)$ задовольняє властивостям білого гаусового шуму з нульовим середнім значенням і коваріаційною матрицею \mathbf{Q} , тобто статистики шуму мають вигляд:

$$\begin{aligned} E[\mathbf{w}(k)] &= 0, \quad \forall k, \\ E[\mathbf{w}(k)\mathbf{w}^T(j)] &= \mathbf{Q}(k)\delta_{kj}, \end{aligned} \quad (7.2.2)$$

де δ_{kj} – дельта-функція Кронекера, що визначається так:

$$\delta_{kj} = \begin{cases} 0 & \text{для } k \neq j \\ 1 & \text{для } k = j \end{cases}; \quad \mathbf{Q}(k) \text{ – додатно визначена коваріаційна матриця}$$

зовнішніх збурень стану розмірності $(n \times n)$. Діагональні елементи матриці представляють собою дисперсії компонент вектора збурень $\mathbf{w}(k)$.

Початковим станом системи \mathbf{x}_0 будемо вважати випадкові змінні з відомими статистиками:

$$E[\mathbf{x}_0] = \bar{\mathbf{x}}_0; \quad E[\mathbf{x}_0 \mathbf{x}_0^T] = \mathbf{M}; \quad E[\mathbf{w}(k) \mathbf{x}_0^T] = 0, \quad \forall k. \quad (7.2.3)$$

Нехай вектор вимірів $\mathbf{z}(k)$ вихідних змінних доступний в будь-який момент часу t_k , а його компоненти лінійно зв'язані з вектором стану і на них впливає шум вимірів, тобто

$$\mathbf{z}(k) = \mathbf{H}(k) \mathbf{x}(k) + \mathbf{v}(k), \quad (7.2.4)$$

де $\mathbf{H}(k)$ – матриця спостережень вимірності $(r \times n)$, $\mathbf{v}(k)$ – r -вимірний вектор випадкових величин шуму вимірів з відомими статистиками:

$$E[\mathbf{v}(k)] = 0, \quad E[\mathbf{v}(k) \mathbf{v}^T(j)] = \mathbf{R}(k) \delta_{kj},$$

де $\mathbf{R}(k)$ – додатно визначена коваріаційна матриця шумів вимірів вимірності $(r \times r)$, діагональні елементи якої є дисперсіями адитивного шуму в кожному каналі вимірів. Шум вимірів також задовольняє властивостям білого гаусового шуму. Він вважається некорельованим із зовнішнім збуренням $\mathbf{w}(k)$ і початковим станом системи, тобто

$$E[\mathbf{v}(k) \mathbf{w}^T(j)] = 0 \quad \forall k, j;$$

$$E[\mathbf{v}(k) \mathbf{x}_0^T] = 0 \quad \forall k.$$

Для визначеності вище системи з вектором стану $\mathbf{x}(k)$ необхідно знайти оцінку стану $\hat{\mathbf{x}}(k)$ в момент t_k як лінійну комбінацію оцінки $\hat{\mathbf{x}}(k-1)$ в момент t_{k-1} і самого останнього виміру (статистичних даних) $\mathbf{z}(k)$.

Оцінка $\hat{\mathbf{x}}(k)$ повинна обчислюватися як найкраща за мінімумом середнього значення суми квадратів оцінок похибок. Іншими словами, оцінка повинна бути такою, щоб

$$E[(\hat{\mathbf{x}}(k) - \mathbf{x}(k))^T (\hat{\mathbf{x}}(k) - \mathbf{x}(k))] = \min_K , \quad (7.2.5)$$

де $\mathbf{x}(k)$ – точне значення вектора стану, яке може бути обчислене за допомогою детермінованої складової математичної моделі процесу; \mathbf{K} – оптимальний матричний коефіцієнт фільтра, який необхідно обчислити в результаті розв'язання оптимізаційної задачі.

Концепція вільної динамічної системи

Вільна динамічна система – це система, простіша ніж повна система (7.2.1); в ній відсутні зовнішні вхідні впливи. Така система описується однорідним лінійним різницевим рівнянням:

$$\mathbf{x}(k) = \mathbf{A}(k) \mathbf{x}(k-1), \quad (7.2.6)$$

Фактично це однорідне рівняння для матричного рівняння стану.

Воно має простішу структуру ніж рівняння (7.2.1) і за його допомогою можна значно легше вивести рівняння фільтра Калмана. Тобто спочатку отримаємо рівняння фільтра для вільної динамічної системи (7.2.6), а потім узагальними їх на випадок системи, збуреної детермінованими та випадковими вхідними впливами.

Формування рівняння фільтрації

Виберемо структуру рівняння для оцінювання лінійної системи. Прогноз оцінки стану в момент часу k на основі попередньої оцінки можна знайти за допомогою матриці динаміки:

$$\hat{\mathbf{x}}(k) = \mathbf{A}(k) \hat{\mathbf{x}}(k-1). \quad (7.2.7)$$

Рівняння (7.2.7) показує, що прогноз оцінки в момент k визначається як результат множення оцінки вектора стану в момент $(k-1)$ на матрицю $\mathbf{A}(k)$, тобто прогноз оцінки визначається тільки динамікою системи.

Для корегування прогнозу оцінки необхідно використати вимір $\mathbf{z}(k)$. З рівняння (7.2.4) видно, що очікуване значення $\hat{\mathbf{z}}(k)$ вихідного вектора вимірів $\mathbf{z}(k)$ в момент t_k буде або

$$\hat{\mathbf{z}}(k) = \mathbf{H}(k) \mathbf{A}(k) \hat{\mathbf{x}}(k-1).$$

Похибка прогнозованого значення виміру може бути знайдена як різниця між виміром та очікуваною величиною, тобто

$$\mathbf{e}_p(k) = \mathbf{v}(k) = \mathbf{z}(k) - \mathbf{H}(k)\mathbf{A}(k)\hat{\mathbf{x}}(k-1).$$

Цю величину називають ще *нєв'язкою* або *інновацією*.

Таким чином, оцінка $\hat{\mathbf{x}}(k)$ є лінійною комбінацією оцінки $\hat{\mathbf{x}}(k-1)$ в попередній момент і значення вихідного вектора $\mathbf{z}(k)$, що вимірюється, в момент t_k :

$$\hat{\mathbf{x}}(k) = \mathbf{A}(k)\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}(k)\mathbf{A}(k)\hat{\mathbf{x}}(k-1)]. \quad (7.2.8)$$

В цьому рівнянні на вибір матриці $\mathbf{K}(k)$ (матричний коефіцієнт фільтра) впливають кілька факторів, які вибираються таким чином, щоб мінімізувати математичне сподівання суми квадратів похибок оцінок вектора стану, тобто задовольняють критерію (7.2.5). Матрицю $\mathbf{K}(k)$ називають матричним оптимальним фільтром Калмана.

Похибка оцінки вектора стану за визначенням дорівнює:

$$\begin{aligned} \mathbf{e}(k) &= \hat{\mathbf{x}}(k) - \mathbf{x}(k), \\ E[(\hat{\mathbf{x}}(k) - \mathbf{x}(k))^T(\hat{\mathbf{x}}(k) - \mathbf{x}(k))] &= E[\mathbf{e}^T(k)\mathbf{e}(k)], \\ E[\mathbf{e}^T(k)\mathbf{e}(k)] &= \text{trace } E[\mathbf{e}(k)\mathbf{e}^T(k)] = \text{trace } [\mathbf{P}(k)], \end{aligned}$$

де $\mathbf{P}(k)$ – коваріаційна матриця похибок оцінок вектора стану, яка асоціюється з оцінкою стану, визначеною рівнянням (7.2.8). Ця матриця необхідна для того щоб знайти в подальшому вираз для матричного оптимального коефіцієнта фільтра.

Для того щоб отримати вираз для матриці $\mathbf{P}(k)$, сформуємо спочатку вираз для вектора похибок оцінок $\mathbf{e}(k)$:

$$\begin{aligned} \mathbf{e}(k) &= [\mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)(\mathbf{z}(k) - \mathbf{H}\hat{\mathbf{x}}(k-1))] - \mathbf{A}(k-1)\mathbf{x}(k-1) = \\ &= \mathbf{A}\mathbf{e}(k-1) - \mathbf{K}(k)\mathbf{H}\mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)\mathbf{z}(k) = \\ &= \mathbf{A}\mathbf{e}(k-1) - \mathbf{K}(k)\mathbf{H}\mathbf{A}[\mathbf{e}(k-1) + \mathbf{x}(k-1)] + \mathbf{K}(k)[\mathbf{H}\mathbf{x}(k) + \mathbf{v}(k)] = \\ &= [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{A}\mathbf{e}(k-1) - \mathbf{K}(k)\mathbf{H}\mathbf{A}\mathbf{x}(k-1) + \\ &\quad + \mathbf{K}(k)\mathbf{H}\mathbf{A}\mathbf{x}(k-1) + \mathbf{K}(k)\mathbf{v}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{A}\mathbf{e}(k-1) + \mathbf{K}(k)\mathbf{v}(k). \end{aligned} \quad (7.2.9)$$

Тепер підставимо (7.2.9) в вираз для коваріаційної матриці $\mathbf{P}(k) = E[\mathbf{e}(k)\mathbf{e}^T(k)]$ і отримаємо:

$$\begin{aligned}\mathbf{P}(k) = & [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{A} E[\mathbf{e}(k-1)\mathbf{e}^T(k-1)]\mathbf{A}^T [\mathbf{I} - \mathbf{H}^T\mathbf{K}^T(k)] + \\ & + \mathbf{K}(k) E[\mathbf{v}(k)\mathbf{e}^T(k-1)]\mathbf{A}^T [\mathbf{I} - \mathbf{H}^T(k)\mathbf{K}^T(k)] + \\ & + [\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)]\mathbf{A} E[\mathbf{e}(k-1)\mathbf{v}^T(k)]\mathbf{K}^T(k) + \mathbf{K}(k) E[\mathbf{v}(k)\mathbf{v}^T(k)]\mathbf{K}^T(k)\end{aligned}.$$

За визначенням,

$$\begin{aligned}E[\mathbf{e}(k-1)\mathbf{e}^T(k-1)] &= \mathbf{P}(k-1), \\ E[\mathbf{v}(k)\mathbf{v}^T(k)] &= \mathbf{R}(k).\end{aligned}\quad (7.2.10)$$

Беручи до уваги, що $E[\mathbf{v}(k)\mathbf{v}^T(k-1)] = 0$ та $E[\mathbf{v}(k)\hat{\mathbf{x}}_0^T] = 0$, отримаємо

$$E[\mathbf{v}(k)\mathbf{e}^T(k-1)] = 0,$$

і

$$E[\mathbf{e}(k-1)\mathbf{v}^T(k)] = 0.$$

Тобто значення шуму $\mathbf{v}(k)$ в момент t_k некорельоване з вектором стану $\mathbf{x}(k-1)$ в момент t_{k-1} . Вектор $\hat{\mathbf{x}}(k-1)$, який залежить від $\mathbf{v}(k-1)$, не корельований з $\mathbf{v}(k)$, тому що немає кореляції між $\mathbf{v}(k-1)$ та $\mathbf{v}(k)$. Звідси отримуємо, що вектор $\mathbf{e}(k-1) = \hat{\mathbf{x}}(k-1) - \mathbf{x}(k-1)$ є некорельованим з $\mathbf{v}(k)$.

Ап'riорна коваріаційна матриця похибок оцінок вектора стану

Беручи до уваги (7.2.10) і останній вираз, запишемо вираз для $\mathbf{P}(k)$ у формі

$$\mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}'(k)[\mathbf{I} - \mathbf{K}(k)\mathbf{H}]^T + \mathbf{K}(k)\mathbf{R}(k)\mathbf{K}^T(k), \quad (7.2.11)$$

де

$$\mathbf{P}'(k) = \mathbf{A}\mathbf{P}(k-1)\mathbf{A}^T.$$

Матрицю $\mathbf{P}'(k)$ називають *ап'riорною* коваріаційною матрицею похибок оцінок вектора стану, тому що вона характеризує оцінку вектора стану до надходження вимірювань $\mathbf{z}(k)$. Тому матрицю $\mathbf{P}(k)$ називають *апостеріорною* коваріаційною матрицею похибок оцінок.

Після деяких алгебраїчних маніпуляцій з рівнянням (7.2.11) його можна представити у вигляді

$$\begin{aligned}
\mathbf{P}(k) &= [\mathbf{P}'(k) - \mathbf{K}(k)\mathbf{H}\mathbf{P}'(k)][\mathbf{I} - \mathbf{H}^T\mathbf{K}^T(k)] + \mathbf{K}(k)\mathbf{R}(k)\mathbf{K}^T(k) = \\
&= \mathbf{P}'(k) - \mathbf{K}(k)\mathbf{H}\mathbf{P}'(k) - \mathbf{P}'(k)\mathbf{H}^T\mathbf{K}^T(k) + \\
&\quad + \mathbf{K}(k)[\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}(k)]\mathbf{K}^T(k).
\end{aligned} \tag{7.2.12}$$

Рівняння такого типу називають рівнянням Ріккаті.

Матриця $[\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}]$ – симетрична, оскільки

$$[\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}] = [\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}]^T,$$

та невід'ємно визначена. Таку матрицю можна представити у вигляді:

$$\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R} = \mathbf{S}(k)\mathbf{S}^T(k), \tag{7.2.13}$$

де матриця $\mathbf{S}(k)$ – на даний момент невідома, але буде визначена нижче.

Таким чином, вираз для $\mathbf{P}(k)$ приймає вигляд:

$$\mathbf{P}(k) = \mathbf{P}'(k) - \mathbf{K}(k)\mathbf{H}\mathbf{P}'(k) - \mathbf{P}'(k)\mathbf{H}^T\mathbf{K}^T(k) + \mathbf{K}(k)\mathbf{S}(k)\mathbf{S}^T(k)\mathbf{K}^T(k),$$

або в спрощеному вигляді без індексів:

$$\mathbf{P} = \mathbf{P}' - \mathbf{K}\mathbf{H}\mathbf{P}' - \mathbf{P}'\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{S}\mathbf{S}^T\mathbf{K}^T. \tag{7.2.14}$$

Праву частину цього рівняння необхідно привести до такого вигляду, щоб можна було:

- визначити коефіцієнт фільтра $\mathbf{K}(k)$ із умови мінімуму сліду матриці $\mathbf{P}(k)$;
- забезпечити додатну напіввизначеність діагональних елементів матриці в правій частині рівняння.

З цією метою введемо до розгляду такий добуток:

$$(\mathbf{K}\mathbf{S} - \mathbf{G})(\mathbf{K}\mathbf{S} - \mathbf{G})^T = \mathbf{K}\mathbf{S}\mathbf{S}^T\mathbf{K}^T - \mathbf{K}\mathbf{S}\mathbf{G}^T - \mathbf{G}\mathbf{S}^T\mathbf{K}^T + \mathbf{G}\mathbf{G}^T,$$

де \mathbf{G} – невідома матриця, що існує за припущенням. Діагональні елементи матриці $(\mathbf{K}\mathbf{S} - \mathbf{G})(\mathbf{K}\mathbf{S} - \mathbf{G})^T$ повинні бути додатно напіввизначеними, тобто $diag\{(\mathbf{K}\mathbf{S} - \mathbf{G})(\mathbf{K}\mathbf{S} - \mathbf{G})^T\} \geq 0$.

Знайдемо матрицю \mathbf{G} . Можна прирівняти члени

$$\mathbf{K}\mathbf{H}\mathbf{P}' - \mathbf{P}'\mathbf{H}^T\mathbf{K}^T = \mathbf{K}\mathbf{S}\mathbf{G}^T + \mathbf{G}\mathbf{S}^T\mathbf{K}^T, \tag{7.2.15}$$

і записати тепер рівняння (7.2.14) у вигляді

$$\mathbf{P} = \mathbf{P}' + (\mathbf{K}\mathbf{S} - \mathbf{G})(\mathbf{K}\mathbf{S} - \mathbf{G})^T - \mathbf{G}\mathbf{G}^T. \quad (7.2.16)$$

Перехід від (7.2.14) до (7.2.16) дозволяє записати наступні рівності (з використанням (7.2.15)):

$$\begin{aligned} \mathbf{K}\mathbf{S}\mathbf{G}^T &= \mathbf{K}\mathbf{H}\mathbf{P}', \\ \mathbf{G}\mathbf{S}^T \mathbf{K}^T &= \mathbf{P}'\mathbf{H}^T \mathbf{K}^T \end{aligned}$$

або $\mathbf{G}\mathbf{S}^T = \mathbf{P}'\mathbf{H}^T$.

Тепер знайдемо \mathbf{G} :

$$\mathbf{G} = \mathbf{P}'\mathbf{H}^T (\mathbf{S}^T)^{-1} = \mathbf{P}'\mathbf{H}^T (\mathbf{S}^{-1})^T. \quad (7.2.17)$$

Обчислення оптимального матричного коефіцієнта фільтра

В рівнянні (7.2.16) від $\mathbf{K}(k)$ залежить тільки середній член, який є добутком матриці $(\mathbf{K}\mathbf{S} - \mathbf{G})$ на її транспоновану. Це забезпечує невід'ємність елементів головної діагоналі середнього члена в (7.2.16).

Таким чином, слід матриці $\mathbf{P}(k)$ буде мінімальним, якщо покласти середній член (7.2.16) рівним нулю, тобто

$$\mathbf{K}\mathbf{S} - \mathbf{G} = 0,$$

або $\mathbf{K}\mathbf{S} = \mathbf{G}$, а звідси $\mathbf{K}(k)$ можна визначити як

$$\mathbf{K}(k) = \mathbf{G}\mathbf{S}^{-1}. \quad (7.2.18)$$

Підставимо в (7.2.18) значення, що визначається рівнянням (7.2.17) і отримаємо:

$$\mathbf{K}(k) = \mathbf{P}'\mathbf{H}^T (\mathbf{S}^{-1})^T \mathbf{S}^{-1} = \mathbf{P}'\mathbf{H}^T (\mathbf{S}\mathbf{S}^T)^{-1},$$

а із врахуванням (7.2.13) остаточно можна записати:

$$\mathbf{K}(k) = \mathbf{P}'\mathbf{H}^T [\mathbf{H}\mathbf{P}'\mathbf{H}^T + \mathbf{R}]^{-1}.$$

Враховуючи часовий аргумент, отримаємо рівняння

$$\mathbf{K}(k) = \mathbf{P}'(k)\mathbf{H}^T(k)[\mathbf{H}(k)\mathbf{P}'(k)\mathbf{H}^T(k) + \mathbf{R}(k)]^{-1}. \quad (7.2.19)$$

Таким чином, коефіцієнт фільтра Калмана (19) знайдено із умови мінімуму математичного сподівання суми квадратів похибок оцінок вектора стану об'єкта.

Апостеріорна коваріаційна матриця похибок оцінок вектора станів

Знайдемо вираз для коваріаційної матриці, яка зв'язана з оптимальною оцінкою станів. Розглянемо знову коваріаційну матрицю, що визначається рівнянням Ріккаті (7.2.12), тобто

$$\mathbf{P}(k) = \mathbf{P}'(k) - \mathbf{K}\mathbf{H}\mathbf{P}'(k) - \mathbf{P}'(k)\mathbf{H}^T\mathbf{K}^T(k) + \mathbf{K}(k)[\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}(k)]\mathbf{K}^T(k).$$

Підставимо в це рівняння значення $K(k)$, що визначається (7.2.19), і отримаємо:

$$\begin{aligned} \mathbf{P}(k) &= \mathbf{P}'(k) - \mathbf{P}'(k)\mathbf{H}^T[\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{H}\mathbf{P}' = \mathbf{P}'(k) - \mathbf{K}(k)\mathbf{H}\mathbf{P}'(k) = \\ &= [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}'(k). \end{aligned} \quad (7.2.20)$$

Тобто апостеріорна коваріаційна матриця похибок оцінок є меншою у порівнянні з апріорною на величину $\mathbf{K}(k)\mathbf{H}\mathbf{P}'(k)$. Це пояснюється тим, що отриманий вимір $\mathbf{z}(k)$ в момент часу t_k сприяє зменшенню невизначеності оцінки вектора станів.

Результати

Матричне рівняння (7.2.8) використовується для обчислення оптимальних оцінок вектора стану в момент t_k за допомогою оптимального коефіцієнта фільтра $\mathbf{K}(k)$, який можна знайти із системи матричних рівнянь. Матричні рівняння (7.2.8), (7.2.11) і (7.2.20) утворюють рекурсивний алгоритм фільтра Калмана. Зрозуміло, що процес оцінювання може продовжуватись до нескінченності. Послідовність обчислюваних кроків показано нижче у вигляді алгоритму. Нагадаємо, що рівняння фільтра Калмана, наведені вище, побудовані для вільної динамічної системи.

Алгоритм оптимального оцінювання стану (фільтрації) вільної динамічної системи):

1. Задати початкові умови \mathbf{x}_0 для вектора стану і коваріаційної матриці похибок оцінок \mathbf{P}_0 . Присвоїти значення коваріаційним матрицям збурень стану \mathbf{Q} та похибок вимірів \mathbf{R} .
2. Обчислити матричний оптимальний коефіцієнт фільтра:

$$\mathbf{K}(k) = \mathbf{P}'(k-1) \mathbf{H}^T [\mathbf{H}\mathbf{P}(k-1)\mathbf{H}^T + \mathbf{R}]^{-1}.$$

3. Скористатись новими вимірами для обчислення поточної оцінки вектора стану:

$$\hat{\mathbf{x}}(k) = \mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}\hat{\mathbf{x}}(k-1)].$$

4. Обчислити апостеріорну коваріаційну матрицю похибок для оновлених оцінок:

$$\mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}'(k).$$

5. Обчислити апріорну (для наступної оцінки вектора стану) коваріаційну матрицю похибок оцінок:

$$\mathbf{P}'(k+1) = \mathbf{A}\mathbf{P}(k)\mathbf{A}^T.$$

і перейти на крок 2 (наступний крок обчислення рівнянь фільтра).

Розглянутий фільтр може продовжувати функціонувати нескінченно довго при надходженні нових вимірів. Після закінчення переходного процесу коефіцієнти фільтра встановлюються, як правило, на деяких постійних значеннях, які залишаються незмінними на довгих проміжках часу якщо процес стаціонарний. Якщо фільтр працює нормально (у відповідності до його теоретичних характеристик), то діагональні елементи коваріаційної матриці похибок оцінок вектора стану залишаються додатно визначеними (оскільки це дисперсії похибок оцінок) і також збігаються до невеликих усталених значень.

Якщо дисперсії похибок оцінок стану не зменшуються, а зростають з часом, то такий режим називають розбіжним. Із зростанням часу розрядна сітка комп'ютера може переповнюватись і функціонування алгоритму повністю припиняється. Очевидно, що виникнення такого режиму

функціонування фільтра необхідно попереджувати. Методи підвищення обчислювальної стійкості алгоритму фільтрації будуть розглянуті нижче в даному розділі, а в наступному параграфі отримаємо рівняння фільтра для динамічної системи, яка функціонує в умовах впливу зовнішніх детермінованих та випадкових впливів (збурень).

7.3 Дискретний фільтр Калмана для лінійної системи з детермінованими і стохастичними входами

Будемо вважати розширою математичною моделлю лінійну систему з детермінованими і стохастичними входами (збуреннями):

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{A}(k)\mathbf{x}(k-1) + \mathbf{B}(k)\mathbf{u}(k-1) + \mathbf{w}(k-1), \\ \mathbf{z}(k) &= \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k).\end{aligned}\tag{7.3.1}$$

Середній член першого рівняння системи (7.3.1) характеризує вплив детермінованого вхідного сигналу $\mathbf{u}(k)$; $\mathbf{w}(k)$ – зовнішні стохастичні збурення, які впливають на функціонування лінійної системи.

Для початку знайдемо прогнозовану оцінку вектора стану, використовуючи попередню оцінку $\hat{\mathbf{x}}(k-1)$ для моменту t_{k-1} , тобто не будемо враховувати вимірювання в момент t_k . Стохастичне збурення стану (або шум) $\mathbf{w}(k-1)$ не залежить від вектора стану в момент t_{k-1} і має нульове середнє значення. Тому вважаємо, що цей шум не буде впливати на оцінку вектора стану в момент t_k . Але $\mathbf{B}(k)\mathbf{u}(k-1) \equiv \mathbf{f}(k-1)$ – це відома векторна функція на інтервалі $[t_{k-1}, t_k]$ і, згідно з рівнянням (7.3.1), прогноз оцінки вектора стану можна представити у вигляді:

$$\hat{\mathbf{x}}'(k) = \mathbf{A}(k)\hat{\mathbf{x}}(k-1) + \mathbf{f}(k-1).\tag{7.3.2}$$

Після появи виміру вихідного вектора $\mathbf{z}(k)$ в момент t_k можна обчислити нову оцінку вектора стану за виразом:

$$\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}'(k) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}'(k)],\tag{7.3.3}$$

де $\mathbf{K}(k)$ – невідома матриця коефіцієнтів фільтра. (В подальшому будемо записувати матриці \mathbf{A}, \mathbf{H} без аргументів для спрощення запису.)

$$\mathbf{e}(k) = \hat{\mathbf{x}}(k) - \mathbf{x}(k) = \{\mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{f}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}\hat{\mathbf{x}}'(k)]\} -$$

$$\begin{aligned}
& -[\mathbf{A}\mathbf{x}(k-1) + \mathbf{f}(k-1) + \mathbf{w}(k-1)] = \\
= & [\mathbf{A}\mathbf{e}(k-1) + \mathbf{K}(k)[\mathbf{H}\mathbf{x}(k) + \mathbf{v}(k)] - \mathbf{K}(k)\mathbf{H}\hat{\mathbf{x}}'(k) - \mathbf{w}(k-1)] = \\
= & \mathbf{A}\mathbf{e}(k-1) + \mathbf{K}(k)\mathbf{H}[\mathbf{A}\mathbf{x}(k-1) + \mathbf{f}(k-1) + \mathbf{w}(k-1)] - \\
& - \mathbf{K}(k)\mathbf{H}[\mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{f}(k-1)] - \mathbf{w}(k-1) + \mathbf{K}(k)\mathbf{v}(k) = \\
= & \mathbf{A}\mathbf{e}(k-1) - \mathbf{K}(k)\mathbf{H}\mathbf{A}\mathbf{e}(k-1) + \mathbf{K}(k)\mathbf{H}\mathbf{w}(k-1) - \mathbf{w}(k-1) + \mathbf{K}(k)\mathbf{v}(k) = \\
= & [\mathbf{I} - \mathbf{K}(k)\mathbf{H}][\mathbf{A}\mathbf{e}(k-1) - \mathbf{w}(k-1)] + \mathbf{K}(k)\mathbf{v}(k).
\end{aligned} \tag{7.3.4}$$

Тепер знайдемо апостеріорну коваріаційну матрицю похибок оцінок

$$\begin{aligned}
\mathbf{P}(k) &= E[\mathbf{e}(k)\mathbf{e}^T(k)] = \\
&= E\left\{\left[[\mathbf{I} - \mathbf{K}(k)\mathbf{H}][\mathbf{A}\mathbf{e}(k-1) - \mathbf{w}(k-1)] + \mathbf{K}(k)\mathbf{v}(k)\right] \times \right. \\
&\quad \left. \left[[\mathbf{I} - \mathbf{K}(k)\mathbf{H}][\mathbf{A}\mathbf{e}(k-1) - \mathbf{w}(k-1) + \mathbf{K}(k)\mathbf{v}(k)]\right]^T\right\}.
\end{aligned} \tag{7.3.5}$$

Введемо апріорну коваріаційну матрицю похибок оцінок вектора стану $\mathbf{P}'(k)$:

$$\begin{aligned}
\mathbf{P}'(k) &= E\left\{[\mathbf{A}\mathbf{e}(k-1) - \mathbf{w}(k-1)][\mathbf{A}\mathbf{e}(k-1) - \mathbf{w}(k-1)]^T\right\} = \\
&= \mathbf{A}\mathbf{P}(k-1)\mathbf{A}^T + \mathbf{Q}(k-1).
\end{aligned} \tag{7.3.6}$$

У відповідності до (7.3.5) апостеріорна коваріаційна матриця $\mathbf{P}(k)$ визначається рівнянням

$$\mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}'(k)[\mathbf{I} - \mathbf{K}(k)\mathbf{H}]^T + \mathbf{K}(k)\mathbf{R}(k)\mathbf{K}^T(k). \tag{7.3.7}$$

Структура цього рівняння ідентична структурі рівняння для апостеріорної матриці, отриманої в попередньому параграфі для вільної динамічної системи. Це означає, що оптимальний матричний коефіцієнт фільтра $\mathbf{K}(k)$ для системи (7.3.1) визначається таким же рівнянням, що і для вільної динамічної системи. Звідси випливає, що оптимальний дискретний фільтр для лінійної системи (7.3.1) визначається рівняннями: (7.3.2), (7.3.3), (7.2.19), (7.2.20) і (7.3.6).

Очевидно, що поява випадкового збурення $\mathbf{w}(k)$ призводить до погіршення якості оцінок, про що свідчить додатковий член $\mathbf{Q}(k-1)$ в правій частині рівняння (7.3.6) для обчислення $\mathbf{P}'(k)$. Детермінований сигнал $\mathbf{f}(k-1)$ впливає на прогноз оцінки вектора стану у відповідності до рівняння (7.3.2).

Якщо середнє значення $E[\mathbf{w}(k-1)]$ випадкового збурення $\mathbf{w}(k-1)$ не дорівнює нулю і може бути виміряне (оцінене) в будь-який момент часу, то його можна досить просто врахувати в рівняннях фільтрації. В такому

випадку значення $\mu_w = E[\mathbf{w}(k-1)]$ додається до детермінованого впливу $\mathbf{f}(k-1)$.

Крім сигналу керування $\mathbf{Bu}(k-1)$, детерміноване збурення $\mathbf{f}(k-1)$ може включати в себе вплив інших вхідних детермінованих сигналів, що впливають на динамічну систему. Analogічно можна врахувати ненульове середнє шуму вимірювань $\mathbf{v}(k)$.

Система рівнянь оптимального фільтра Калмана для лінійної системи, яка функціонує в умовах впливу детермінованих та випадкових збурень, наведена нижче у вигляді алгоритму фільтрації.

Алгоритм оптимального оцінювання стану для лінійної системи з детермінованими та випадковими входами:

1. Математична модель процесу (лінійної системи):

$$\begin{aligned} \mathbf{x}(k) &= \mathbf{A}(k)\mathbf{x}(k-1) + \mathbf{B}(k)\mathbf{u}(k-1) + \mathbf{w}(k-1), \\ \mathbf{z}(k) &= \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k); \\ E[\mathbf{w}(k)] &= 0, \quad E[\mathbf{v}(k)] = 0, \quad E[\mathbf{x}_0] = \bar{\mathbf{x}}_0, \quad E[\bar{\mathbf{x}}_0 \bar{\mathbf{x}}_0^T] = \mathbf{P}_0 = \mathbf{P}'_0; \\ E[\mathbf{w}(k)\mathbf{w}^T(l)] &= \mathbf{Q}(k)\delta(k-l), \quad E[\mathbf{v}(k)\mathbf{v}^T(l)] = \mathbf{R}(k)\delta(k-l); \\ E[\mathbf{w}(k)\mathbf{v}^T(l)] &= E[\mathbf{w}(k)\mathbf{x}_0^T] = E[\mathbf{v}(k)\mathbf{x}_0^T] = 0. \end{aligned}$$

2. Матричний коефіцієнт фільтра:

$$\mathbf{K}(k) = \mathbf{P}'(k) \mathbf{H}^T [\mathbf{H}\mathbf{P}(k)\mathbf{H}^T + \mathbf{R}]^{-1}.$$

3. Рівняння фільтрації:

$$\hat{\mathbf{x}}(k) = \mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{B}\mathbf{u}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}\mathbf{A}\hat{\mathbf{x}}(k-1)].$$

4. Апостеріорна коваріаційна матриця похибок оцінок:

$$\mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}'(k-1).$$

5. Апріорна коваріаційна матриця похибок оцінок для наступного циклу:

$$\mathbf{P}'(k+1) = \mathbf{A}\mathbf{P}(k)\mathbf{A}^T + \mathbf{Q}(k).$$

6. Перейти на крок 2.

Якщо значення коваріаційних матриць \mathbf{Q} і \mathbf{R} неможливо оцінити на основі наявної інформації про процес, то необхідно застосувати алгоритм адаптивної фільтрації, який дозволяє одночасно із станом оцінювати невідомі статистичні параметри процесу. Однак, необхідно пам'ятати, що ускладнення алгоритму фільтрації потребує особливої уваги до забезпечення збіжності, оскільки він дуже чутливий до похибок моделі процесу.

Статистичне моделювання алгоритму фільтрації

Перед практичним застосуванням оптимального фільтра для обробки реальних даних, необхідно виконати його статистичне моделювання, яке полягає в багатократному прогоні алгоритму з використанням множини реалізацій випадкових процесів, що описують вплив збурень стану та шуму вимірювань. Суть статистичного моделювання полягає у виконанні наступних кроків:

1. Задати число реалізацій алгоритму фільтрації, якому відповідає число реалізацій випадкових процесів. Як правило вибирають число реалізацій $M \geq 100$. Покласти число реалізацій (циклів) алгоритму фільтрації рівним M : $C := M$.

2. Згенерувати вибірки випадкових процесів $\mathbf{w}(k)$ і $\mathbf{v}(k)$, довжина яких дорівнює довжині вибірки експериментальних (статистичних) даних, і врахувати їх у вимірюваннях (згенерованих чи реальних).

3. Застосувати алгоритм фільтрації до вибірки вимірювань довжиною N і зберегти на кожному кроці ($k = 1, \dots, N$) діагональні елементи коваріаційної матриці похибок фільтрації $\mathbf{P}(k)$ (тобто, дисперсії похибок оцінок вектора стану). При кожному наступному прогоні алгоритму фільтрації до цих дисперсій необхідно додати нові, отримані для наступного прогону. Тобто якщо вимірність вектора стану дорівнює n , то в результаті моделювання отримаємо n векторів дисперсій похибок оцінок вимірністю $[N \times 1]$.

4. Зменшити число реалізацій на 1: $C := C - 1$. Якщо $C \neq 0$, то перейти на крок 2. Інакше, перейти на крок 5. Знайти середнє значення по ансамблю реалізацій для кожного діагонального елемента похибок фільтрації $\mathbf{P}(k)$, тобто,

$$\bar{p}_{ii}(k) = \frac{1}{M} \sum_{l=1}^M p_{ii}(l), \quad k = 1, \dots, N,$$

де N – довжина вибірки даних.

5. Побудувати графіки залежностей усереднених дисперсій похибок оцінок від часу k , тобто, $p_{ii}(k), k = 1, \dots, N$. Якщо алгоритм реалізовано коректно, то усереднені дисперсії будуть експоненціально спадати до невеликих усталених значень.

Статистичне моделювання алгоритму фільтрації дає можливість дослідити характеристики його функціонування на множині реалізацій випадкових процесів і тим самим отримати гарантію його нормального функціонування в системі обробки фактичних експериментальних даних.

7.4 Причини розбіжності та заходи щодо підвищення обчислювальної стійкості оптимального фільтра

Показники нормального функціонування оптимального фільтра

Нормальне (коректне) функціонування фільтра характеризується такими показниками:

1. Фільтр Калмана відноситься до рекурсивних алгоритмів, характерною ознакою яких є наявність перехідного процесу оцінювання. На протязі перехідного процесу значення оцінок стану можуть досить сильно коливатись, а похиби оцінок можуть приймати великі значення. Для стаціонарного процесу дисперсії похибок оцінок вектора стану повинні бути спадаючими функціями часу і збігатись до деяких невеликих постійних усталених значень після закінчення перехідного процесу. Таким чином, контролювати функціонування фільтра можна за допомогою діагональних елементів матриці $\mathbf{P}(k)$.
2. Для стаціонарного процесу коефіцієнти підсилення фільтра $\mathbf{K}(k)$ повинні приймати постійні усталені значення.
3. Графіки змінних стану $\mathbf{x}(k)$ не повинні мати значних імпульсів чи викидів по закінченню перехідного процесу.
4. Послідовність інновацій $\{v(k) = \mathbf{z}(k) - \mathbf{A}\hat{\mathbf{x}}(k-1)\}$ повинна задовольняти властивостям білого шуму: $E[v(k)] = 0$ і $E[v(k)v(k-l)] = 0, l \geq 1$.

Остання властивість досить часто використовується в системах автоматичної діагностики для виявлення аномальних режимів функціонування процесів і систем в реальному часі. Як правило, вона використовується в комбінації з іншими процедурами аналізу, наприклад,

процедурами формування ймовірнісного висновку. Системи такого типу надзвичайно популярні, оскільки сприяють значному підвищенню якості діагностики та скороченню часу, необхідного для аналізу функціонування системи.

Причини розбіжності алгоритму оптимальної фільтрації

В практиці впровадження алгоритмів оптимальної фільтрації досить часто зустрічаються випадки, коли фільтр характеризується “розбіжністю”. Тобто елементи коваріаційної матриці похибок оцінок вектора стану не збігаються до невеликих усталених значень, а з часом набувають таких великих значень, які можуть перевищувати максимально допустимі для вибраного способу зберігання чисел в комп’ютері. Розбіжність алгоритму фільтрації може бути зумовлена наступними причинами:

1. *Математична модель процесу* (матриці $\mathbf{A}, \mathbf{B}, \mathbf{H}$) має недостатню ступінь адекватності динамічній системі, а тому в процесі її побудови необхідно застосувати декілька статистичних критеріїв для аналізу якості моделі. Особливу увагу необхідно звернути на якість даних, на основі яких оцінено коефіцієнти (параметри) моделі. Якщо дані недостатньо інформативні, то побудувати модель, адекватну досліджуваному режиму функціонування процесу, практично неможливо.
2. Некоректне визначення початкових умов – коваріаційної матриці похибок оцінок початкового стану системи $\mathbf{P}(0)$, тобто, використовується наближене значення $\mathbf{P}_c(0)$, де індекс “*c*” означає субоптимальність. Початкові умови особливо сильно впливають на якість функціонування фільтра в перехідному режимі.
3. *Значення коваріаційних матриць \mathbf{Q} і \mathbf{R}* невідомі або відомі із значними похибками. В такому випадку необхідно збирати додаткову інформацію про процес та/або застосовувати алгоритм адаптивної фільтрації. В окремих випадках можна спробувати “підібрати” невідомі статистичні параметри, виходячи з конкретних значень вимірюваних сигналів (статистичних даних). Якість фільтрації можна оцінювати за допомогою діагональних елементів коваріаційної матриці похибок фільтрації $\mathbf{P}(k)$.
4. *Вимірювальні (статистичні) дані* не відповідають вимогам якості (наявні значні пропуски, великі викиди, існують значні періоди з постійними значеннями).
5. *Припущення щодо розподілу та/або статистичних характеристик* збурень стану та похибок (шуму) вимірювань не відповідають дійсності.
6. Іноді для спрощення алгоритму при застосуванні його в реальному часі використовують *постійні коефіцієнти фільтра*, що також

- може призводити до великих похибок оцінок стану досліджуваного процесу.
7. Комбінації вказаних причин та помилки програмування алгоритму.

Можливі заходи щодо підвищення обчислюваної стійкості алгоритму

1. *Підвищення адекватності* математичної моделі процесу шляхом уточнення та ускладнення її структури.
2. Застосування спеціальних методів *попередньої обробки даних* з метою їх нормування, заповнення пропусків та зменшення великих викидів.
3. *Застосування алгоритмів адаптивної фільтрації* у випадку, коли невідомі параметри коваріаційних матриць \mathbf{Q} і \mathbf{R} .
4. Якщо шум вимірів відсутній, то його необхідно ввести штучно (на деякому невеликому рівні по відношенню до рівня вимірів), що сприяє покращенню обчислюваної обумовленості матриці $[\mathbf{H}\mathbf{P}(k)\mathbf{H}^T + \mathbf{R}]$, для якої обчислюється обернена при визначенні коефіцієнта фільтра.
5. Застосування спеціальних алгоритмів *обчислення обернених матриць* при визначенні матричного коефіцієнта фільтра. Наприклад, алгоритму квадратного кореня, LU – факторизації та інших. При можливості можна використовувати *алгоритм послідовної фільтрації*, який не потребує операції обернення матриці.
6. Використання *статистичного (Монте-Карло) моделювання* з метою підвищення достовірності визначення характеристик алгоритму фільтрації перед його використанням за призначенням.
7. Необхідно постійно *контролювати процес оцінювання* шляхом аналізу значень діагональних елементів матриці $\mathbf{P}(k)$. Допускається примусове зменшення значень цих елементів, якщо це не пов’язано з похибками самого алгоритму оцінювання. Крім того, *інноваційна послідовність* $\mathbf{v}(k) = \mathbf{z}(k) - \mathbf{H}\hat{\mathbf{x}}(k-1)$ повинна задовольняти вимогам: $E[\mathbf{v}(k)] = 0$ і $E[\mathbf{v}(k)\mathbf{v}^T(l)] = \mathbf{G}\delta(k-l)$, тобто, повинна бути некорельованою послідовністю з нульовим середнім.
8. Якщо конкретне застосування фільтра припускає примусове вилучення перехідного процесу оцінювання (в якому спостерігаються максимальні похибки оцінок), то необхідно використовувати *фільтри з постійними коефіцієнтами*, які визначаються наперед під час статистичного моделювання.

9. Необхідно уникати випадків одночасного оцінювання невідомих параметрів моделі та стану процесу, тому що це приводить до появи нелінійностей в моделі і розбіжності алгоритму.
10. У випадку мікропроцесорної реалізації необхідно уникати використання цілочислової арифметики.

7.5 Приклади побудови оптимального фільтра

Приклад 7.5.1. Розглянемо просту скалярну модель випадкового кроку

$$\begin{aligned}x(k) &= x(k-1) + w(k), \quad x(0) = x_0, \\z(k) &= x(k) + v(k),\end{aligned}$$

де $w(k)$ і $v(k)$ – процеси білого шуму з нульовим середнім та дисперсіями Q і R , відповідно. Оскільки процес скалярний, то модель і рівняння фільтрації суттєво спрощуються. Так, рівняння фільтрації має вигляд:

$$\hat{x}(k) = \hat{x}(k-1) + K(k)[z(k) - \hat{x}(k-1)].$$

Апріорна коваріація оцінки вектора стану і коефіцієнт фільтра

$$P'(k) = P(k-1) + Q, \quad K(k) = \frac{P(k-1) + Q}{P(k-1) + Q + R}.$$

З останнього рівняння видно, що значення коефіцієнта $K(k)$ є обернено пропорціональним дисперсії шуму вимірів. Це означає, що чим вищий шум вимірів, тим нижчим буде значення коефіцієнта фільтра і оцінка стану $\hat{x}(k)$ в значній мірі визначається попередньою оцінкою $\hat{x}(k-1)$. Внаслідок малого значення коефіцієнта $K(k)$, інновація $v(k) = z(k) - \hat{x}(k-1)$ не буде суттєво впливати на оцінки стану. У випадку малого рівня шуму вимірів (або похибок статистичних даних) значення $K(k)$ буде зростати і на поточне значення оцінки стану $x(k)$ будуть суттєво впливати інновації (тобто, поточні виміри).

Ця властивість фільтра Калмана співпадає з поведінкою людини, яка рухається вночі по неосвітленій або погано освітленій дорозі. В такій ситуації “виміри” стану дороги (впадини, випадкові предмети) дуже ненадійні внаслідок поганого освітлення. Як правило, людина рухається в

таких випадках обережно, більше покладаючись на свою власну “внутрішню прогнозуючу машину”, ніж на вимірювальну інформацію від візуальної системи. Якщо надворі зовсім темно, то можна сказати, що $R \rightarrow \infty$ і коефіцієнт “фільтра” візуальної системи людини $K(k) \rightarrow 0$. У такому випадку людина може рухатись вперед тільки завдяки своїй “прогнозуючій” машині. Фільтр Калмана, реалізований програмно чи апаратно, фактично функціонує за подібною логікою.

Вплив збурення стану $w(k)$ на коефіцієнт фільтра також можна простежити за допомогою отриманих рівнянь. Однак його вплив не такий чіткий як шуму вимірювань внаслідок того, що дисперсія збурення стану є в чисельнику і знаменнику рівняння для $K(k)$. Очевидно, що $w(k)$ безпосередньо впливає на ап'яорну коваріацію похибок оцінок $P'(k)$. Чим більше значення має дисперсія Q , тим значнішими будуть коваріації $P'(k)$ і $P(k)$. Це означає, що Q безпосередньо впливає на похибки оцінок і, таким чином, зменшує надійність оцінок $\hat{x}(k)$ в цілому.

Приклад 7.5.2. Класичний технічний приклад застосування фільтра в навігаційних системах [19].

Розглянемо прямолінійний рух літаючого апарату (літака) в повітряному просторі. На рух апарату впливає випадкове прискорення, зумовлене неоднорідністю густини атмосфери. Положення літаючого апарату (ЛА) вимірюється в дискретні моменти часу з періодом дискретизації T_s . При цьому кожний вимірювання містить шумову складову, появу якої зумовлена методичними похибками вимірювального пристрою (радара) та впливом атмосфери на корисний радіосигнал [23].

Будемо вважати, що положення ЛА вимірюється в прямокутних координатах, а точність вимірювань не залежить від часу та координат; похибки вимірювань мають нульове середнє і некорельовані між собою. Таким чином, положення $z(k)$ можна описати рівнянням:

$$z(k) = x(k) + v(k), \quad (7.5.1)$$

де $z(k) = z(kT_s)$ – вимірювання положення ЛА; $x(k) = x(kT_s)$ – дійсне положення ЛА; $v(k)$ – випадкова завада вимірювань, яка має такі статистичні характеристики:

$$E[v(k)] = 0, \quad \forall k; \quad E[v^2(k)] = \sigma_x^2 = \text{const}, \quad \forall k; \quad E[v(k)v(l)] = 0, \quad k \neq l. \quad (7.5.2)$$

Припустимо, що координати ЛА вимірюються незалежно, що дає можливість описати динаміку руху апарату по кожній координаті окремо за допомогою наступних рівнянь:

$$x(k) = x(k-1) + v(k-1)T_s + \frac{1}{2}a(k-1)T_s^2, \quad (7.5.3)$$

$$v(k) = v(k-1) + a(k-1)T_s, \quad (7.5.4)$$

де $x(k)$ – положення ЛА в поточний момент kT_s ; $v(k) = \dot{x}(k)$ – швидкість руху ЛА; $a(k)$ – випадкове прискорення ЛА, величина якого покладається незмінною на протязі одного періоду дискретизації вимірювань. Будемо також вважати, що середнє значення випадкового прискорення дорівнює нулю, а кореляція між значеннями $a(k)$ в різних періодах дискретизації відсутня, тобто

$$E[a(k)] = 0; E[a^2(k)] = \sigma_a^2 = \text{const}, \forall k; E[a(k)a(l)] = 0, k \neq l. \quad (7.5.5)$$

Для того щоб скористатись рівняннями оптимальної фільтрації, запишемо рівняння руху ЛА у формі простору станів:

$$\mathbf{x}(k) = \mathbf{F}\mathbf{x}(k-1) + \mathbf{G}a(k-1), \quad (7.5.6)$$

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + v(k). \quad (7.5.7)$$

де

$$\mathbf{F} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} T_s^2 \\ \frac{1}{2} \end{bmatrix}, \quad \mathbf{H} = [1 \ 0], \quad (7.5.8)$$

$\mathbf{z}(k) = [x(k) \ v(k)]^T$ – вектор стану ЛА.

Рівняння дискретного фільтра Калмана приймає вигляд:

$$\hat{\mathbf{x}}(k) = \mathbf{F}\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}\mathbf{F}\hat{\mathbf{x}}(k-1)]. \quad (7.5.9)$$

Величина $\hat{\mathbf{x}}(k, k-1) = \mathbf{F}\hat{\mathbf{x}}(k-1)$ – є прогнозом значення вектора стану на основі попередньої оптимальної оцінки. Із врахуванням цього позначення рівняння фільтрації (7.5.9) можна переписати у вигляді:

$$\hat{\mathbf{x}}(k, k) = \hat{\mathbf{x}}(k, k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \hat{\mathbf{x}}(k, k-1)], \quad (7.5.10)$$

де $\hat{\mathbf{x}}(k, k)$ – оцінка вектора стану в момент k на основі інформації, що є в розпорядженні на цей момент включно.

Матричний коефіцієнт підсилення фільтра $\mathbf{K}(k)$ визначається за формулою:

$$\mathbf{K}(k) = \mathbf{P}'(k)\mathbf{H}^T[\mathbf{H}\mathbf{P}'(k)\mathbf{H}^T + \mathbf{R}]^{-1}, \quad (7.5.11)$$

де $\mathbf{R} = \sigma_v^2$ – дисперсія шуму вимірів; $\mathbf{P}'(k)$ – априорна (до отримання останнього виміру $z(k)$) коваріаційна матриця похибок оцінювання, що оцінюється за наступною формулою:

$$\mathbf{P}'(k) = \mathbf{F}\mathbf{P}(k-1, k-1)\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T, \quad (7.5.12)$$

де

$$\mathbf{P}(k-1, k-1) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}'(k-1) \quad (7.5.13)$$

– апостеріорна коваріаційна матриця похибок оцінювання, отримана на попередньому кроці виконання алгоритму; $\mathbf{Q} = \sigma_a^2$ – дисперсія випадкового процесу (прискорення руху ЛА) $a(k)$.

Із врахуванням виразу (7.5.11) для обчислення коефіцієнта фільтра, (7.5.12) приймає вигляд:

$$\mathbf{P}(k-1, k-1) = \{\mathbf{I} - \mathbf{P}'\mathbf{H}^T[\mathbf{H}\mathbf{P}'\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{H}\}\mathbf{P}'. \quad (7.5.14)$$

Матриці \mathbf{P}' і \mathbf{P} мають різні значення, оскільки після врахування останнього виміру похибки оцінок зменшуються. Тобто, між цими матрицями існує таке співвідношення:

$$\|\mathbf{P}\| < \|\mathbf{P}'\|.$$

Однак, у відповідності до рівняння (7.5.12), випадкові прискорення ЛА між моментами дискретизації вимірів приводять до збільшення похибок оцінювання. Фільтр переходить в усталений стан функціонування тоді, коли збільшення похибок на протязі періоду дискретизації внаслідок випадкових прискорень буде компенсуватися їх зменшенням завдяки приходу чергового виміру.

Аналіз фільтра в усталеному стані

В усталеному стані справедливі співвідношення:

$$\mathbf{P}'(k, k) = \mathbf{P}'(k-1, k-1) = \mathbf{P}', \quad \mathbf{P}(k, k) = \mathbf{P}(k-1, k-1) = \mathbf{P}, \quad (7.5.15)$$

а тому (7.5.12), (7.5.14) приймають вигляд:

$$\mathbf{P}' = \mathbf{F}\mathbf{P}\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T, \quad (7.5.16)$$

$$\mathbf{P} = \{\mathbf{I} - \mathbf{P}'\mathbf{H}^T[\mathbf{H}\mathbf{P}'\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{H}\}\mathbf{P}', \quad (7.5.17)$$

Перепишемо (7.5.16) як

$$\mathbf{P}' - \mathbf{GQG}^T = \mathbf{FPF}^T,$$

і підставимо в праву частину останньої рівності замість \mathbf{P} (7.5.17):

$$\mathbf{P}' - \mathbf{GQG}^T = \mathbf{F}\{\mathbf{I} - \mathbf{P}'\mathbf{H}^T[\mathbf{HP}'\mathbf{H}^T + \mathbf{R}]^{-1}\mathbf{H}\}\mathbf{P}'\mathbf{F}^T. \quad (7.5.18)$$

Якщо ввести позначення

$$\mathbf{P}' = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad \text{де} \quad p_{12} = p_{21} \quad (7.5.19)$$

і врахувати значення матриць, задані (7.5.8), то (7.5.18) можна переписати у вигляді

$$\begin{bmatrix} p_{11} - \frac{\sigma_a^2 T_s^4}{4} & p_{12} - \frac{\sigma_a^2 T_s^3}{2} \\ p_{12} - \frac{\sigma_a^2 T_s^3}{2} & p_{22} - \sigma_a^2 T_s^2 \end{bmatrix} = \frac{1}{1 + \frac{p_{11}}{\sigma_x^2}} \times \quad (7.5.20)$$

$$\times \begin{bmatrix} p_{11} + 2p_{12}T_s + p_{22}T_s^2 + \frac{(p_{11}p_{22} - p_{12}^2)T_s^2}{\sigma_x^2} & p_{12} + p_{22}T_s + \frac{(p_{11}p_{22} - p_{12}^2)T_s}{\sigma_x^2} \\ p_{12} + p_{22}T_s + \frac{(p_{11}p_{22} - p_{12}^2)T_s}{\sigma_x^2} & p_{22} + \frac{p_{11}p_{22} - p_{12}^2}{\sigma_x^2} \end{bmatrix}.$$

Якщо прирівняти елементи матриці зліва (7.5.20) елементам справа, то можна записати систему алгебраїчних рівнянь відносно усталених коваріацій похибок оцінок стану ЛА:

$$\frac{p_{11}}{\sigma_x^2} = \frac{\sqrt{1+2r}(\sqrt{1+r}+1)^2}{r^2}, \quad (7.5.21)$$

$$\frac{p_{12}}{\sigma_x \sigma_a T_s} = \frac{1}{2} \frac{(\sqrt{1+2r}+1)^2}{r}, \quad (7.5.22)$$

$$\frac{p_{22}}{\sigma_a^2 T_s^2} = \frac{1}{2} (\sqrt{1+2r}+1), \quad (7.5.23)$$

де $r = 4\sigma_x/(\sigma_a T_s^2)$ – параметр, який можна розглядати як відношення “шум/сигнал”, оскільки σ_x – середня квадратична похибка датчика, що вимірює положення ЛА; $\sigma_a T_s^2/2$ – середнє квадратичне відхилення ЛА щодо прискорення, яке зумовлене його випадковим прискоренням.

За допомогою рівнянь (7.5.21)-(7.5.24) можна знайти безрозмірні відношення, які характеризують елементи апостеріорної матриці дисперсій похибок оцінювання:

$$\frac{\hat{p}_{11}}{\sigma_x^2} = \frac{\sqrt{1+2r}}{r^2} (\sqrt{1+2r}-1)^2, \quad (7.5.25)$$

$$\frac{\hat{p}_{12}}{\sigma_x \sigma_a T_s} = \frac{1}{2} \frac{(\sqrt{1+2r}-1)^2}{r}, \quad (7.5.26)$$

$$\frac{\hat{p}_{22}}{\sigma_a^2 T_s^2} = \frac{1}{2} (\sqrt{1+2r}-1). \quad (7.5.27)$$

Величини p_{11}/σ_x^2 і \hat{p}_{11}/σ_x^2 є відношеннями середнього значення квадрата похибки оцінки положення ЛА до дисперсії похибки σ_x^2 датчика положення до і після моменту вимірювання положення ЛА, відповідно.

Приклад 7.5.3. Припустимо, що лінійна система описується дискретною функцією передачі

$$G(z) = \frac{Y(z)}{U(z)} = \frac{b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} = \frac{0,6 z^{-1}}{1 - 1,6 z^{-1} + z^{-2}}, \quad (7.5.28)$$

де u і y – вхід і вихід системи, відповідно; z^{-1} – оператор затримки в часі на один крок (ми використовували його раніше при формуванні функцій прогнозування методом мінімізації дисперсії оцінки прогнозу).

Модель (7.5.28) необхідно представити у просторі станів. Методика представлення моделей такого типу у просторі станів буде розглянута в наступному розділі, а зараз наведено тільки кінцевий результат – матриці \mathbf{A} і \mathbf{B} , побудовані на основі коефіцієнтів моделі:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -a_2 & -a_1 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ a_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix};$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 1,6 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ -1,6 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,6 \\ 0,96 \end{bmatrix}.$$

Модель (7.5.28) у просторі станів має вигляд:

$$\mathbf{x}(k) = \begin{bmatrix} 0 & 1 \\ -1 & 1,6 \end{bmatrix} \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \end{bmatrix} + \begin{bmatrix} 0,6 \\ 0,96 \end{bmatrix} u(k-1), \quad (7.5.29)$$

де $x_1(k) = y(k)$; $x_2(k) = y(k-1)$ – компоненти вектора стану. Рівняння вимірів:

$$\mathbf{z}(k) = \mathbf{Hx}(k) + \mathbf{v}(k) = [1 \ 0] \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + v(k) = x_1(k) + v(k). \quad (7.5.30)$$

Для того щоб врахувати випадкові збурення, що діють на процес, в праву частину рівняння (7.5.29) необхідно ввести змінну $\mathbf{w}(k)$. Такі збурення, як правило, впливають на протікання реальних процесів, наприклад, у вигляді похибок неідеальної моделі, похибок виконання арифметичних операцій, неврахованих ефектів при моделюванні та інших факторів. Статистичні параметри процесу $\mathbf{w}(k)$ необхідно визначати апріорно, до впровадження фільтра в функціонуючу систему обробки даних. Якщо неможливо визначити точні значення статистичних параметрів, то використовують наближені, визначені в процесі статистичного моделювання фільтра.

Нехай коваріації похибок оцінок початкового стану процесу, а також коваріації збурень стану і похибок вимірів мають наступні значення:

$$\mathbf{P}_0 = \mathbf{P}(0,0) = \begin{bmatrix} 0,5 & 0,2 \\ 0,2 & 0,5 \end{bmatrix}; \quad \mathbf{Q} = \begin{bmatrix} 0,3 & 0 \\ 0 & 0,1 \end{bmatrix}; \quad \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Використовуючи ці дані, обчислимо коефіцієнт фільтра для одного кроку:

$$\begin{aligned} \mathbf{P}(1,0) &= \mathbf{A} \mathbf{P}(0,0) \mathbf{A}^T + \mathbf{Q} = \begin{bmatrix} 0,8 & 0,6 \\ 0,6 & 1,24 \end{bmatrix}; \\ \mathbf{K}(1) &= \mathbf{P}(1,0) \mathbf{H}^T [\mathbf{H} \mathbf{P}(1,0) \mathbf{H}^T + \mathbf{R}]^{-1} = \begin{bmatrix} 0,55 & 0 \\ 0 & 0,33 \end{bmatrix}; \\ \mathbf{P}(1,1) &= [\mathbf{I} - \mathbf{K}(1) \mathbf{H}] \mathbf{P}(1,0) = \begin{bmatrix} 0,36 & 0,27 \\ 0,336 & 1,042 \end{bmatrix}. \end{aligned}$$

Далі можна обчислювати оцінку вектора стану з використанням виміру, який поступає в момент часу 1, і переходити до наступного циклу реалізації рівнянь фільтра.

Приклад 7.5.4. Розглянемо неперервний стаціонарний процес Гаусса-Маркова з автокореляційною функцією

$$r_x(\tau) = \exp(-|\tau|).$$

Для одиничної дисперсії і нульового середнього спектральною функцією цього процесу є

$$S_x(s) = \frac{2}{1-s^2} = \frac{\sqrt{2}}{1+s} \cdot \frac{\sqrt{2}}{1-s},$$

а формуючий фільтр, який перетворює білий шум в процес Гаусса-Маркова, має функцію передачі

$$G(t) = \frac{x(t)}{w(t)} = \frac{\sqrt{2}}{1+s},$$

де $w(t)$ – білий гаусівський шум з одиничною дисперсією; $x(t)$ – процес Гаусса-Маркова; s – змінна Лапласа. Рівняння стану для цього процесу має вигляд:

$$\dot{x}(t) = -x(t) + \sqrt{2} w(t).$$

Для того щоб застосувати фільтр Калмана необхідно згенерувати послідовність вимірів даного процесу. Виберемо період дискретизації вимірів $T_s = 0,045$ секунди з нульовим початковим значенням при $t = 0$.

Нехай дисперсія похибок вимірів складає $R = 0,65$. Перехідна матриця станів в даному випадку є скаляром і може бути визначена за допомогою експоненти як

$$A(k) = \exp(-0,045) \approx 0,956.$$

Оскільки між виміром вихідної величини процесу і станом $x(t)$ нема різниці, то матриця (скаляр) вимірів $H(k) = 1$. Дисперсію збурення (шуму) стану

$$Q(k) = E[w^2(k)] = E\left[\int_0^{0,045} \sqrt{2} \exp(-u) w(u) du \times \int_0^{0,045} \sqrt{2} \exp(-u) w(u) du \right] =$$

$$= \int_0^{0,045} (\sqrt{2} \exp(-v))^2 dv \approx 0,165.$$

Для того щоб розпочати процес оцінювання необхідно задати початкові умови x_0 і P_0 . Оскільки процес починається в момент $t = 0$, має нульове середнє та одиничну дисперсію, то покладемо початкові умови рівними $x_0 = 0$, $P_0 = 1$ при $t = 0$. Тепер всі параметри фільтра відомі і можна розпочинати рекурсивну процедуру оцінювання.

В спеціальній літературі з проблем оцінювання параметрів і станів динамічних систем наведено багато прикладів успішного застосування оптимального фільтра. Спектр реальних задач, де він може бути використаний, є надзвичайно широким. Зокрема, це навігаційні системи, аерокосмічна техніка, обробка часових рядів в експериментальній фізиці, сейсмологічних дослідженнях, економетриці, біології та інших галузях.

7.6 Оцінювання невимірюваних компонент вектора стану за допомогою оптимального фільтра

При моделюванні динамічних систем виникають проблеми оцінювання невимірюваних компонент вектора стану. Наприклад, якщо до вектора стану входить третя похідна або вища, то вимірювати такі величини дуже складно або неможливо. Іншими прикладами можуть бути необхідність вимірювання високих температур при плавленні тугоплавких матеріалів або вимірювання відстаней до космічних об'єктів. Температури сплавів можуть бути настільки високими, що неможливо створити датчик, який може витримати таку температуру. Проблема оцінювання невимірюваних компонент вектора стану досить часто виникає при моделюванні економічних систем. Така задача виникає у випадках, коли відповідні статистичні дані не були зібрані або ці дані мають нерегулярний характер. Задачу оцінювання невимірюваних компонент можна розв'язати за допомогою оптимального фільтра.

Оцінювання невимірюваних компонент вектора стану можливе у випадку, коли коваріаційна матриця похибок оцінок має ненульові відповідні елементи, які дозволяють обчислити коефіцієнти фільтра, зв'язані з невимірюваними компонентами. Розглянемо цей механізм оцінювання. В даному випадку в рівнянні фільтрації

$$\hat{\mathbf{x}}(k) = \mathbf{A}(k)\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}\mathbf{A}(k)\hat{\mathbf{x}}(k-1)]$$

вимірність вектора вимірів $\mathbf{z}(k)$ є меншою вимірності вектора стану $\mathbf{x}(k)$, тобто, $\dim[\mathbf{z}] < \dim[\mathbf{x}]$.

Оптимальний коефіцієнт фільтра обчислюється за рівнянням:

$$\mathbf{K}(k) = \mathbf{P}'(k) \mathbf{H}^T [\mathbf{H} \mathbf{P}'(k) \mathbf{H}^T + \mathbf{R}]^{-1},$$

де $\dim[\mathbf{P}'(k)] = [n \times n]$; $\dim[\mathbf{H}^T] = [n \times r]$ за визначенням; n – вимірність вектора стану і $\dim[\mathbf{H} \mathbf{P}'(k) \mathbf{H}^T + \mathbf{R}]^{-1} = [r \times r]$. Таким чином, $\dim[\mathbf{K}(k)] = [n \times r]$.

Вектор нев'язок (інновацій) $\mathbf{v}(k)$ в рівнянні оцінювання

$$\mathbf{v}(k) = \mathbf{z}(k) - \mathbf{H} \mathbf{A} \hat{\mathbf{x}}(k)$$

має вимірність $[r \times 1]$, а вимірність добутку $\dim[\mathbf{K}(k) \mathbf{v}(k)] = [n \times 1]$. Наприклад, якщо $\dim[\mathbf{x}] = [3 \times 1]$, а $\dim[\mathbf{z}] = [2 \times 1]$, то добуток $\mathbf{K}(k) \mathbf{v}(k)$ має вигляд:

$$\mathbf{K}(k) \mathbf{v}(k) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \\ K_{31} & K_{32} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

і оцінка вектора стану визначається за рівняннями:

$$\begin{aligned} \hat{x}_1(k) &= \hat{x}_1(k, k-1) + K_{11}v_1 + K_{12}v_2, \\ \hat{x}_2(k) &= \hat{x}_2(k, k-1) + K_{21}v_1 + K_{22}v_2, \\ \hat{x}_3(k) &= \hat{x}_3(k, k-1) + K_{31}v_1 + K_{32}v_2. \end{aligned}$$

Таким чином, невимірювана третя компонента вектора стану може бути оцінена, якщо відповідні елементи матриць $\mathbf{P}'(k)$, $\mathbf{P}(k)$ і $\mathbf{K}(k)$ мають ненульові значення.

7.7 Функція прогнозування на основі оптимального фільтра

Застосування фільтра Калмана розпочалось з навігаційних систем, в яких його успішно використовують для фільтрації вимірів радіолокаційних станцій, оцінювання невимірюваних компонент сигналів, екстраполяції, розв'язання задачі комплексування інформації (об'єднання потоків даних, що поступають з різних джерел) з метою підвищення якості оцінок стану

об'єкта та деяких інших задач. Успіх такого застосування пояснюється високою ступінню адекватності математичної моделі рухомого об'єкта, стан якого оцінюється, в тому числі досить високою точністю статистичного описання випадкових збурень. Область застосування алгоритмів оптимальної фільтрації постійно розширяється. На сьогодні фільтри застосовують, в тій чи іншій формі, практично в усіх областях техніки, експериментальній фізиці, біофізиці, економіці, фінансах та інших галузях досліджень. Одним із напрямів сучасного застосування фільтра є оптимальне оцінювання та короткострокове прогнозування динаміки розвитку процесів різної природи. Така нагальна потреба застосування методів оптимальної обробки вибірок даних зумовлена тим, що в природі немає процесів (об'єктів), які функціонували б у ідеальних умовах. Фактично всі вони функціонують в умовах впливу різномірних випадкових збурень та наявності шумів вимірювань.

Структура рівняння в просторі станів є дуже зручною для використання з метою обчислення короткострокового прогнозу. Розглянемо дискретне рівняння динаміки стохастичної системи в просторі станів (детермінований керуючий вплив для простоти викладок не враховуємо):

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{w}(k), \quad (7.7.1)$$

де $\mathbf{w}(k)$ – збурюючий процес білого шуму з нульовим середнім та скінченою постійною коваріацією $\mathbf{Q}(k)$. Функцію прогнозування на один крок можна знайти як умовне математичне сподівання

$$\hat{\mathbf{x}}(k+1|k) = E_k [\mathbf{x}(k+1)] = \mathbf{F}\mathbf{x}(k), \quad (7.7.2)$$

де $\hat{\mathbf{x}}(k+1)$ – прогноз на один крок на основі інформації на момент k , включно. Функцією (7.7.2) можна скористатись для обчислення прогнозу на довільне число кроків. Так, прогноз на два крохи має вигляд:

$$\hat{\mathbf{x}}(k+2) = \mathbf{F}\hat{\mathbf{x}}(k+1) = \mathbf{F} \cdot \mathbf{F}\mathbf{x}(k) = \mathbf{F}^2\mathbf{x}(k),$$

і на довільне число кроків s :

$$\hat{\mathbf{x}}(k+s) = \mathbf{F}^s\mathbf{x}(k). \quad (7.7.3)$$

Очевидно, що дисперсія похибки прогнозу буде зростати пропорціонально числу кроків s . Похибка прогнозу на один і два крохи складає

$$\mathbf{e}_f(1) = \mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{w}(k) - \mathbf{F}\mathbf{x}(k) = \mathbf{w}(k),$$

$$\begin{aligned}\mathbf{e}_f(2) &= \mathbf{x}(k+2) - \hat{\mathbf{x}}(k+2) = \mathbf{F}\mathbf{x}(k+1) + \mathbf{w}(k+1) - \mathbf{F}[\mathbf{F}\mathbf{x}(k)] = \\ &= \mathbf{F}\mathbf{x}(k+1) + \mathbf{w}(k+1) - \mathbf{F}[\mathbf{x}(k+1) - \mathbf{w}(k)] = \mathbf{w}(k+1) + \mathbf{w}(k).\end{aligned}$$

Тобто, дисперсія прогнозу на s кроків визначається як

$$\begin{aligned}Var[\mathbf{e}_f(s)] &= E\{[\mathbf{w}(k+s-1) + \mathbf{w}(k+s-2) + \dots + \mathbf{w}(k)] \times \\ &\quad \times [\mathbf{w}^T(k+s-1) + \mathbf{w}^T(k+s-2) + \dots + \mathbf{w}^T(k)]\} = \\ &= s \mathbf{Q}_w.\end{aligned}$$

Отриманий результат аналогічний тому, що був отриманий для прогнозування на основі різницевих рівнянь, оскільки в обох випадках ми користуємося лінійними моделями одного класу. Тобто (7.7.1) можна інтерпретувати як стохастичну авторегресію першого порядку. Незважаючи на те, що модель досить проста, вона дає можливість визначити високоякісну оцінку однокрокового прогнозу завдяки тому, що оцінки стану формуються з використанням оптимального коефіцієнта фільтра.

Прогнозування навантаження на телефонну лінію (неперервний час)

Приклад 7.7.1. Задача побудови моделі та прогнозування навантаження на телефонний канал розглядається в багатьох роботах. Навантаження на канал має коливальний характер з незначним позитивним трендом. Максимальне навантаження припадає на зимові місяці, а мінімальне на середину літа, що пояснюється піком відпусткового сезону. На лінійний тренд процесу накладається шумова компонента, зумовлена тим, що коливання не носять чисто гармонічний характер. В цілому процес описується сумою коливальної та лінійної компонент з адитивним білим шумом. Лінійна частина неперервної моделі має вигляд:

$$\ddot{x}(t) = f_1(t), \tag{7.7.4}$$

де $f_1(t)$ – білий шум з нульовим середнім. Коливання описані рівнянням:

$$\ddot{y}(t) + \omega^2 y(t) = f_2(t), \tag{7.7.5}$$

де $f_2(t)$ – білий шум, незалежний від $f_1(t)$. Рівняння (7.7.4) і (7.7.5) записані у векторній формі:

$$x_1(t) = x(t), \tag{7.7.6}$$

$$x_2(t) = \dot{x}(t), \tag{7.7.7}$$

$$x_3(t) = y(t), \tag{7.7.8}$$

$$x_4(t) = \dot{y}(t). \quad (7.7.9)$$

Модель у просторі станів:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ f_1(t) \\ 0 \\ f_2(t) \end{bmatrix}. \quad (7.7.10)$$

Оскільки виміри мають дискретний характер з періодом дискретизації T_s , то модель перетворена в дискретну форму:

$$\mathbf{x}(k+1) = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\omega T_s) & \sin(\omega T_s) \\ 0 & 0 & -\sin(\omega T_s) & \cos(\omega T_s) \end{bmatrix} \mathbf{x}(k) + \mathbf{w}(k). \quad (7.7.11)$$

Вигляд елементів правої нижньої субматриці вимірності $[2 \times 2]$ зумовлений виглядом розв'язку диференціального рівняння (7.7.5). Значення елементів коваріаційної матриці шуму стану об'єкта $\mathbf{Q}(k)$ залежать від амплітуд випадкових величин $f_1(t)$ і $f_2(t)$. З визначенням елементів матриці $\mathbf{Q}(k)$, як правило, існують проблеми, а тому частіше всього значення її елементів підбирають в процесі обчислювальних експериментів.

Рівняння вимірів в даному випадку має досить просту форму завдяки тому, що загальний вимір представляє собою суму лінійної та коливальної частин. Таким чином, $z(k)$ – це скаляр, який визначається рівнянням:

$$z(k) = [1 \ 0 \ 1 \ 0] \mathbf{x}(k) + \mathbf{v}(k). \quad (7.7.12)$$

Оскільки виміри навантаження на телефонний канал є досить точними, то елементи коваріаційної матриці $\mathbf{R}(k)$ шумів вимірів $\mathbf{v}(k)$ в даному випадку мають відносно малі значення. Після визначення елементів матриць $\mathbf{F}(k)$, $\mathbf{Q}(k)$, $\mathbf{H}(k)$ і $\mathbf{R}(k)$ розв'язувалась задача прогнозування навантаження на телефонний канал на один крок. Як зазначалось вище, прогноз визначається добутком $\hat{\mathbf{x}}(k+1|k) = \mathbf{F}(k) \hat{\mathbf{x}}(k-1)$, тобто, для прогнозування можна скористатись звичайним алгоритмом фільтрації, який складається з наступних кроків:

1. Початкові умови (ініціалізація): задати $\hat{\mathbf{x}}(0)$, $\mathbf{P}(0)$, \mathbf{Q} і \mathbf{R} .
2. Обчислити коефіцієнт фільтра:
$$\mathbf{K}(k) = \mathbf{P}'(k) \mathbf{H}^T [\mathbf{H} \mathbf{P}'(k) \mathbf{H}^T + \mathbf{R}(k)]^{-1}.$$
3. Обчислити нову оцінку вектора стану:
$$\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(k | k-1) + \mathbf{K}(k)[\mathbf{z}(k) + \mathbf{H}\hat{\mathbf{x}}(k | k-1)].$$
4. Визначити елементи апостеріорної коваріаційної матриці:
$$\mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}] \mathbf{P}'(k).$$
5. Обчислити априорну коваріаційну матрицю для наступного кроку і прогнозовану оцінку вектора стану:
$$\mathbf{P}'(k+1) = \mathbf{F}(k) \mathbf{P}(k) \mathbf{F}^T(k) + \mathbf{Q},$$

$$\hat{\mathbf{x}}(k+1 | k) = \mathbf{F}(k) \hat{\mathbf{x}}(k).$$

Перейти на крок 2.

Для того щоб стабілізувати функціонування фільтра в перехідному процесі, було використано перших 16 вимірів навантаження на канал для оцінювання елементів коваріаційної матриці похибок оцінок вектора стану $\mathbf{P}(0)$ (фактично, це означає адаптацію моделі за допомогою експериментальних даних). Для отримання значень похибок оцінок використано зважений метод найменших квадратів з одночасним використанням всіх 16 вимірів.

Фактично використання наведеного алгоритму фільтрації розпочато з 17-го кроку і продовжувалось до того моменту, на який були відомі виміри. Іншим варіантом ініціалізації фільтра є використання довільних великих значень діагональних елементів матриці $\mathbf{P}(0)$. В результаті встановлено, що фільтр Калмана дозволяє обчислити досить точний (прийнятний) однокроковий прогноз для выбраної змінної – навантаження на телефонну лінію. Загалом, в спеціальній літературі можна знайти численні приклади застосування фільтра в різних галузях техніки, керуванні процесами різної природи і т. ін.

Фільтр Калмана також успішно застосовується для прогнозування динаміки змінних в багатьох сферах діяльності, в тому числі для прогнозування фінансових та економічних змінних і їх дисперсії, що підтверджено численними дослідженнями. Існуючі різновидності фільтра дають можливість охопити широкий клас стаціонарних та нестаціонарних процесів різної природи, а також нелінійні процеси, які можна апроксимувати лінійними моделями з прийнятною точністю.

Прогнозування фінансової змінної

Приклад 7.7.2. Розглянемо ряд, що описує дисперсію вартості акцій однієї з компаній, що входить до числа провідних на Нью-Йоркській

фондовій біржі. На основі часового ряду, що характеризує динаміку ціни акцій, побудовано модель дисперсії ціни

$$\varepsilon_1^2(k) = 0,0714 + 0,1187 \varepsilon_1^2(k-1) + 0,1123 \varepsilon_1^2(k-2) - 0,096 \varepsilon_1^2(k-4) + 0,1257 \varepsilon_1^2(k-5) + \varepsilon_2(k),$$

яка представлена також у просторі станів.

Прогноз дисперсії вартості акцій на 5 кроків виконано за допомогою трьох методів: безпосередньо за отриманою моделлю, за допомогою фільтра Калмана (який використовує ту ж модель) та методом подібних траєкторій.

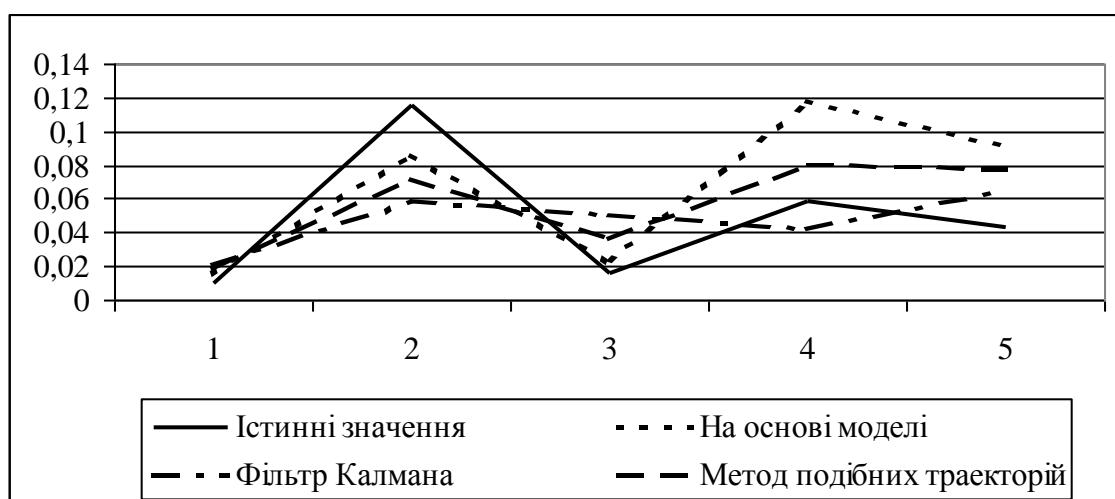


Рис. 7.2. Порівняльний графік прогнозу дисперсії вартості акцій компанії на Нью-Йоркській фондовій біржі

Таблиця 7.1

Метод прогнозування	Максимальне відхилення		Мінімальне відхилення		Сума квадратів похибок
	Абсолютне	%	Абсолютне	%	
За моделлю	0,055	73	0,003	8	0,0338
Фільтр Калмана	0,059	78	0,007	12	0,0337
Метод подібних траєкторій	0,058	77	0,007	10	0,0335

Графік, який ілюструє якість прогнозу, наведено на рисунку 7.1 (по

осі абсцис – число кроків прогнозування, а по осі ординат – значення похибки прогнозу.

Прогноз на один крок отримано за допомогою фільтра. З таблиці 7.1 видно (вона містить лише результати оцінювання якості прогнозу), що фільтр дозволяє також обчислити прийнятні значення прогнозу на п'ять кроків для досить складного процесу. Він є кращим за сумою квадратів похибок прогнозу. Численні обчислювальні експерименти з фільтром свідчать про те, що він, як правило, дає можливість знайти високоякісну оцінку однокрокового прогнозу. Однак, прогнозування на більше число кроків супроводжується високими похибками оцінок. Це можна пояснити використанням відносно простої моделі у просторі станів, яка враховує тільки попередній стан системи. Основною проблемою застосування оптимального фільтра залишається побудова математичної моделі процесу високого ступеня адекватності. Досить непростою є також задача оцінювання і прогнозування станів нелінійних систем, яка потребує коректної лінеаризації нелінійної моделі.

Запитання і вправи до сьомого розділу

1. Запишіть вираз для поточного середнього і поясніть принцип рекурсивного оцінювання. Чи можна застосовувати рекурсивне оцінювання на довгих часових інтервалах?
2. Яка форма математичної моделі використовується в методі оптимальної фільтрації Калмана, які її переваги? Наведіть приклад моделі процесу, яка може бути використана для побудови алгоритму оптимальної фільтрації?
3. Дайте визначення вільної динамічної системи? Що є причиною руху динамічної системи в даному випадку?
4. Запишіть модель у просторі станів загального вигляду для нестационарного динамічного процесу? Поясніть суть можливих нестационарностей.
5. Сформулуйте мету застосування оптимального фільтра. В чому полягає різниця між оптимальним і цифровим фільтром?
6. Яким чином враховуються статистичні параметри зовнішнього випадкового збурення $w(k)$ і шумів вимірювань $v(k)$ в алгоритмі оптимальної фільтрації?
7. Яким чином можна отримати оцінки коваріаційних матриць Q і R ? Що представляють собою діагональні елементи цих матриць?
8. Поясніть суть квадратичного функціоналу

$$J = E\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k)]^T [\mathbf{x}(k) - \hat{\mathbf{x}}(k)]\},$$

на мінімізації якого ґрунтуються обчислення оптимального матричного коефіцієнта фільтра Калмана? Яким чином можна отримати точні значення вектора стану $\mathbf{x}(k)$?

9. Які початкові умови необхідно задати оптимальному фільтру? Як їх можна отримати?
10. В чому полягає різниця між апріорною та апостеріорною коваріаційними матрицями похибок оцінок вектора стану динамічної системи?
11. Якого типу диференціальне рівняння необхідно розв'язувати для знаходження оптимального коефіцієнта фільтра?
12. Яка суть діагональних елементів коваріаційної матриці похибок оцінок вектора стану? Як повинні змінюватись їх значення у випадку нормального функціонування фільтра?
13. Сформулюйте своїми словами суть процесу статистичного моделювання оптимального фільтра. З якою метою виконується статистичне моделювання?
14. Запишіть повністю алгоритм статистичного моделювання процесу оптимальної фільтрації для моделі AR(1).
15. Назвіть фактори, які можуть спричинити розбіжність процедури оптимальної фільтрації? Поясніть, що означає розбіжність на практиці?
16. Назвіть, які заходи необхідно застосовувати для підвищення обчислювальної стійкості алгоритму фільтрації? Чи можна побудувати алгоритм оптимальної фільтрації без операції обернення матриці? У яких випадках?
17. Чому окремі компоненти вектора стану динамічної системи можуть невимірюватись?
18. Поясніть, яким чином можна оцінити невимірювані компоненти вектора стану системи?
19. Рівняння для обчислення коефіцієнта фільтра можна отримати за допомогою диференціального числення. Для цього необхідно мінімізувати елементи головної діагоналі коваріаційної матриці похибок оцінок вектора стану, яка визначається рівнянням Ріккаті:

$$\mathbf{P} = \mathbf{P}' - \mathbf{K}\mathbf{H}\mathbf{P}' - \mathbf{P}'\mathbf{H}^T\mathbf{K}^T + \mathbf{K}(\mathbf{H}\mathbf{P}'\mathbf{H}^T + \mathbf{R})\mathbf{K}^T.$$

Елементи головної діагоналі матриці \mathbf{P} необхідно мінімізувати по відношенню до коефіцієнта фільтра \mathbf{K} . Тобто можна мінімізувати слід цієї матриці $trace \mathbf{P} = p_{11} + p_{22} + \dots + p_{nn}$. При знаходженні похідних будуть корисними формули матричного числення:

$$(1) \quad \frac{d[trace(\mathbf{AB})]}{d\mathbf{A}} = \mathbf{B}^T, \quad (\text{матриця } \mathbf{AB} \text{ повинна бути квадратною});$$

$$(2) \quad \frac{d[\operatorname{trace}(\mathbf{A}\mathbf{C}\mathbf{A}^T)]}{d\mathbf{A}} = 2\mathbf{AC}, \quad (\text{матриця } \mathbf{C} \text{ повинна бути симетричною}).$$

Знайдіть вираз для оптимального коефіцієнта фільтра за цим методом.

Розділ 8

ПРЕДСТАВЛЕННЯ МАТЕМАТИЧНИХ МОДЕЛЕЙ У ПРОСТОРІ СТАНІВ

8.1 Вступ

В попередньому розділі ми використали представлення моделей у просторі станів (ПС) при формуванні рівнянь оптимального фільтра. Таке представлення є дуже зручним, а тому часто використовується в теорії моделювання, оцінювання, прогнозування і автоматичного керування. В цьому розділі розглянемо представлення у просторі станів моделей різних типів, зокрема, авторегресії та авторегресії з ковзним середнім, які вже були використані вище для описання динаміки процесів та побудови функцій прогнозування. Таке уніфіковане представлення математичних моделей в ПС спрощує їх подальше використання.

Від моделей у вигляді різницевих чи диференціальних рівнянь можна перейти до простору станів, який представляє собою уніфіковану форму представлення математичних моделей, зручною для застосування у багатьох випадках, наприклад, при прогнозуванні значень змінних та системі систем керування. Для неперервного часу лінійна модель у просторі станів має такий вигляд:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{w}(t), \quad (8.1.1)$$

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t), \quad (8.1.2)$$

де $\mathbf{x}(t)$ – вектор змінних стану вимірності n ; $\dot{\mathbf{x}}(t)$ – перша похідна вектора стану відносно часу (швидкість зміни значень вектора стану); $\mathbf{A}(t)$ –

матриця динаміки об'єкта (суб'єкта) чи процеса вимірності $[n \times n]$, яка в загальному випадку залежить від часу; $\mathbf{B}(t)$ – матриця коефіцієнтів керування вимірності $[n \times m]$, яка в загальному випадку також залежить від часу; $\mathbf{u}(t)$ – вектор сигналів керування вимірності $[m \times 1]$; $\mathbf{w}(t)$ – вектор збурень стану вимірності $[n \times 1]$; $\mathbf{z}(t)$ – вектор вимірювань (експериментальних даних) вимірності $[r \times 1]$; $\mathbf{H}(t)$ – матриця коефіцієнтів вимірювань вимірності $[r \times n]$; $\mathbf{v}(t)$ – вектор похибок (шумів) вимірів вимірності r . Наведена вище модель може бути, при необхідності, частково модифікована з метою її адаптації до конкретного типу процесів. Досить часто приймається припущення (яке повинне бути доведеним для конкретного випадку) про те, що збурення стану $\mathbf{w}(t)$ та шуми вимірів $\mathbf{v}(t)$ є незалежними гаусівськими процесами з постійними дисперсіями та нульовими середніми значеннями, тобто

$$E[\mathbf{w}(t)] = 0, \quad E[\mathbf{w}(t)\mathbf{w}^T(\tau)] = \mathbf{Q}(t)\delta(t - \tau); \quad (8.1.3)$$

$$E[\mathbf{v}(t)] = 0, \quad E[\mathbf{v}(t)\mathbf{v}^T(\tau)] = \mathbf{R}(t)\delta(t - \tau), \quad (8.1.4)$$

де $E[\cdot]$ – символ математичного сподівання; $\mathbf{w}^T(t)$ – означає операцію транспонування вектора або матриці; $\mathbf{Q}(t)$, $\mathbf{R}(t)$ – коваріаційні матриці збурень стану та шумів вимірів, відповідно (діагональні елементи цих матриць є дисперсіями сигналів $\mathbf{w}(t)$ і $\mathbf{v}(t)$). Модель в просторі станів для дискретного часу має такий же вигляд як і (8.1.1), (8.1.2), але є різниця у методах отримання матриць \mathbf{A} , \mathbf{B} , \mathbf{H} . Нижче буде розглянуто приклади переходу від диференціальних та різницевих рівнянь до простору станів.

Приклад 8.1.1. Розглянемо модель моторної функції людини, яку часто описують диференціальним рівнянням другого порядку

$$a_2 \frac{d^2y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + y(t) = bu(t), \quad (8.1.5)$$

де $y(t)$ – сигнал (реакція) на виході суб'єкта; $u(t)$ – вхідний сигнал (через систему візуального або слухового сприйняття). Для того щоб представити рівняння (8.1.5) в просторі станів, приведемо його спочатку до вигляду:

$$\frac{d^2y(t)}{dt^2} + \frac{a_1}{a_2} \frac{dy(t)}{dt} + \frac{1}{a_2} y(t) = \frac{b}{a_2} u(t), \quad (8.1.6)$$

і введемо позначення $x_1 = y(t)$, $x_2 = \dot{y}(t)$. Тепер

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ i \quad \dot{x}_2(t) &= \ddot{y}(t) = -\frac{a_1}{a_2} \dot{y}(t) - \frac{1}{a_2} y(t) + \frac{b}{a_2} u(t). \end{aligned}$$

При таких позначеннях можна записати рівняння (8.1.5) в матричній формі як

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1/a_2 & -a_1/a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ b/a_2 \end{bmatrix} u(t).$$

Таким чином, звичайне диференціальне рівняння другого порядку представлено у вигляді системи двох рівнянь першого порядку або у формі Коши, або у просторі станів. Рівняння в частинних похідних також можна представити в такій формі.

Розглянемо лінійне диференціальне рівняння n -го порядку, що описує систему з одним входом і одним виходом в неперервному часі:

$$(p^n + \alpha_{n-1}p^{n-1} + \dots + \alpha_1 p + \alpha_0)y(t) = (\beta_n p^n + \beta_{n-1}p^{n-1} + \dots + \beta_1 p + \beta_0)u(t), \quad (8.1.7)$$

де $u(t)$, $y(t)$ – вход і вихід системи, відповідно; $p = d/dt$ – оператор диференціювання; $\alpha_i, \beta_i, i = 1, \dots, n$ – параметри моделі. На рис.8.1 наведена структурна схема обчислень, яка відповідає цьому рівнянню.

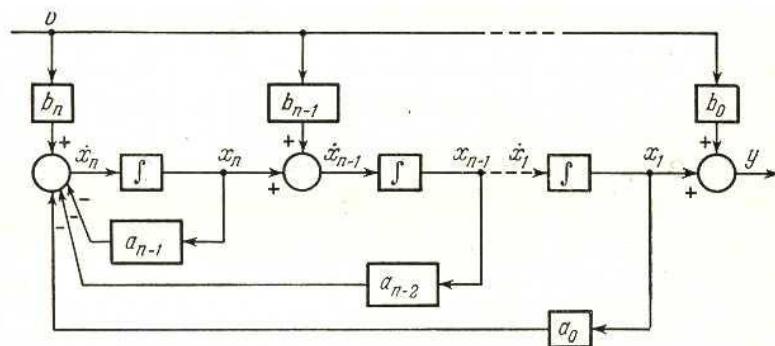


Рис.8.1. Структурна схема для рівняння (8.1.7)

Безпосередньо із цієї схеми можна записати, що

$$\begin{aligned} y(t) &= x_1(t) + b_0 u(t), \\ \dot{x}_k(t) &= x_{k+1}(t) + b_k u(t), \quad \forall k < n, \\ \dot{x}_n(t) &= -[a_0 x_1(t) + a_1 x_2(t) + \dots + a_{n-1} x_n(t)] + b_n u(t). \end{aligned} \quad (8.1.8)$$

Запишемо похідну від $y(t)$

$$p[y(t)] = \dot{y}(t) = \dot{x}_1(t) + b_0 \dot{u}(t)$$

і підставимо замість $\dot{x}_1(t)$ відповідний вираз із (8.1.8):

$$p[y(t)] = \dot{y}(t) = x_2(t) + b_1 u(t) + b_0 \dot{u}(t). \quad (8.1.9)$$

Тепер, по аналогії, запишемо старші похідні від $y(t)$:

$$\begin{aligned} p^2[y(t)] &= \dot{x}_2(t) + b_1 \dot{u}(t) + b_0 \ddot{u}(t) = x_3(t) + b_2 u(t) + b_1 \dot{u}(t) + b_2 \ddot{u}(t), \\ &\vdots \\ p^{n-1}[y(t)] &= x_n + b_{n-1} u(t) + b_{n-2} p[u(t)] + \dots + b_0 p^{n-1}[u(t)], \\ p^n[y(t)] &= -[a_0 x_1(t) + a_1 x_2(t) + \dots + a_{n-1} x_n(t)] + \\ &\quad + b_n u(t) + b_{n-1} p[u(t)] + \dots + b_0 p^n[u(t)]. \end{aligned} \quad (8.1.10)$$

Якщо підставити $y(t)$, $p[y(t)]$, ..., $p^{n-1}[y(t)]$, які визначаються рівняннями (8.1.8)–(8.1.10) в рівняння (8.1.7) і співставити результат з виразом для $p^n[y(t)]$, що задається рівнянням (8.1.10), то отримаємо вирази для коефіцієнтів a_i , b_i у вигляді:

$$a_i = \alpha_i, \quad (8.1.11)$$

$$b_0 = \beta_n,$$

$$b_1 = \beta_{n-1} - \alpha_{n-1} b_0,$$

$$b_2 = \beta_{n-2} - \alpha_{n-1} b_1 - \alpha_{n-2} b_0,$$

\vdots

$$b_n = \beta_0 - \alpha_{n-1} b_{n-1} - \alpha_{n-2} b_{n-2} - \dots - \alpha_0 b_0. \quad (8.1.12)$$

Рівняння (8.1.12) представляє собою зручну форму запису виразу для b_i , які можна знайти шляхом послідовних підстановок. Так, з рівняння (8.1.12) випливає, що β_i можна записати у вигляді:

$$\begin{bmatrix} \beta_n \\ \beta_{n-1} \\ \beta_{n-2} \\ \vdots \\ \beta_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha_{n-1} & 1 & 0 & \cdots & 0 \\ \alpha_{n-2} & \alpha_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \alpha_0 & \alpha_1 & \cdots & \alpha_{n-1} & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

а звідси можна визначити коефіцієнти b_i за допомогою оберненої матриці. Таким чином, згідно із отриманими виразами, однією з можливих форм матриць **A**, **B** і **H** моделі в просторі станів є така:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-1} \end{bmatrix}, \quad (8.1.13)$$

$$\mathbf{B} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha_{n-1} & 1 & 0 & \cdots & 0 \\ \alpha_{n-2} & \alpha_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ \alpha_0 & \alpha_1 & \cdots & \alpha_{n-1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \beta_n \\ \beta_{n-1} \\ \beta_{n-2} \\ \vdots \\ \beta_0 \end{bmatrix}, \quad (8.1.14)$$

$$\mathbf{H} = [1 \ 0 \ 0 \ \cdots \ 0]. \quad (8.1.15)$$

Приклад 8.1.2. Привести до форми простору станів диференціальне рівняння третього порядку:

$$\ddot{y}(t) + 3\dot{y}(t) + 4y(t) = 2\ddot{u}(t) + 3\dot{u}(t) + u(t).$$

За допомогою рівняння (8.1.12) отримаємо:

$$\begin{aligned} b_0 &= \beta_n = 2, \\ b_1 &= \beta_{n-1} - \alpha_{n-1}b_0 = 3 - 3 \cdot 2 = -3, \\ b_2 &= \beta_{n-2} - \alpha_{n-1}b_1 - \alpha_{n-2}b_0 = 1 - 3 \cdot (-3) - 4 \cdot 2 = 2, \\ b_3 &= \beta_{n-3} - \alpha_{n-1}b_2 - \alpha_{n-2}b_1 - \alpha_{n-3}b_0 = 2 - 3 \cdot 2 - 4 \cdot (-3) - 1 \cdot 2 = 6. \end{aligned}$$

Згідно з представленням (8.1.13), (8.1.14) запишемо модель в просторі станів:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -4 & -3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} -3 \\ 2 \\ 6 \end{bmatrix} u(t).$$

З попередніх розділів видно, що наші методи спрямовані на дослідження дискретних систем у просторі станів. Перед тим як перейти до представлення дискретних систем у просторі станів, розглянемо перехід від лінійних неперервних до лінійних дискретних систем.

8.2 Перехід від лінійних неперервних до лінійних дискретних систем

Багатовимірну лінійну стаціонарну систему можна описати системою лінійних диференціальних рівнянь у векторній формі

$$\dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t), \quad (8.2.1)$$

де \mathbf{A} – матриця з постійними коефіцієнтами вимірності $[n \times n]$; $\mathbf{x}(t)$ – вектор-стовпчик змінних стану вимірності $[n \times 1]$. Подібно розв'язку скалярних диференціальних рівнянь розв'язок рівняння (8.2.1) записується у вигляді:

$$\mathbf{x}(t) = e^{\mathbf{A}(t-\tau)} \mathbf{x}(\tau), \quad (8.2.2)$$

де матриця $e^{\mathbf{A}t}$ може бути представлена у вигляді розкладу в нескінчений ряд

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \dots . \quad (8.2.3)$$

При $t = \tau$ матриця $e^{\mathbf{A}(\tau-\tau)}$ стає одиничною, тобто граничні умови виконуються

$$\mathbf{x}(\tau) = e^{\mathbf{A}0} \mathbf{x}(\tau) = \mathbf{I} \mathbf{x}(\tau) = \mathbf{x}(\tau).$$

Матрицю $\mathbf{F}(t) = e^{\mathbf{A}t}$ називають *перехідною матрицею стану* або *фундаментальною матрицею* системи (8.2.1).

Перехідна матриця стану “описує” рух кінця вектора стану в просторі станів з початкового значення до наступного (заданого кінцевим моментом часу), а тому вона характеризує (описує) перехід стану системи. Очевидно, що вектор стану $\mathbf{x}(t) = [x_1(t) \dots x_n(t)]^T$ містить в собі значну інформацію про лінійну систему. Перехідну матрицю стану $\mathbf{F}(t)$ можна визначати різними методами. Основні підходи базуються на теоремі Сільвестра та

Келі-Гамільтона, методі розкладання в нескінчений ряд, перетворенні Лапласа та методі передавальної функції.

Метод розкладання в нескінчений ряд. За визначенням $e^{\mathbf{A}t}$ перехідна матриця обчислюється за допомогою нескінченого ряду

$$\mathbf{e}^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \dots$$

Цей метод потребує найбільше обчислювальних витрат, якщо значення \mathbf{A}^k не збігаються до нуля при збільшенні k . Після обчислення суми необхідно знайти ще прийнятні (замкнені) вирази для кожного елемента матриці $\mathbf{F}(t)$.

Приклад 8.2.1. Знайти перехідну матрицю стану для матриці

$$\mathbf{A} = \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix}.$$

Зайдемо \mathbf{A}^2 і \mathbf{A}^3 :

$$\mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A} = \begin{bmatrix} -2 & 6 \\ -3 & 7 \end{bmatrix}, \quad \mathbf{A}^3 = \mathbf{A}^2 \cdot \mathbf{A} = \begin{bmatrix} 6 & 14 \\ 7 & -15 \end{bmatrix}.$$

Таким чином, перехідна матриця визначається як

$$\mathbf{F}(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix}t + \begin{bmatrix} -2 & 6 \\ -3 & 7 \end{bmatrix} \frac{t^2}{2!} + \begin{bmatrix} 6 & 14 \\ 7 & -15 \end{bmatrix} \frac{t^3}{3!} + \dots$$

або

$$\mathbf{F}(t) = \begin{bmatrix} 1 - \frac{2t^2}{2!} + \frac{6t^3}{3!} + \dots & -2t + \frac{6t^2}{2!} + \frac{14t^3}{3!} + \dots \\ t - \frac{3t^2}{2!} + \frac{7t^3}{3!} + \dots & 1 - 3t + \frac{7t^2}{2!} - \frac{15t^3}{3!} + \dots \end{bmatrix}.$$

Обчислюючи нескінченні ряди для кожного з елементів матриці (це основний недолік даного методу), отримаємо

$$\mathbf{F}(t) = \begin{bmatrix} 2e^{-t} - e^{2t} & 2(e^{-2t} - e^{-t}) \\ e^{-t} - e^{-2t} & 2e^{-2t} - e^{-t} \end{bmatrix}.$$

Якщо позначити період дискретизації вимірювань через T_s , то перехід лінійної системи з одного стану в інший на протязі періоду дискретизації легко визначити за допомогою підстановки $t = T_s$.

Метод на основі перетворення Лапласа (метод комплексної частоти). Перетворення Лапласа для рівняння динаміки (8.2.1) має вигляд: $s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s)$, тобто,

$$\mathbf{X}(s) = [s\mathbf{I} - \mathbf{A}]^{-1} \mathbf{x}(0). \quad (8.2.4)$$

Застосуємо до (8.2.4) зворотне перетворення Лапласа

$$\mathbf{x}(t) = L^{-1}\{[s\mathbf{I} - \mathbf{A}]^{-1}\} \mathbf{x}(0). \quad (8.2.5)$$

Порівнюючи (8.2.5) з (8.2.2), отримаємо

$$\mathbf{F}(t) = e^{\mathbf{A}t} = L^{-1}\{[s\mathbf{I} - \mathbf{a}]^{-1}\}. \quad (8.2.6)$$

Даний метод є одним із самих зручних при розв'язуванні багатьох задач. Труднощі можуть виникати при знаходженні оберненої матриці.

Приклад 8.2.2. Обчислити $\mathbf{F}(t)$ для матриці \mathbf{A} з попереднього прикладу за методом перетворення Лапласа. Для матриці

$$[s\mathbf{I} - \mathbf{A}] = \begin{bmatrix} s & 2 \\ -1 & s+3 \end{bmatrix}$$

обернена має вигляд

$$\Phi(s) = [s\mathbf{I} - \mathbf{A}]^{-1} = \frac{1}{\Delta} \begin{bmatrix} s+3 & -2 \\ 1 & s \end{bmatrix},$$

де $\Delta = s^2 + 3s + 2$, а матриця $\mathbf{F}(t) = L^{-1}[\Phi(s)]$.

Крім розглянутих методів, переходну матрицю станів можна знайти за допомогою деяких інших методів, наприклад, передаточної функції.

8.3 Дискретні рівняння типу АРКС в просторі станів

В дискретному часі динаміка систем описується різницевими рівняннями, для яких також розроблено процедури переходу до простору станів. Наприклад, можна побудувати авторегресійні моделі для змінних, що характеризують поточний стан людини і представити отриману модель у просторі станів.

Приклад 8.3.1. Нехай для характеристики поточного стану людини вибрано наступні змінні: $x_1(k)$ – температура тіла ($k=0,1,2,\dots$ – дискретний час); $x_2(k)$ – артеріальний тиск крові; $x_3(k)$ – частота пульсу; $x_4(k)$ –

електричний опір шкіри; $x_5(k)$ – вміст цукру в крові. Якщо вимірювати ці змінні два рази на добу, то за місяць отримаємо для кожної із них вибірку із 60 вимірювань, яку можна використати для побудови моделей у вигляді авторегресійних рівнянь, що описуватимуть динаміку отриманих часових рядів. Очевидно, що більш глибокий аналіз процесів в організмі людини дозволить вибрати складніші структури моделей, які будуть містити в собі незалежні змінні в правій частині рівнянь та можливі нелінійності. Але якщо зупинитися на авторегресії першого порядку, то отримаємо наступну систему рівнянь:

$$\begin{aligned}x_1(k) &= a_{10} + a_{11}x_1(k-1) + \varepsilon_1(k), \\x_2(k) &= a_{20} + a_{21}x_2(k-1) + \varepsilon_2(k), \\x_3(k) &= a_{30} + a_{31}x_3(k-1) + \varepsilon_3(k), \\x_4(k) &= a_{40} + a_{41}x_4(k-1) + \varepsilon_4(k), \\x_5(k) &= a_{50} + a_{51}x_5(k-1) + \varepsilon_5(k),\end{aligned}$$

яку можна представити у просторі станів у вигляді

$$\mathbf{x}(k) = \mathbf{F}(k)\mathbf{x}(k-1) + \mathbf{a} + \mathbf{w}(k),$$

де $\mathbf{x}^T(k) = [x_1(k) \ x_2(k) \ x_3(k) \ x_4(k) \ x_5(k)]$ – вектор стану; $\mathbf{a}^T = [a_{10} \ a_{20} \ a_{30} \ a_{40} \ a_{50}]$ – вектор постійних значень; $\mathbf{F}(k)$ – матриця динаміки дискретної системи; $\mathbf{w}^T(k) = [\varepsilon_1(k) \ \varepsilon_2(k) \ \varepsilon_3(k) \ \varepsilon_4(k) \ \varepsilon_5(k)]$. Матриця динаміки і вектор $\mathbf{x}(k-1)$ для даної системи мають вигляд:

$$\begin{aligned}\mathbf{F}(k) &= diag[a_{11} \ a_{21} \ a_{31} \ a_{41} \ a_{51}]; \\ \mathbf{x}^T(k) &= [x_1(k) \ x_2(k) \ x_3(k) \ x_4(k) \ x_5(k)].\end{aligned}$$

Очевидно, що така модель, побудована для відхилень від середнього, не буде містити вектора констант \mathbf{a} .

Цей приклад можна розглядати як спрощене представлення в просторі станів, оскільки керуючого впливу в правій частині нема. В цілому отриману систему в просторі станів не можна вважати вільною, тому що в правій частині наявні вектори зміщення \mathbf{a} та збурень $\mathbf{w}(k)$. Нагадаємо, що вільною називають таку динамічну систему, рух якої відбувається під впливом тільки початкових умов.

Розглянемо лінійну систему, яка описується авторегресією з ковзним середнім довільного порядку. Нехай зв'язок між входом і виходом процесу визначається за допомогою різницевого рівняння

$$\begin{aligned} y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) &= \\ &= b_0 u(k) + b_1 u(k-1) + b_2 u(k-2) + \dots + b_m u(k-m), \end{aligned}$$

або

$$\begin{aligned} y(k+n) + a_1 y(k+n-1) + \dots + a_n y(k) &= \\ &= b_0 u(k+n) + b_1 u(k+n-1) + b_2 u(k+n-2) + \dots + b_m u(k), \end{aligned} \quad (8.3.1)$$

де $u(k)$, $y(k)$ – вхід та вихід процесу, відповідно. Якщо ввести позначення $y(k-1) = z^{-1}y(k)$, наведене рівняння можна переписати у вигляді:

$$(1 + a_1 z^{-1} + \dots + a_n z^{-n}) y(k) = (b_0 + b_1 z^{-1} + \dots + b_m z^{-m}) u(k).$$

Оператор z^{-1} називають оператором зсуву (в часі). Рівняння (8.3.1) можна представити у вигляді дискретної функції передачі між виходом і входом як

$$G(z) = \frac{y(z)}{u(z)} = \frac{b_0 z^n + b_1 z^{n-1} + \dots + b_m z^{-m}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}}.$$

або при $m=n$:

$$G(z) = \frac{y(z)}{u(z)} = \frac{b_0 z^n + b_1 z^{n-1} + \dots + b_n}{z^n + a_1 z^{n-1} + \dots + a_n}. \quad (8.3.2)$$

Зазначимо, що для процесів, які можуть бути фізично реалізованими, необхідно виконувати умову $m \leq n$, що забезпечує стійкість їх функціонування.

Введемо такі змінні стану:

$$\left. \begin{array}{l} x_1(k) = y(k) \\ x_2(k) = x_1(k+1) = y(k+1) \\ x_3(k) = x_2(k+1) = y(k+2) \\ \vdots \\ x_n(k) = x_{n-1}(k+1) = y(k+n-1) \\ x_n(k+1) = y(k+n) \end{array} \right\} \quad (8.3.3)$$

Підставимо вирази (8.3.3) в рівняння (8.3.1) при $b_n = 1$ і $b_0, b_1, \dots, b_{n-1} = 0$. В результаті отримаємо:

$$y(k+n) = x_n(k+1) = -a_1 x_n(k) - a_2 x_{n-1}(k) - \dots - a_n x_1(k) + 1 u(k). \quad (8.3.4)$$

Це співвідношення можна представити у формі простору станів як

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} u(k), \quad (8.2.5)$$

а рівняння вимірів має вигляд:

$$y(k) = [1 \ 0 \ \cdots \ 0] \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix}. \quad (8.3.6)$$

Тепер рівняння (8.3.1) можна представити у просторі станів наступним чином:

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k), \quad (8.3.7)$$

$$\mathbf{y}(k) = \mathbf{H}^T \mathbf{x}(k), \quad (8.3.8)$$

де матриці \mathbf{F} , \mathbf{B} , \mathbf{H} визначаються у відповідності до (8.3.5) та (8.3.6).

Якщо $b_n = 1$, а $b_0, b_1, \dots, b_{n-1} = 0$, то рівняння (8.3.2), (8.3.3) можна представити у вигляді

$$y(z) = \frac{1}{z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n} u(z) = x_1(z). \quad (8.3.9)$$

Якщо ж $b_n \neq 1$, а $b_0, b_1, \dots, b_{n-1} \neq 0$, то із умови стійкості можна записати, що

$$b_n + b_{n-1}z + b_{n-2}z^2 + \dots + b_0z^n = 1. \quad (8.3.10)$$

Якщо помножити ліву і праву частини (8.3.10) на $y(z)$, то отримаємо:

$$y(z) = b_n y(z) + b_{n-1}z y(z) + \dots + b_0 z^n y(z),$$

а із врахуванням того, що $y(z) = x_1(z)$ маємо:

$$y(z) = b_n x_1(z) + b_{n-1} z x_1(z) + \dots + b_0 z^n x_1(z)$$

або

$$y(k) = b_n x_1(k) + b_{n-1} x_1(k+1) + \dots + b_0 x_1(k+n). \quad (8.3.11)$$

Використовуючи позначення (8.3.3), перейдемо до наступного запису (8.3.11):

$$y(k) = b_n x_1(k) + b_{n-1} x_2(k) + b_{n-2} x_3(k) + \dots + b_0 x_n(k+1). \quad (8.3.12)$$

З рівняння (8.3.4) отримуємо:

$$x_n(k+1) = -a_1 x_n(k) - a_2 x_{n-1}(k) - \dots - a_n x_1(k) + 1 u(k), \quad (8.3.13)$$

а в результаті підстановки правої частини (8.3.13) в (8.3.12) замість $x_n(k+1)$ отримаємо:

$$y(k) = (b_n - b_0 a_n) x_1(k) + (b_{n-1} - b_0 a_{n-1}) x_2(k) + \dots + (b_1 - b_0 a_1) x_n(k) + b_0 u(k). \quad (8.3.14)$$

Останнє рівняння представляє собою рівняння вимірів, в якому враховано вплив чисельника передаточної функції, тобто воно приймає вигляд:

$$y(k) = [(b_n - b_0 a_n) \dots (b_1 - b_0 a_1)] \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix} + b_0 u(k),$$

або

$$y(k) = \mathbf{H}^T \mathbf{x}(k) + d u(k). \quad (8.3.15)$$

При $b_0 = 0$, тобто для систем без прямої передачі керуючого сигналу, рівняння (8.3.15) приймає спрощений вигляд

$$y(k) = [b_n \dots b_1] \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix}. \quad (8.3.16)$$

Другий підхід до представлення моделі (8.3.1) у просторі станів

В даному випадку матриця \mathbf{F} в (8.3.7) визначається так само як і в попередньому випадку, а матриця \mathbf{B} визначається за виразом:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ a_1 & 1 & 0 & 0 & \dots & 0 \\ a_2 & a_1 & 1 & 0 & & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot \\ a_{n-1} & \cdot & \cdot & \cdot & \cdot & a_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix} = (\mathbf{A}')^{-1} \mathbf{b}. \quad (8.3.17)$$

Матрицю \mathbf{F} називають коагульованою матричною формою; \mathbf{A}' – нижня трикутна матриця. Покажемо, що представлення у просторі станів

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} u(k), \quad (8.3.18)$$

відповідає початковій моделі (8.3.1). В (8.3.18)

$$\mathbf{x}^T = [x_1 \ x_2 \ \dots \ x_n], \quad \mathbf{B}^T = [\beta_1 \ \beta_2 \ \dots \ \beta_n]. \quad (8.3.19)$$

Вище було показано, що

$$x_1(k) = y(k). \quad (8.3.20)$$

У відповідності до (8.3.18) та (8.3.19) можна записати

$$\begin{aligned} x_1(k) &= x_2(k-1) + \beta_1 u(k-1) \\ \text{або} \quad x_2(k) &= x_1(k+1) - \beta_1 u(k) = y(k+1) - \beta_1 u(k). \end{aligned}$$

По аналогії можна отримати

$$x_i(k) = y(k+i-1) - \sum_{j=2}^i \beta_{j-1} u(k+i-j), \quad i \leq n. \quad (8.3.21)$$

Відповідно, $x_n(k)$ визначається рівнянням

$$x_n(k) = y(k+n-1) - \beta_1 u(k+n-2) - \dots - \beta_{n-1} u(k-1). \quad (8.3.22)$$

Прирівняємо нижній рядок матричного рівняння (8.3.18) до виразу (8.3.22) і скористаємося матрицею \mathbf{F} , що визначається (8.3.18). В результаті отримаємо

$$\begin{aligned} -a_n x_1(k-1) - \dots - a_1 x_n(k-1) + \beta_n u(k-1) = \\ = y(k+n-1) - \beta_1 u(k+n-2) - \dots - \beta_{n-1} u(k). \end{aligned} \quad (8.3.23)$$

Тепер підставимо вираз для $x_i(k)$, $i=1, \dots, n-1$ із рівняння (8.3.21) в останнє рівняння і отримаємо

$$\begin{aligned} y(k+n-1) - \beta_1 u(k+n-2) - \dots - \beta_{n-1} u(k) = & -a_n y(k-1) - \dots - a_1 y(k+n-2) + \\ & + a_{n-1} \beta_1 u(k-1) + a_{n-2} [\beta_1 u(k) + \beta_2 u(k-1)] + \\ & + \dots + a_1 [\beta_1 u(k+n-3) + \dots + \beta_{n-1} u(k-1)] + \beta_n u(k-1). \end{aligned} \quad (8.3.24)$$

Зробимо перегрупування в останньому рівнянні

$$\begin{aligned} y(k+n-1) + a_1 y(k+n-2) + a_2 y(k+n-3) + \dots + a_n y(k-1) = \\ = a_{n-1} \beta_1 u(k-1) + a_{n-2} [\beta_1 u(k) + \beta_2 u(k-1)] + \dots + \\ + a_1 [\beta_1 u(k+n-3) + \dots + \beta_{n-1} u(k-1)] + \\ + \beta_1 u(k+n-2) + \dots + \beta_n u(k-1) = \\ = b_1 u(k+n-2) + b_2 u(k+n-3) + \dots + b_m u(k+n-m-1). \end{aligned} \quad (8.3.25)$$

Таким чином,

$$\begin{aligned} b_1 &= \beta_1, \\ b_2 &= \beta_2 + a_1 \beta_1, \\ &\vdots \\ b_m &= \beta_m + a_1 \beta_{m-1} + \dots + a_{m-1} \beta_1, \\ 0 &= \beta_j + a_1 \beta_{j-1} + \dots + a_{j-1} \beta_1, \quad n \geq j > m. \end{aligned} \quad (8.3.26)$$

Рівняння (8.3.26) можна записати таким чином:

$$\begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ b_m \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdot & \cdot & 0 \\ a_1 & 1 & 0 & 0 & \cdot & \cdot & 0 \\ a_2 & a_1 & 1 & 0 & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & & 0 & \cdot \\ a_{n-1} & \cdot & \cdot & \cdot & \cdot & a_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \beta_n \end{bmatrix} = \mathbf{A}' \mathbf{B} \quad (8.3.27)$$

або

$$\mathbf{B} = (\mathbf{A}')^{-1} \begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ b_m \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = (\mathbf{A}')^{-1} \mathbf{b}, \quad (8.3.28)$$

де \mathbf{b} – n – вимірний вектор; $m \leq n$. Оскільки рівняння (8.3.25) задовольняє рівнянням (8.3.20)-(8.3.21), то перетворення (8.3.1) до простору станів закінчено.

Приклад 8.3.2. Розглянемо функцію передачі другого порядку в дискретному часі

$$\begin{aligned} G(z) &= \frac{0,5z}{z^2 - 1,3z + 0,9} = \frac{0,5z^{-1}}{1 - 1,3z^{-1} + 0,9z^{-2}} = \\ &= \frac{b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} = \frac{y(z)}{u(z)}. \end{aligned}$$

Для того щоб отримати дискретну модель цієї системи в просторі станів, обчислимо матриці \mathbf{F} , \mathbf{B} у відповідності до (8.3.5) та (8.3.6):

$$\begin{aligned} \mathbf{F} &= \begin{bmatrix} 0 & 1 \\ -0,9 & +1,3 \end{bmatrix}; \\ \mathbf{B} &= \begin{bmatrix} 1 & 0 \\ -1,3 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1,3 & 1 \end{bmatrix} \begin{bmatrix} 0,5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,5 \\ 0,65 \end{bmatrix}. \end{aligned}$$

В результаті отримаємо:

$$\begin{aligned}x_1(k) &= x_2(k-1) + 0,5u(k-1), \\x_2(k) &= -0,9x_1(k-1) + 1,3x_2(k-2) + 0,65u(k-1), \\y(k) &= x_1(k).\end{aligned}$$

Приклад 8.3.3. Розглянемо різницеве рівняння третього порядку

$$y(k) = 0,5y(k-1) - 0,25y(k-2) + 0,4y(k-3) + 0,8u(k-1) - 0,6u(k-2),$$

якому відповідає функція передачі

$$G(z) = \frac{y(z)}{u(z)} = \frac{0,8z^{-1} - 0,6z^{-2}}{1 - 0,5z^{-1} + 0,25z^{-2} - 0,4z^{-3}}.$$

Отримаємо модель цієї системи в просторі станів у відповідності до (8.3.5) та (8.3.6):

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0,4 & -0,25 & +0,5 \end{bmatrix}; \quad \mathbf{A}' = \begin{bmatrix} 1 & 0 & 0 \\ -0,5 & 1 & 0 \\ 0,25 & -0,5 & 1 \end{bmatrix};$$

$$(\mathbf{A}')^{-1} = \frac{1}{\det(\mathbf{A}')} \begin{bmatrix} \text{транспонована матриця} \\ \text{алгебраїчних доповнень} \\ \text{для } \mathbf{A}' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0,5 & 1 & 0 \\ 0 & 0,5 & 1 \end{bmatrix};$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0,5 & 1 & 0 \\ 0 & 0,5 & 1 \end{bmatrix} \begin{bmatrix} 0,8 \\ -0,6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,8 \\ -0,2 \\ -0,3 \end{bmatrix}.$$

В результаті отримаємо:

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{F}\mathbf{x}(k-1) + \mathbf{B}u(k-1) = \\&= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0,4 & -0,25 & +0,5 \end{bmatrix} \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \end{bmatrix} + \begin{bmatrix} 0,8 \\ -0,2 \\ -0,3 \end{bmatrix} u(k-1)\end{aligned}$$

або

$$\begin{aligned}x_1(k) &= x_2(k-1) + 0,8u(k-1), \\x_2(k) &= x_3(k-1) - 0,2u(k-1), \\x_3(k) &= 0,4x_1(k-1) - 0,25x_2(k-1) + 0,5x_3(k-1) - 0,3u(k-1),\end{aligned}$$

$$y(k) = x_1(k).$$

Розглянуті вище математичні моделі складають лише незначну частину того, що використовується і може використовуватись в економетричному аналізі і технічних системах.

8.4 Приклади представлення моделей у просторі станів

В теорії аналізу часових рядів часто використовують таку форму представлення дискретних моделей у просторі станів (рівняння стану і рівняння вимірів):

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}\mathbf{x}(k) + \mathbf{w}(k), \\ \mathbf{z}(k) &= \mathbf{A}^T \mathbf{u}(k) + \mathbf{H}^T \mathbf{x}(k) + \mathbf{v}(k),\end{aligned}$$

де $\mathbf{F}, \mathbf{A}^T, \mathbf{H}^T$ – матриці параметрів вимірності $[n \times n], [r \times p]$ і $[r \times n]$, відповідно; $\mathbf{u}(k)$ – вектор регресорів (екзогенних змінних) вимірності $[p \times 1]$. Як і раніше, $\mathbf{w}(k), \mathbf{v}(k)$ – вектори збурень стану і шумів вимірів, відповідно, з такими статистичними характеристиками:

$$E[\mathbf{w}(k) \mathbf{w}^T(l)] = \begin{cases} \mathbf{Q}, & k = l \\ 0, & k \neq l \end{cases}, \quad E[\mathbf{v}(k) \mathbf{v}^T(l)] = \begin{cases} \mathbf{R}, & k = l \\ 0, & k \neq l \end{cases},$$

де $\mathbf{Q}[n \times n], \mathbf{R}[r \times r]$ – коваріаційні матриці збурень та шумів вимірів, відповідно. Послідовності $\{\mathbf{w}(k)\}, \{\mathbf{v}(k)\}$ некорельовані для будь-яких моментів часу:

$$E[\mathbf{w}(k) \mathbf{v}^T(l)] = 0, \quad \forall k, l.$$

Припускається, що вектор екзогенних змінних $\mathbf{u}(k)$ не містить інформації щодо $\mathbf{x}(k+s)$ та $\mathbf{v}(k+s)$ для $s = 0, 1, 2, \dots$ окрім тієї, що міститься в $y(k-1), y(k-2), \dots, y(1)$. Наприклад, $\mathbf{u}(k)$ може містити в собі затримані у часі значення у або змінні, що некорельовані з $\mathbf{x}(j)$ або $\mathbf{v}(j), \forall j$.

Модель у просторі станів використовується, як правило, для описання динаміки деякого скінченного числа спостережень $\{y(1), y(2), \dots, y(T)\}$ з припущенням, що відоме початкове значення вектора стану $\mathbf{x}(k)$. Як правило, для зручності подальшого аналізу робиться припущення щодо

некорельованості початкового значення $\mathbf{x}(1)$ вектора стану $\mathbf{x}(k)$ з будь-якою реалізацією векторів $\mathbf{w}(k), \mathbf{v}(k)$, тобто:

$$E[\mathbf{x}(1)\mathbf{w}^T(k)] = 0, \quad \forall k \in [1, \dots, T], \quad E[\mathbf{x}(1)\mathbf{v}^T(k)] = 0, \quad \forall k \in [1, \dots, T].$$

З рівняння станів витікає, що $\mathbf{x}(k)$ можна записати як лінійну функцію від $[\mathbf{x}(1), \mathbf{w}(2), \mathbf{w}(3), \dots, \mathbf{w}(k)]$:

$$\begin{aligned} \mathbf{x}(k) &= \mathbf{w}(k) + \mathbf{F}\mathbf{w}(k-1) + \mathbf{F}^2\mathbf{w}(k-2) + \dots + \mathbf{F}^{k-2}\mathbf{w}(2) + \mathbf{F}^{k-1}\mathbf{x}(1), \\ \forall k &= 2, 3, \dots, T. \end{aligned}$$

Таким чином, із $E[\mathbf{w}(k)\mathbf{w}^T(l)] = 0, k \neq l$ і $E[\mathbf{x}(1)\mathbf{w}^T(k)] = 0, k = 1, 2, \dots, T$ випливає, що $\mathbf{w}(k)$ некорельовано із затриманими значеннями \mathbf{x} :

$$E[\mathbf{x}(l)\mathbf{w}^T(k)] = 0 \quad \text{при } l = k-1, k-2, \dots, 1.$$

Аналогічно:

$$\begin{aligned} E[\mathbf{x}(l)\mathbf{v}^T(k)] &= 0 \quad \text{при } l = 1, 2, \dots, T, \\ E[\mathbf{y}(l)\mathbf{v}^T(k)] &= E\{[\mathbf{A}^T\mathbf{u}(l) + \mathbf{H}^T\mathbf{x}(l) + \mathbf{v}(l)]\mathbf{v}^T(k)\} = 0 \quad \text{при } l = k-1, k-2, \dots, 1, \\ E[\mathbf{y}(l)\mathbf{w}^T(k)] &= 0 \quad \text{при } l = k-1, k-2, \dots, 1. \end{aligned}$$

Очевидно, що матриці параметрів $\mathbf{F}, \mathbf{Q}, \mathbf{A}, \mathbf{H}, \mathbf{R}$ можуть бути функціями часу для нестационарних об'єктів та процесів.

Приклад 8.4.1. Розглянемо процес автoreгресії p -го порядку:

$$y(k+1) = a_0 + a_1 y(k) + a_2 y(k-1) + \dots + a_p y(k-p+1) + \varepsilon(k+1), \quad (8.4.1)$$

де $\{\varepsilon(k)\}$ – нормально розподілений процес незалежних величин, скорочено можна записати так: $\{\varepsilon(k)\} \sim HHP(0, \sigma_\varepsilon^2)$, тобто $E[\varepsilon(k)] = 0$, а

$$E[\varepsilon(k)\varepsilon(l)] = \begin{cases} \sigma_\varepsilon^2, & k = l \\ 0, & k \neq l \end{cases}.$$

Процес (8.4.1) можна представити також у вигляді:

$$y(k+1) - \mu = a_1[y(k) - \mu] + a_2[y(k-1) - \mu] + \dots + a_p[y(k-p+1) - \mu] + \varepsilon(k+1), \quad (8.4.2)$$

де μ – відоме середнє значення процесу.

Даний процес можна записати у формі простору станів наступним чином:

Рівняння стану:

$$\begin{bmatrix} y(k+1) - \mu \\ y(k) - \mu \\ y(k-1) - \mu \\ \vdots \\ y(k-p+2) - \mu \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} y(k) - \mu \\ y(k-1) - \mu \\ y(k-2) - \mu \\ \vdots \\ y(k-p+1) - \mu \end{bmatrix} + \begin{bmatrix} \varepsilon(k+1) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (8.4.3)$$

Тобто, вектор стану має вигляд:

$$x(k) = [y(k) - \mu, y(k-1) - \mu, y(k-2) - \mu, \dots, y(k-p+1) - \mu]^T.$$

Рівняння вимірів:

$$z(k) = y(k) = \mu + [1 \ 0 \ \dots \ 0] \begin{bmatrix} y(k) - \mu \\ y(k-1) - \mu \\ \vdots \\ y(k-p+1) - \mu \end{bmatrix}. \quad (8.4.4)$$

В рівнянні вимірів шуми відсутні $\mathbf{R} = 0$, а коваріаційна матриця збурень стану має вигляд:

$$\mathbf{Q} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Приклад 8.4.2. Процес ковзного середнього першого порядку, КС(1):

$$y(k) = \mu + \varepsilon(k) + \beta_1 \varepsilon(k-1), \quad (8.4.5)$$

який може бути представлений у просторі станів наступним чином:

Рівняння стану:

$$\begin{bmatrix} x(k+1) \\ x(k) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon(k) \\ \beta_1 \varepsilon(k-1) \end{bmatrix} + \begin{bmatrix} \varepsilon(k+1) \\ 0 \end{bmatrix}. \quad (8.4.6)$$

Рівняння вимірів:

$$z(k) = y(k) = \mu + [1 \ \ \beta_1] \begin{bmatrix} \varepsilon(k) \\ \varepsilon(k-1) \end{bmatrix}, \quad (8.4.7)$$

тобто,

$$\mathbf{x}(k) = \begin{bmatrix} \varepsilon(k) \\ \varepsilon(k-1) \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{w}(k+1) = \begin{bmatrix} \varepsilon(k+1) \\ 0 \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{z}(k) = y(k), \quad \mathbf{v}(k) = 0, \quad \mathbf{R} = 0.$$

Процес КС(1) можна представити в просторі станів і в такий спосіб:

Рівняння стану:

$$\begin{bmatrix} \varepsilon(k+1) + \beta_1 \varepsilon(k) \\ \beta_1 \varepsilon(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon(k) + \beta_1 \varepsilon(k-1) \\ \beta_1 \varepsilon(k) \end{bmatrix} + \begin{bmatrix} \varepsilon(k+1) \\ \beta_1 \varepsilon(k+1) \end{bmatrix}. \quad (8.4.8)$$

Рівняння вимірів:

$$z(k) = y(k) = \mu + [1 \ \ 0] \begin{bmatrix} \varepsilon(k) + \beta_1 \varepsilon(k-1) \\ \beta_1 \varepsilon(k) \end{bmatrix}. \quad (8.4.9)$$

Приклад 8.4.3. Процес авторегресії з ковзним середнім довільного порядку, АРКС(p, q). В даному випадку вимірність вектора стану буде $n \equiv \max\{p, q+1\}$. В рівнянні

$$\begin{aligned} y(k) - \mu &= a_1[y(k-1) - \mu] + a_2[y(k-2) - \mu] + \dots + a_n[y(k-n) - \mu] + \\ &+ \varepsilon(k) + \beta_1 \varepsilon(k-1) + \beta_2 \varepsilon(k-2) + \dots + \beta_{n-1} \varepsilon(k-n+1) \end{aligned} \quad (8.4.10)$$

$a_j = 0, j > p$ і $\beta_j = 0, j > q$. Розглянемо наступне представлення у просторі станів:

Рівняння стану ($n = \max\{p, q+1\}$):

$$\mathbf{x}(k+1) = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} \varepsilon(k+1) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (8.4.11)$$

Рівняння вимірів:

$$z(k) = y(k) = \mu + [1 \ \beta_1 \ \beta_2 \ \dots \ \beta_{n-1}] \mathbf{x}(k). \quad (8.3.12)$$

Покажемо, що (8.4.10) і (8.4.11), (8.4.12) описують один і той же процес. Нехай $x_j(k)$ – елемент вектора стану $\mathbf{x}(k)$. Відповідно до другого рядка рівняння стану (8.4.11) можна записати

$$x_2(k+1) = x_1(k).$$

З третього рядка випливає, що $x_3(k+1) = x_2(k) = x_1(k-1)$, а з довільного j -го рядка випливає, що

$$x_j(k+1) = L^{j-1} x_1(k+1).$$

Таким чином, для першого рядка рівняння стану справедливе рівняння

$$x_1(k+1) = (a_1 + a_2 L + a_3 L^2 + \dots + a_n L^{n-1}) x_1(k) + \varepsilon(k+1)$$

або

$$(1 - a_1 L - a_2 L^2 - a_3 L^3 - \dots - a_n L^n) x_1(k+1) = \varepsilon(k+1). \quad (8.4.13)$$

З рівняння вимірів можна записати:

$$z(k) = y(k) = \mu + (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_{n-1} L^{n-1}) x_1(k). \quad (8.4.14)$$

Якщо останнє рівняння домножити на многочлен $(1 - a_1 L - a_2 L^2 - \dots - a_n L^n)$ і скористатись рівнянням (8.3.13), то отримаємо:

$$(1 - a_1 L - a_2 L^2 - \dots - a_n L^n)[y(k) - \mu] = (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_{n-1} L^{n-1})\varepsilon(k),$$

що збігається з рівнянням (8.4.10).

Приклад 8.4.4. *Динаміка реальної процентної ставки.* Реальну процентну ставку (РПС) можна розглядати як різницю між номінальною процентною ставкою і очікуваним індексом індексом інфляції $\pi^e(k)$. РПС не спостерігається (тобто, не вимірюється), оскільки неможливо отримати дані щодо індексу інфляції, який очікується на фондовому ринку.

Таким чином, змінна стану в даному випадку є скаляром:

$$x(k) = i(k) - \pi^e(k) - \mu,$$

де μ – середнє значення РПС. В літературі для описання динаміки очікуваної РПС запропонована модель AP(1):

$$x(k+1) = a x(k) + w(k+1), \quad (8.4.15)$$

де $w(k)$ – шум стану. Рівняння вимірів можна записати із врахуванням того, що за минулий час можна зібрати дані, які характеризують відомі значення РПС (номінальна процентна ставка $i(k)$ мінус дійсне значення індексу інфляції $\pi(k)$), тобто, $i(k) - \pi(k)$. Таким чином, рівняння вимірів має вигляд:

$$i(k) - \pi(k) = [i(k) - \pi^e(k)] + [\pi^e(k) - \pi(k)] = \mu + x(k) + v(k), \quad (8.4.16)$$

де $v(k) \equiv [\pi^e(k) - \pi(k)]$ – похибка прогнозу індекса інфляції. Якщо прогнозоване значення індекса інфляції обчислюється за оптимальним алгоритмом, то $v(k)$ не повинно корелюватись їз своїми минулими значеннями і очікуваними значеннями РПС. Таким чином, для даного прикладу матриці моделі в просторі станів мають вигляд:

$$\mathbf{F} = a, \quad z(k) = y(k) = i(k) - \pi(k), \quad \mathbf{H} = 1, \quad v(k) = [\pi^e(k) - \pi(k)].$$

Приклад 8.4.5. Для описання динаміки ділового циклу в модель вводять скаляр $c(k)$, значення якого не спостерігається (тобто, не вимірюється). Припускається, що на множину n різних вимірюваних макроекономічних змінних $[y_1(k), y_2(k), \dots, y_n(k)]$ впливає діловий цикл, а також те, що кожна з цих компонент має свою специфічну структурну

складову, яку ми позначимо через $\chi_i(k)$ і яка не зв'язана із значеннями $y_j(k)$ при $i \neq j$. Якщо діловий цикл і кожна із специфічних структурних складових можуть бути описані за допомогою рівняння AP(1), то отримаємо вектор стану

$$\mathbf{x}(k) = [c(k) \ \chi_1(k) \ \chi_2(k) \dots \chi_n(k)]^T$$

вимірністю $[(n+1) \times 1]$ і рівняння стану:

$$\begin{bmatrix} c(k+1) \\ \chi_1(k+1) \\ \chi_2(k+1) \\ \vdots \\ \chi_n(k+1) \end{bmatrix} = \begin{bmatrix} a_c & 0 & 0 & \dots & 0 \\ 0 & a_1 & 0 & \dots & 0 \\ 0 & 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{bmatrix} \begin{bmatrix} c(k) \\ \chi_1(k) \\ \chi_2(k) \\ \vdots \\ \chi_n(k) \end{bmatrix} + \begin{bmatrix} w_c(k+1) \\ w_1(k+1) \\ w_2(k+1) \\ \vdots \\ w_n(k+1) \end{bmatrix}; \quad (8.4.17)$$

рівняння вимірів:

$$\begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} + \begin{bmatrix} \gamma_1 & 1 & 0 & \dots & 0 \\ \gamma_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \gamma_n & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} c(k) \\ \chi_1(k) \\ \chi_2(k) \\ \vdots \\ \chi_n(k) \end{bmatrix}. \quad (8.4.18)$$

Параметр γ_i описує чутливість i -го ряду до ділового циклу. Для того щоб описати динаміку p -го порядку, величини $c(k)$ і $\chi_i(k)$ у векторі стану замінюють на $(p \times 1)$ -вимірні вектори $[c(k) \ c(k-1) \dots \ c(k-p+1)]^T$ та $[\chi_i(k) \ \chi_i(k-1) \dots \ \chi_i(k-p+1)]^T$. Таким чином вектор стану буде мати вимірність $[(n+1)p \times 1]$. Скаляри a_i в рівняння стану (8.4.17) замінюються матрицями \mathbf{F}_i вимірності $[p \times p]$, а між стовпчиками матриці \mathbf{H}^T в рівнянні вимірів (8.4.18) вставляються блоки нулів вимірностю $[n \times (p-1)]$.

Приклад 8.3.6. Лінійна гаусівська модель в просторі станів:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{d}(k) + \mathbf{T}(k)\mathbf{x}(k) + \mathbf{B}(k)\boldsymbol{\varepsilon}(k), \quad \{\mathbf{x}(k)\} \sim N\{\boldsymbol{\mu}_x, \mathbf{P}\}; \\ \mathbf{x}_1 &\sim N(\mathbf{a}, \mathbf{P}_1); \quad \{\boldsymbol{\varepsilon}(k)\} \sim NID(0, \mathbf{I}); \end{aligned} \quad (8.3.19)$$

$$\mathbf{y}(k) = \mathbf{c}(k) + \mathbf{H}(k)\mathbf{x}(k), \quad (8.4.20)$$

$$\mathbf{z}(k) = \mathbf{y}(k) + \mathbf{G}(k)\varepsilon(k), \quad (8.4.21)$$

де μ_x – вектор середніх значень нормально розподіленої послідовності векторів $\{\mathbf{x}(k)\}$; матриця спостережень вимірності $[N \times n]$ задається векторами $[\mathbf{z}_1, \dots, \mathbf{z}_n]$; вектор стану $\mathbf{x}(k)$ вимірності $[m \times 1]$ містить стохастичний процес, що не спостерігається, а також невідомі ефекти від постійних впливів; $\mathbf{x}(k)$ – послідовність нормально розподілених векторів з середнім значенням μ_x та коваріаційною матрицею \mathbf{P} ; початковий стан \mathbf{x}_1 – нормально розподілений вектор із середнім a та коваріаційною матрицею \mathbf{P}_1 . Рівняння (8.4.19) має так звану марковську структуру, яка є досить ефективною для описання структури послідовної кореляції часового ряду $\mathbf{z}(k)$.

Сигнал $\mathbf{y}(k)$ – деякий проміжний сигнал між вектором стану $\mathbf{x}(k)$ і вектором вимірів $\mathbf{z}(k)$; $\varepsilon(k)$ – нормальну розподілений вектор збурень стану з нульовим середнім значенням та одиничною дисперсією.

Детерміновані матриці $\mathbf{T}(k)$, $\mathbf{B}(k)$, $\mathbf{H}(k)$ і $\mathbf{G}(k)$ називають матрицями системи; вони мають, як правило, розріджену структуру. Вектори $\mathbf{d}(k)$ і $\mathbf{c}(t)$ – частіше всього відомі константи або мають нульові значення.

Якщо система (об'єкт) інваріантний в часі (стаціонарний), то розглянуті вище матриці системи записують без часового аргумента: $\mathbf{T}, \mathbf{B}, \mathbf{H}$ і \mathbf{G} . В такому випадку систему в просторі станів (8.4.19)-(8.4.21) називають інваріантною відносно часу.

Запитання і вправи до восьмого розділу

- Поясніть переваги представлення математичних моделей динамічних систем у просторі станів? Яким чином враховується шум вимірів у моделі АРКС та в моделі, представлений у просторі станів? Яка з цих моделей представляється вам більш адекватною модельованому процесу і чому?
- Поясніть, яку роль відіграють матриці моделі $\mathbf{A}, \mathbf{B}, \mathbf{H}$? Запишіть розмірності цих матриць, якщо розмірність вектора стану дорівнює n .
- Яке співвідношення існує, як правило, між розмірністю вектора стану $\mathbf{x}(k)$ і вектора вимірів $\mathbf{z}(k)$? Поясніть це на прикладах.
- Яка причина існування, в деяких випадках, невимірюваних компонент вектора стану?
- Якщо матриця вимірів \mathbf{H} діагональна, то що це означає по відношенню до вимірів? Поясніть характер взаємозв'язку між векторами стану та вимірів у випадку, коли вона недіагональна?

6. У якому випадку для аналізу процесу можна скористатись моделлю вільної динамічної системи? Наскільки вона зручна для практичного застосування?
7. Наведіть приклади випадкових збурень стану динамічної системи, які можна описати нормальним розподілом?
8. Якщо на динамічну систему впливає випадкове збурення, яке має розподіл, відмінний від нормального, то яким чином це вплине на модель у просторі станів?
9. Запишіть у формі простору станів звичайне диференціальне рівняння третього порядку:

$$a_3 \frac{d^3 y(t)}{dt^3} + a_2 \frac{d^2 y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + y(t) = b_1 u_1(t) + b_2 u_2(t).$$

10. Запишіть у формі простору станів рівняння АРКС(3, 2):

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + a_3 y(k-3) + b_1 \varepsilon(k-1) + b_2 \varepsilon(k-2) + \varepsilon(k).$$

11. Як можна визначити стійкість динамічної системи, яка описується моделлю у просторі станів? Зверніть увагу на відмінності аналізу стійкості в неперервному та дискретному часі.
12. Які труднощі щодо використання моделі в ПС можуть виникнути у випадку, якщо розмірність вектора стану висока ($n = 25-30$)?
13. Запропонуйте структуру моделі у просторі станів для опису сукупності макроекономічних процесів. Яким чином можна знайти оптимальні оцінки стану такої системи? Поясніть природу шумів вимірювань у випадку побудови такої математичної моделі. Опишіть можливі збурення стану, особливо зверніть увагу на збурення для макроекономічних процесів переходного періоду.
14. Для розв'язання яких задач можна скористатись математичною моделлю у просторі станів, побудованою для макроекономіки країни?
15. Виберіть три макроекономічних процеси (наприклад, формування валового внутрішнього продукту, інфляцію та надходження від податку на додану вартість) і побудуйте для них векторно-матричну модель у просторі станів на основі статистичних даних. Визначте особливості цієї моделі у порівнянні із скалярним випадком.

Розділ 9

МЕТОДИКА ПРОГНОЗУВАННЯ ДИНАМІКИ ПРОЦЕСІВ НА ОСНОВІ МАТЕМАТИЧНИХ МОДЕЛЕЙ ЧАСОВИХ РЯДІВ З РІЗНОТЕМПОВОЮ ДИСКРЕТИЗАЦІЄЮ

9.1 Вступ

Розглянемо методику прогнозування вихідних координат динамічних процесів у стохастичному середовищі при дискретизації вхідних збурень з малими періодами квантування T_0 , а вихідних координат з великими періодами h . До указаних процесів можна віднести фінансово- і соціально-економічні процеси, в яких вихідні координати вимірюються в дискретні моменти з періодами квантування, рівними декаді, місяцю, кварталу, півріччю або одному року. В той час періоди квантування фіксованих вхідних збурень відповідно дорівнюють одній добі, тижню або місяцю. У цьому випадку для опису динаміки даних процесів в дискретному часі небхідно застосовувати математичні моделі з різнометровою дискретизацією координат. Динаміка стаціонарних процесів представлена стохастичними моделями авторегресії і ковзного середнього (APKC), а нестаціонарних процесів — моделями авторегресії і інтегрованого ковзного середнього (APIMA) з різнометровою дискретизацією.

9.2 Розробка різнометрової моделі авторегресії і ковзного середнього

9.2.1 Розглянемо вихідну однотемпову модель АРСС $(1, n_e)$

$$(1 + \alpha_1 z^{-1}) y(kT_0) = (1 + b_1 z^{-1} + b_2 z^{-2} + \dots)$$

$$\dots + b_{n_e} z^{-n_e}) \xi(kT_0) + \delta, \quad (9.2.1)$$

де $\{\xi(kT_0)\}$ — послідовність збурень у вигляді дискретного білого шуму з нульовим середнім. Параметр δ визначає рівень процесу. Вибираємо збільшений період квантування h для дискретизації вихідної координати, який визначається співвідношенням

$$h = mT_0, \quad (9.2.2)$$

де T_0 — базовий період квантування, а m — ціле число, більше одиниці.

Розробимо рекурентну процедуру перетворення вихідної моделі (9.2.1) в різнометрову при дискретизації вихідної координати у з періодом квантування h . З приводу цього вихідну модель (9.1) запишемо у різницевій формі:

$$\begin{aligned} y(kT_0) = & -\alpha_1 y[(k-1)T_0] + \xi(kT_0) + b_1 \xi[(k-1)T_0] + \\ & + b_2 \xi[(k-2)T_0] + \dots + b_{n_e} \xi[(k-n_e)T_0] + \delta; \end{aligned}$$

a) при $k=1$ будемо мати:

$$y(1 \cdot T_0) = -\alpha_1 y(0 \cdot T_0) + \xi(1 \cdot T_0) + b_1 \xi(0 \cdot T_0) + \delta;$$

б) при $k=2$ після врахування попереднього кроку:

$$\begin{aligned} y(2 \cdot T_0) = & -\alpha_1 y(1 \cdot T_0) + \xi(2 \cdot T_0) + b_1 \xi(1 \cdot T_0) + b_2 \xi(0 \cdot T_0) + \delta = \\ = & \alpha_1^2 y(0 \cdot T_0) + \xi(2 \cdot T_0) + (b_1 - \alpha_1) \xi(1 \cdot T_0) + \\ & + (b_2 - \alpha_1 b_1) \xi(0 \cdot T_0) + (1 - \alpha_1) \delta; \end{aligned}$$

в) при $k=3$ після врахування попереднього кроку при умові $n_e \geq 3$:

$$\begin{aligned} y(3 \cdot T_0) = & -\alpha_1 y(2 \cdot T_0) + \xi(3 \cdot T_0) + b_1 \xi(2 \cdot T_0) + b_2 \xi(1 \cdot T_0) + \\ & + b_3 \xi(0 \cdot T_0) + \delta = -\alpha_1^3 y(0 \cdot T_0) + \xi(3 \cdot T_0) + (b_1 - \alpha_1) \xi(2 \cdot T_0) + \\ & + (\alpha_1^2 - \alpha_1 b_1 + b_2) \xi(1 \cdot T_0) + (\alpha_1^2 b_1 - \alpha_1 b_2 + b_3) \xi(0 \cdot T_0) + [1 - \alpha_1(1 - \alpha_1)] \delta; \end{aligned}$$

г) при $k=4$ після врахування попереднього кроку при умові $n_e \geq 4$

дістанемо:

$$\begin{aligned} y(4 \cdot T_0) = & -\alpha_1 y(3 \cdot T_0) + \xi(4 \cdot T_0) + b_1 \xi(3 \cdot T_0) + b_2 \xi(2 \cdot T_0) + \\ & + b_3 \xi(1 \cdot T_0) + b_4 \xi(0 \cdot T_0) + \delta = -\alpha_1^4 y(0 \cdot T_0) + \xi(4 \cdot T_0) + \\ & + (b_1 - \alpha_1) \xi(3 \cdot T_0) + (\alpha_1^2 - \alpha_1 b_1 + b_2) \xi(2 \cdot T_0) + \\ & + (\alpha_1^3 + \alpha_1^2 b_1 - \alpha_1 b_2 + b_3) \xi(1 \cdot T_0) + (-\alpha_1^3 b_1 + \end{aligned}$$

$$+ \alpha_1^2 b^2 - \alpha_1 b_3 + b_4) \xi(0 \cdot T_0) + [1 + \alpha_1^2 (1 - \alpha_1) - \alpha_1] \delta.$$

Продовжуючи рекурентну процедуру до $k = m$ за умови, що $m \geq n_\beta$, дістанемо різнотемпову модель в узагальненій різницевій формі:

$$\begin{aligned} y\left[\left[\frac{k}{m}\right]h\right] &= -a_1 y\left[\left(\left[\frac{k}{m}\right]-1\right)h\right] + \xi(kT_0) + c_1 \xi[(k-1)T_0] + \\ &+ c_2 \xi[(k-2)T_0] + \dots + c_m \xi[(k-m)T_0] + a_0, \end{aligned} \quad (9.2.3)$$

де $\left[\frac{k}{m}\right]$ — ціла частина частки від ділення номера дискретного відліку k на коефіцієнт m .

Вираз (9.2.3) можна записати у формі різнотемпової моделі АРКС:

$$A(z_1^{-1})y(rh) = c(z^{-1})\xi(kT_0) + a_0, \quad (9.2.4)$$

де

$$A(z_1^{-1}) = 1 + a_1 z_1^{-1}, \quad (9.2.5)$$

$$c(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_m z^{-m}, \quad (9.2.6)$$

в яких z_1^{-1} — оператор зворотного зсуву на один період квантування h , а z^{-1} — оператор зворотного зсуву на період квантування T_0 .

На основі виразу (9.2.2) співвідношення операторів зворотного зсуву в поліномах (9.2.5), (9.2.6) буде

$$z_1^{-1} = z^{-m}. \quad (9.2.7)$$

Тоді рівність (9.2.5) прийме такий вигляд:

$$A(z^{-1}) = 1 + a_1 z^{-m}. \quad (9.2.8)$$

Вихідна однотемпова модель АРКС відповідно до (9.2.1) є стійкою, тобто корінь полінома $(1 + \alpha_1 z^{-1}) = 0$ $|z| < 1$. При розробці різнотемпової моделі АРСС (9.2.4) в процесі рекурентної процедури необхідно, щоби модель (9.2.4) також була стійкою. Ця умова виконується, якщо корінь полінома $A(z_1^{-1}) = 1 + a_1 z_1^{-1} = 0$ буде знаходитися в середині круга одиничного радіуса. Це виконується при $|a_1| < 1$, тому що корінь дорівнює $z = -a_1$.

Внаслідок того, що вихідна однотемпова модель (9.1) має характеристичне рівняння $A(z^{-1}) = 0$ першого порядку, то її стійкість забезпечується при $|\alpha_1| < 1$. При розробці різнотемпової моделі на основі приведеної рекурентної процедури дістанемо у виразі (9.2.5)

$$a_1 = (-1)^{m-1} \cdot \alpha_1^m.$$

Таким чином, якщо вихідна однотемпова модель (9.2.1) буде стійкою, то різнатемпова модель (9.2.4)–(9.2.6) також буде стійкою. В цьому випадку відліки вихідних координат $y\left[\left[\frac{k}{m}\right]h\right]$ згідно (9.2.3) будуть співпадати з відліками $y(kT_0)$ у вихідній моделі (9.1) при $kT_0 = \left[\frac{k}{m}\right]h$.

9.2.2 Розглянемо вихідну однотемпову модель АРКС (2,2)

$$(1 + \alpha_1 z^{-1} + \alpha_2 z^{-2})y(kT_0) = (1 + b_1 z^{-1} + b_2 z^{-2})\xi(kT_0) + \delta. \quad (9.2.9)$$

Виконаємо рекурентну процедуру перетворення цієї моделі в різнатемпову при дискретизації вихідної координати y з періодом квантування $h = 2 \cdot T_0$. Для цього модель запишемо в різницевій формі:

$$\begin{aligned} y(kT_0) &= -\alpha_1 y[(k-1)T_0] - \alpha_2 y[(k-2)T_0] + \\ &+ \xi(kT_0) + b_1 \xi[(k-1)T_0] + b_2 \xi[(k-2)T_0] + \delta. \end{aligned}$$

На основі рекурентної процедури, приведеної в підрозділі 9.2.1, дістанемо різнатемпову модель при $h = 2T_0$ в різницевій формі:

$$\begin{aligned} y\left[\left[\frac{k}{2}\right]h\right] &= a_1 y\left[\left(\left[\frac{k}{2}\right]-1\right)h\right] + a_2 y\left[\left(\left[\frac{k}{2}\right]-2\right)h\right] + \\ &+ \xi\left[\left[\frac{k}{2}\right]h\right] + c_1 \xi\left[\left[\frac{k}{2}\right]h - T_0\right] + c_2 \xi\left[\left[\frac{k}{2}\right]h - 2T_0\right] + \\ &+ c_3 \xi\left[\left[\frac{k}{2}\right]h - 3T_0\right] + c_4 \xi\left[\left[\frac{k}{2}\right]h - 4T_0\right] + a_0, \end{aligned} \quad (9.2.10)$$

де $\left[\frac{k}{2}\right]$ — ціла частина частки від ділення номера дискретного відліку k на 2.

Коефіцієнти цієї моделі дорівнюють $a_1 = (\alpha_1^2 - 2\alpha_2)$; $a_2 = -\alpha_2^2$; $c_1 = (b_1 - \alpha_1)$; $c_2 = (-\alpha_1 b_1 + \alpha_2 + b_2)$; $c_3 = (\alpha_2 b_1 - \alpha_1 b_2)$; $a_0 = (1 - \alpha_1 + \alpha_2)\delta$.

Координати $y\left[\left[\frac{k}{2}\right]h\right]$ різнатемпової моделі (9.2.10) і $y(kT_0)$ вихідної моделі (9.2.9) будуть співпадати в узлових точках відліку $\left[\frac{k}{2}\right]h$. На основі

запропонованої рекурентної процедури модель (9.2.10) можна перетворити в різнатемпову модель при $h = 4T_0$.

Розглянемо створення різнатемпової моделі при дискретизації вихідної координати y з періодом квантування $h = 3T_0$. Застосовуючи рекурентну процедуру, яка приведена у підрозділі 9.2.1, перетворимо однотемпову модель (9.2.9) у різнатемпову модель АРСС:

$$\begin{aligned} y\left[\left[\frac{k}{3}\right]h\right] &= a'_1 y\left[\left(\left[\frac{k}{3}\right]-1\right)h\right] + a'_2 y\left[\left(\left[\frac{k}{3}\right]-2\right)h\right] + \xi y\left[\left[\frac{k}{3}\right]h\right] + \\ &+ c'_1 \xi \left[\left[\frac{k}{3}\right]h - T_0\right] + c'_2 \xi \left[\left[\frac{k}{3}\right]h - 2T_0\right] + c'_3 \xi \left[\left[\frac{k}{3}\right]h - 3T_0\right] + \\ &+ c'_4 \xi \left[\left[\frac{k}{3}\right]h - 4T_0\right] + c'_5 \xi \left[\left[\frac{k}{3}\right]h - 5T_0\right] + c'_6 \xi \left[\left[\frac{k}{3}\right]h - 6T_0\right] + a'_0, \end{aligned}$$

де $\left[\frac{k}{3}\right]$ — ціла частина частки від ділення номера дискретного відліку k на 3.

Коефіцієнти цієї моделі дорівнюють: $a'_1 = (-\alpha_1^3 + 3\alpha_1\alpha_2)$; $a'_2 = -\alpha_2^3$; $c'_1 = (b_1 - \alpha_1)$; $c'_2 = -(\alpha_1^2 - \alpha_1 b_1 - \alpha_2 + b_2)$; $c'_3 = -(\alpha_1^2 b_1 - \alpha_1 \alpha_2 - \alpha_2 b_1 - \alpha_1 b_2)$; $c'_4 = (\alpha_2^2 - \alpha_1 \alpha_2 b_1 + \alpha_1^2 b_2 - \alpha_2 b_2)$; $c'_5 = (-\alpha_1 \alpha_2 b_2 + \alpha_2^2 b_1)$; $c'_6 = (\alpha_2^2 b_2)$; $a'_0 = (\alpha_1^2 + \alpha_2^2 - \alpha_1 - \alpha_2 + 1)\delta$.

На основі запропонованої методики вищепереліченої моделі можна перетворити в різнатемпову з подвоєною частотою при $h = 6T_0$ або з потроєною частотою при $h = 9T_0$.

У загальному випадку для вихідної однотемпової моделі АРКС (2,2) типу (9.2.9) різнатемпова модель при $h = mT_0$ буде мати такий вигляд:

$$\begin{aligned} y\left[\left[\frac{k}{m}\right]h\right] &= a'_1 y\left[\left(\left[\frac{k}{m}\right]-1\right)h\right] + a'_2 y\left[\left(\left[\frac{k}{m}\right]-2\right)h\right] + \xi \left[\left[\frac{k}{3}\right]h\right] + \\ &+ c'_1 \xi \left[\left[\frac{k}{m}\right]h - T_0\right] + c'_2 \xi \left[\left[\frac{k}{m}\right]h - 2T_0\right] + \dots c'_{2_m} \xi \left[\left[\frac{k}{m}\right]h - 2T_0\right] + a'_0. \end{aligned}$$

9.2.3 Загальний метод розробки різнатемпової моделі на основі однотемпової моделі АРКС (p, q)

Розглянемо однотемпову модель АРКС (p, q):

$$A(z^{-1})y(z) = c(z^{-1})\xi(z), \quad (9.2.11)$$

де y , ξ — відповідно вихідна і вхідна координата моделі АРКС з малим періодом дискретизації T_0 , тобто $z^{-1}y(kT_0) = y[(k-1)T_0]$, $z^{-1}u(kT_0) = u[(k-1)T_0]$. Визначимо також передаточну функцію для цієї системи

$$W(z^{-1}) = \frac{y(z)}{u(z)} = \frac{C(z^{-1})}{A(z^{-1})}, \text{ де корені поліномів } A(z^{-1}), C(z^{-1}) \text{ відносно оператора}$$

z за модулем менші від одиниці. Нехай тепер для дискретизації вихідної координати вибраний збільшений період квантування (9.2.2) $h = mT_0$, а вхідна координата залишається з малим періодом T_0 . Необхідно виконати перетворення моделі (9.2.11) у різnotемпову так, щоби значення вихідної координати у вузлових точках при великому періоді квантування h збігалися з відповідними значеннями у вузлових точках при малому періоді квантування T_0 .

Припустимо, що знайдені корені $\delta_1, \dots, \delta_p$ поліному $A(z^{-1})$ (p -го порядку) відносно z , є полюси передаточної функції $w(z^{-1})$. Тоді $A(z^{-1})$ можна представити у формі $A(z^{-1}) = \prod_{i=1}^p (1 - \delta_i z^{-1})$. Вводимо спеціальний поліном [50]:

$$T(z^{-1}) = \prod_{i=1}^p (1 + \delta_i z^{-1} + \delta_i^2 z^{-2} + \dots \delta_i^{m-1} z^{-m+1})$$

та помножимо на нього ліву і праву частини рівності (9.2.11):

$$T(z^{-1})A(z^{-1})y(z) = T(z^{-1})c(z^{-1})\xi(z).$$

При цьому врахуємо очевидний факт, що $(1 - \delta_i z^{-1})(1 + \delta_i z^{-1} + \delta_i^2 z^{-2} + \dots \delta_i^{m-1} z^{-m+1}) = 1 - \delta_i^m z^{-m}$, тоді дістанемо:

$$T(z^{-1})A(z^{-1}) = \prod_{i=1}^p (1 - \delta_i^m z^{-m}).$$

У лівій частині цього виразу усі степені z будуть кратними m . Іншими словами, вихідна координата тепер дискретизується з періодом $h = mT_0$. Позначимо, відповідно до (9.7), $z_1^{-1} = z^{-m}$, $z^{-m}y(kT_0) = z_1^{-1}y(rh) = y[(k-m)T_0] = y[(r-1)h]$, де $r = \left[\frac{k}{m} \right]$. У результаті отримаємо таку різnotемпову модель АРКС:

$$\tilde{A}(z_1^{-1})y(z_1) = \tilde{c}(z^{-1})u(z), \quad (9.2.12)$$

$$\text{де } \tilde{A}(z_1^{-1}) = \prod_{i=1}^p (1 - \delta_i^m z^{-m}), \tilde{c}(z^{-1}) = T(z^{-1})c(z^{-1}).$$

Розглянемо передаточну функцію різnotемпової системи $\tilde{w}(z_1^{-1}, z^{-1}) =$

$= \frac{y(z_1)}{u(z)} = \frac{\tilde{c}(z_1)}{\tilde{A}(z_1^{-1})}$. Відзначимо її повну еквівалентність початкової передаточній функції $w(z^{-1})$ у вузлових точках при періоді h :

$$\begin{aligned}\tilde{w}(z_1^{-1}, z^{-1}) &= \frac{\tilde{c}(z^{-1})}{\tilde{A}(z_1^{-1})} = \frac{T(z^{-1})c(z^{-1})}{\prod_{i=1}^p (1 - \delta_i^m z_1^{-1})} = \frac{T(z^{-1})c(z^{-1})}{\prod_{i=1}^p (1 - \delta_i^m z^{-m})} = \\ &= \frac{T(z^{-1})c(z^{-1})}{T(z^{-1})A(z^{-1})} = \frac{c(z^{-1})}{A(z^{-1})} = W(z^{-1}).\end{aligned}$$

9.3 Прогнозування одновимірних стаціонарних процесів на основі моделей авторегресії і ковзного середнього з різнометровою дискретизацією

9.3.1 Розробка функції прогнозування на один період квантування h

При врахуванні співвідношення (9.2.8) запишемо рівняння (9.2.4) відносно $y(rh)$:

$$y(rh) = \frac{C(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{a_0}{A(z^{-1})}. \quad (9.3.1)$$

Виконаємо зсув вперед за часом $h = mT_0$. Тоді на основі (9.2.7) будемо мати:

$$y[(r+1)] = y[(k+m)T_0] = z^m \frac{C(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{a_0}{A(z^{-1})} \quad (9.3.2)$$

Для розділення впливу вхідних збурень, які діяли до поточного часу і на майбутній час, складемо діофантове рівняння [23, 48]:

$$C(z^{-1}) = A(z^{-1})L(z^{-1}) + z^{-m}F(z^{-1}), \quad (9.3.3)$$

на основі якого рівність (9.3.2) можна записати таким чином:

$$\begin{aligned}y[(r+1)h] = y[(k+m)T_0] &= z^m L(z^{-1}) \xi(kT_0) + \\ &+ \frac{F(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{1}{A(z^{-1})} a_0.\end{aligned} \quad (9.3.4)$$

При $L(z^{-1}) = C(z^{-1})$ із діофантового рівняння (9.3.3) при врахуванні (9.2.8) однозначно визначається поліном:

$$F(z^{-1}) = -a_1 C(z^{-1}) = -a_1(1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_m z^{-m}).$$

Позаяк порядок полінома $L(z^{-1})$ дорівнює m , то перший член у правій частині рівності (9.3.4) представляє майбутню помилку прогнозування на один період квантування h :

$$e[(r+1)h] = e[(k+m)T_0] = z^m L(z^{-1}) \xi(kT_0). \quad (9.3.5)$$

Тоді на основі (9.3.4), (9.3.5) можна записати функцію прогнозування вихідної координати у на один період квантування h :

$$y^*[(r+1)h | rh] = \frac{F(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{a_0}{A(z^{-1})} = y^*[(k+m)T_0 | kT_0]. \quad (9.3.6)$$

При цьому вважаємо, що збурення $\zeta(kT_0)$ вимірюється і вважається відомим до моменту у часі kT_0 включно.

9.3.2 Розробка функції прогнозування на p періодів квантування h

При врахування співвідношення (9.2.2) виконуємо зсів уперед за часом на $ph = pmT_0$ у рівності (9.3.1):

$$y[(r+p)h] = y[(k+pm)T_0] = z^{pm} \frac{C(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{a_0}{A(z^{-1})}, \quad (9.3.7)$$

де $p = 2, 3, \dots$. При цьому вважаємо, що $a_0 = \text{const}$ упродовж ph .

Складемо діофантове рівняння:

$$C(z^{-1}) = A(z^{-1})L(z^{-1}) + z^{-pm} F_p(z^{-1}), \quad (9.3.8)$$

на основі якого вираз (9.3.7) можна перетворити до такого вигляду:

$$\begin{aligned} y[(r+p)h] &= y[(k+pm)T_0] = z^{pm} L(z^{-1}) \xi(kT_0) + \\ &+ \frac{F_p(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{a_0}{A(z^{-1})}. \end{aligned}$$

Тоді функція прогнозування на p періодів квантування h буде аналогічною (9.3.6), а саме:

$$y^*[(r+p)h | rh] = y^*[(k+pm)T_0 | kT_0] = \frac{F_p(z^{-1})}{A(z^{-1})} \xi(kT_0) + \frac{a_0}{A(z^{-1})}. \quad (9.3.9)$$

а помилка прогнозування:

$$e[(r+p)h] = e[(k+pm)T_0] = z^{pm} L(z^{-1}) \xi(kT_0).$$

Визначення поліному $F_p(z^{-1})$ розглянемо на основі різотемпової моделі процесу (9.2.4), (9.2.5), (9.2.8):

$$(1 + a_1 z^{-m}) y(rh) = C(z^{-1}) \xi(kT_0) + a_0. \quad (9.3.10)$$

При прогнозуванні на $2m$ періодів T_0 при $p=2$ діофантове рівняння (9.3.8) запищеться у вигляді

$$C(z^{-1}) = (1 + a_1 z^{-m}) L(z^{-1}) + z^{-2m} F_p(z^{-1}).$$

Вибираємо поліном $L(z^{-1}) = (1 - a_1 z^{-m}) \cdot C(z^{-1})$. Тоді з діофантового рівняння однозначно визначається $F_p(z^{-1}) = a_1^2 C(z^{-1})$.

При прогнозуванні на $3m$ періодів T_0 при $p=3$ запишемо діофантове рівняння

$$C(z^{-1}) = (1 + a_1 z^{-m}) L(z^{-1}) + z^{-3m} F_p(z^{-1}).$$

При виборі поліному $L(z^{-1}) = (1 - a_1 z^{-m} + a_1^2 z^{-2m}) C(z^{-1})$ поліном $F_p(z^{-1})$ однозначно визначається $F_p(z^{-1}) = -a_1^3 C(z^{-1})$.

При прогнозуванні на $4m$ періодів T_0 при $p=4$ діофантове рівняння (9.3.8) буде

$$C(z^{-1}) = (1 + a_1 z^{-m}) L(z^{-1}) + z^{-4m} F_p(z^{-1}).$$

Вибираємо поліном $L(z^{-1}) = (1 - a_1 z^{-m} + a_1^2 z^{-2m} - a_1^3 z^{-3m}) C(z^{-1})$, тоді поліном $F_p(z^{-1}) = a_1^4 C(z^{-1})$.

Таким чином, при прогнозуванні вихідної координати процесу (9.3.10) на p періодів квантування h відповідно до функції (9.3.9) поліном $F_p(z^{-1})$ буде визначатися однозначно з діофантового рівняння (9.3.8) в такому вигляді:

$$F_p(z^{-1}) = (1)^p a_1^p C(z^{-1}). \quad (9.3.11)$$

На основі (9.3.7), (9.3.9) визначимо дисперсію помилки прогнозування при $a_0 = 0$:

$$\begin{aligned} M\{y[(r+p)h] - y^*[(r+p)h | rh]\}^2 &= \\ &= M\left\{z^{pm} \frac{C(z^{-1})}{A(z^{-1})} \xi(kT_0) - \frac{F_p(z^{-1})}{A(z^{-1})} \xi(kT_0)\right\}^2 = \end{aligned}$$

$$\begin{aligned}
&= M \left\{ z^{pm} \frac{[A(z^{-1})L(z^{-1}) + z^{-pm}F_p(z^{-1})]}{A(z^{-1})} \xi(kT_0) - \right. \\
&\quad \left. - \frac{F_p(z^{-1})}{A(z^{-1})} \xi(kT_0) \right\}^2 = M \{L(z^{-1})\xi[(k+pm)T_0]\}^2 + \\
&\quad \left. \frac{F_p(z^{-1})}{A(z^{-1})} \xi(kT_0) - \frac{F_p(z^{-1})}{A(z^{-1})} \xi(kT_0) \right\}^2 = M \{L(z^{-1})\xi[(k+pm)T_0]\}^2,
\end{aligned}$$

де M — оператор математичного сподівання.

Таким чином, використання функції прогнозування (9.3.9) забезпечує мінімізацію дисперсії помилки прогнозування вихідної координати у на p періодів квантування h .

Приклад 9.1. Динаміка процесу зображена однотемповою моделлю АРСС (1,1):

$$(1 + \alpha_1 z^{-1})y(kT_0) = (1 + b_1 z^{-1})\xi(kT_0),$$

де $\alpha_1 = -0,9802$; $b_1 = -0,5$. Ця модель розроблена з неперервної моделі першого порядку з постійною часу $T = 800$ с при вибраному базовому періоді квантування $T_0 = 16$ с. Необхідно розробити процедуру прогнозування координати у на період квантування $h = 5T_0 = 80$ с.

Запишемо вихідну модель у різницевій формі:

$$y(kT_0) = -\alpha_1 y[(k-1)T_0] + \xi(kT_0) + b_1 \xi[(k-1)T_0].$$

На основі рекурентної процедури, приведеної в підрозд. 9.2.1, при $k = 5$ дістанемо різнатемпову модель в узагальненій різницевій формі:

$$\begin{aligned}
y\left[\left[\frac{k}{5}\right]h\right] &= -\alpha_1^5 y\left[\left(\left[\frac{k}{5}\right]-1\right)h\right] + \xi(kT_0) + \\
&\quad + (b_1 - \alpha_1) \xi[(k-1)T_0] - \alpha_1(b_1 - \alpha_1) \xi[(k-2)T_0] + \\
&\quad + \alpha_1^2(b_1 - \alpha_1) \xi[(k-3)T_0] - \alpha_1^3(b_1 - \alpha_1) \xi[(k-4)T_0] + \alpha_1^4 b_1 \xi[(k-5)T_0], \quad (9.3.12)
\end{aligned}$$

де $\left[\frac{k}{5}\right]$ — ціла частина частки від ділення номера відліку k на 5. При підстановці коефіцієнтів α_1 , b_1 ця модель прийме вигляд:

$$\begin{aligned}
y\left[\left[\frac{k}{5}\right]h\right] &= 0,9048 y\left[\left(\left[\frac{k}{5}\right]-1\right)h\right] + \xi(kT_0) + \\
&\quad + 0,4802 \xi[(k-1)T_0] + 0,47069 \xi[(k-2)T_0] +
\end{aligned}$$

$$+ 0,46137\xi[(k-3)T_0] + 0,45223\xi[(k-4)T_0] - 0,46156\xi[(k-5)T_0].$$

Запишемо цей вираз у формі різнометрової моделі АРСС:

$$A(z^{-1})y\left[\left[\frac{k}{5}\right]h\right] = C(z^{-1})\xi(kT_0),$$

$$\text{де } A(z^{-1}) = 1 - 0,9048z^{-5}; \quad C(z^{-1}) = (1 + 0,4802z^{-1} + 0,47069z^{-2} + 0,46137z^{-3} + \\ + 0,45223z^{-4} - 0,46156z^{-5}).$$

Складемо діофантове рівняння:

$$C(z^{-1}) = A(z^{-1})L(z^{-1}) + z^{-5}F(z^{-1}),$$

$$\text{з якого при } L(z^{-1}) = C(z^{-1}) \text{ визначаємо поліном } F(z^{-1}) = 0,9048 + 0,4344z^{-1} + \\ + 0,42588z^{-2} + 0,41744z^{-3} + 0,40917z^{-4} - 0,41762z^{-5}.$$

Тоді на основі (9.3.6) можна виконати прогнозування вихідної координати на $1h = 5T_0$ таким чином:

$$\begin{aligned} y^*[(r+1)h | rh] &= y^*[(k+5)T_0 | kT_0] = \\ &= 0,9048y(rh) + 0,9048\xi(kT_0) + 0,4344\xi[(k-1)T_0] + \\ &\quad + 0,42588\xi[(k-2)T_0] + 0,41744\xi[(k-3)T_0] + \\ &\quad + 0,40917\xi[(k-4)T_0] - 0,41762\xi[(k-5)T_0]. \end{aligned} \quad (9.3.13)$$

Результати цифрового моделювання різнометрової моделі (9.3.12) та прогнозованого значення $y^*[(r+1)h | rh]$ згідно функції прогнозування (9.3.13) приведені на рис. 9.1 при подачі на вхід додаткового одиничного ступінчастого збурення.

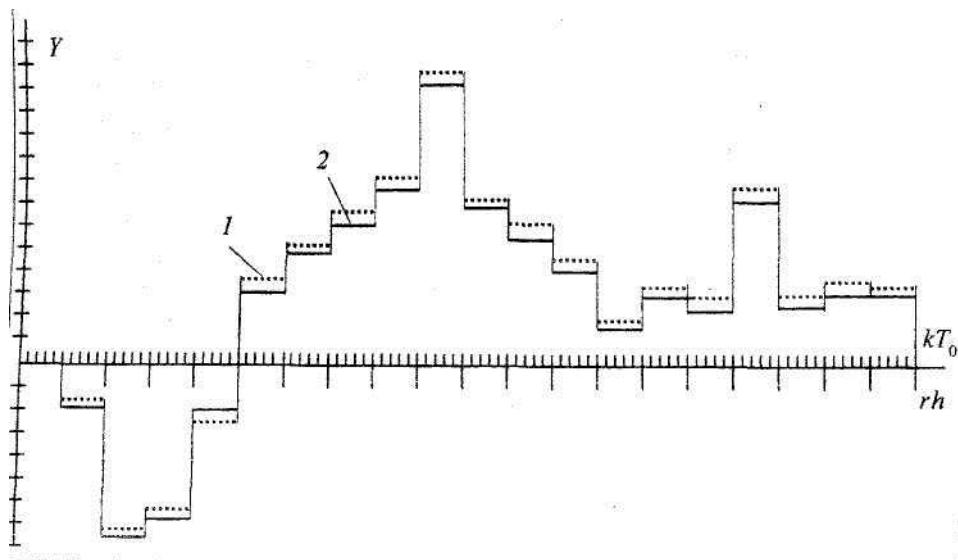


Рис. 9.1. Графіки моделювання вихідної змінної (1) та її прогнозованого значення (2)

різнометрової одновимірної системи

9.4 Прогнозування багатовимірних стаціонарних процесів з різнометровою дискретизацією

Різнометрова багатовимірна дискретна модель АРСС представлена у вигляді

$$\begin{aligned} & \begin{bmatrix} A_{11}(z_1^{-1}) & 0 & \dots & 0 \\ 0 & A_{22}(z_2^{-1}) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & A_{nn}(z_n^{-1}) \end{bmatrix} \begin{bmatrix} y_1(r_1 h_1) \\ y_2(r_2 h_2) \\ \vdots \\ y_n(r_n h_n) \end{bmatrix} = \\ & = \begin{bmatrix} C_{11}(z^{-1}) & C_{12}(z^{-1}) & \dots & C_{1n}(z^{-1}) \\ C_{21}(z^{-1}) & C_{22}(z^{-1}) & \dots & C_{2n}(z^{-1}) \\ \vdots & \vdots & \vdots & \vdots \\ C_{n1}(z^{-1}) & C_{n2}(z^{-1}) & \dots & C_{nn}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \xi_2(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} + \begin{bmatrix} a_{01} \\ a_{02} \\ \vdots \\ a_{04} \end{bmatrix}. \quad (9.4.1) \end{aligned}$$

При цьому співвідношення періодів квантування для дискретних відліків вихідних координат буде таке:

$$h_i = m_i T_0, \quad i = 1, 2, \dots, n, \quad (9.4.2)$$

де m_i — ціле число, більше від одиниці. Тоді співвідношення операторів зворотного зсуву:

$$z_i^{-1} = z^{-m_i}, \quad (9.4.3)$$

де z^{-1} — оператор зворотного зсуву на один період квантування T_0 ; z_i^{-1} — оператор зворотного зсуву на один період h_i . Структура поліномів в моделі (9.4.1) має вигляд:

$$A_{ii}(z_i^{-1}) = 1 + a_{1i} z_i^{-1}; \quad (9.4.4)$$

$$C_{ij}(z^{-1}) = 1 + c_{1ij} z^{-1} + c_{2ij} z^{-2} + \dots + c_{m_{ij}} z^{-m},$$

$$i = 1, 2, \dots, n; \quad j = 1, 2, \dots, n. \quad (9.4.5)$$

При цьому припускаємо, що $\xi_i(kT_0)$ — це дискретний білий шум, а a_{0i} — зміщення i -ї вихідної координати, яке дорівнює $A_{ii}(1) \cdot y_{i_{cep}}$.

На основі рівності (9.4.3) динамічні поліноми $A_{ii}(z_i^{-1})$ можна записати через базовий оператор z таким чином:

$$A_{ii}(z^{-1}) = 1 + a_{1i} z^{-m_i}. \quad (9.4.6)$$

Розв'яжемо дискретну модель (9.4.1) відносно $y_i(r_i h_i)$:

$$\begin{aligned} \begin{bmatrix} y_1(r_1 h_1) \\ \vdots \\ y_n(r_n h_n) \end{bmatrix} &= \begin{bmatrix} 1 & \cdots & 0 \\ \frac{1}{A_{11}(z_1^{-1})} & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \begin{bmatrix} C_{11}(z^{-1}) & \cdots & C_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ C_{n1}(z^{-1}) & \cdots & C_{nn}(z^{-1}) \end{bmatrix} \times \\ &\times \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} + \begin{bmatrix} 1 & \cdots & 0 \\ \frac{1}{A_{11}(z_1^{-1})} & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix}. \end{aligned} \quad (9.4.7)$$

Перемножимо ліву і праву частину на $\text{diag}\{z^{p_i m_i}\}$:

$$\begin{aligned} \begin{bmatrix} y_1[(r_1 + p_1)h_1] \\ \vdots \\ y_n[(r_n + p_n)h_n] \end{bmatrix} &= \begin{bmatrix} z^{p_1 m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{p_n m_n} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 0 \\ \frac{1}{A_{11}(z_1^{-1})} & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times \\ &\times \begin{bmatrix} C_{11}(z^{-1}) & \cdots & C_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ C_{n1}(z^{-1}) & \cdots & C_{nn}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} + \\ &+ \begin{bmatrix} 1 & \cdots & 0 \\ \frac{1}{A_{11}(z_1^{-1})} & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \begin{bmatrix} a_{01} \\ \vdots \\ a_{0n} \end{bmatrix}. \end{aligned} \quad (9.4.8)$$

З урахуванням (9.4.3), при якому $A_{ii}(z^{-1})$ визначається відповідно до (9.4.6), складемо таке матричне діофантове рівняння:

$$\begin{bmatrix} C_{11}(z^{-1}) & \cdots & C_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ C_{n1}(z^{-1}) & \cdots & C_{nn}(z^{-1}) \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \frac{1}{A_{11}(z_1^{-1})} & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times$$

$$\begin{aligned}
& \times \begin{bmatrix} L_{11}(z^{-1}) & \cdots & L_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ L_{n1}(z^{-1}) & \cdots & L_{nn}(z^{-1}) \end{bmatrix} + \begin{bmatrix} z^{p_1 m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{p_n m_n} \end{bmatrix} \times \\
& \times \begin{bmatrix} F_{11}(z^{-1}) & \cdots & F_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ F_{n1}(z^{-1}) & \cdots & F_{nn}(z^{-1}) \end{bmatrix}. \tag{9.4.9}
\end{aligned}$$

За допомогою (9.4.9) та з урахуванням (9.4.6) визначимо першу складову правої частини моделі (9.4.8):

$$\begin{aligned}
& \begin{bmatrix} z^{p_1 m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{p_n m_n} \end{bmatrix} \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times \\
& \times \begin{bmatrix} C_{11}(z^{-1}) & \cdots & C_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ C_{n1}(z^{-1}) & \cdots & C_{nn}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} = \\
& = \begin{bmatrix} z^{p_1 m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{p_n m_n} \end{bmatrix} \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times \\
& \times \left\{ \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \begin{bmatrix} L_{11}(z^{-1}) & \cdots & L_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ L_{n1}(z^{-1}) & \cdots & L_{nn}(z^{-1}) \end{bmatrix} + \right. \\
& \left. + \begin{bmatrix} z^{p_1 m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{p_n m_n} \end{bmatrix} \begin{bmatrix} F_{11}(z^{-1}) & \cdots & F_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ F_{n1}(z^{-1}) & \cdots & F_{nn}(z^{-1}) \end{bmatrix} \right\} \times \\
& \times \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} = \begin{bmatrix} z^{p_1 m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{p_n m_n} \end{bmatrix} \begin{bmatrix} L_{11}(z^{-1}) & \cdots & L_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ L_{n1}(z^{-1}) & \cdots & L_{nn}(z^{-1}) \end{bmatrix} \times
\end{aligned}$$

$$\begin{aligned}
& \times \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} + \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times \\
& \times \begin{bmatrix} F_{11}(z^{-1}) & \cdots & F_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ F_{n1}(z^{-1}) & \cdots & F_{nn}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix}. \tag{9.4.10}
\end{aligned}$$

Перша складова в правій частині цього виразу представляє вектор майбутніх помилок $\{e_1[(k + p_1 m_1)T_0], \dots, e_n[(k + p_n m_n)T_0]\}^T$.

Якщо вираз (9.4.10) підставити в рівняння (9.4.8), то дістанемо майбутнє значення вихідних координат:

$$\begin{aligned}
& \begin{bmatrix} y_1(r_1 h_1) \\ \vdots \\ y_n(r_n h_n) \end{bmatrix} = \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times \\
& \times \begin{bmatrix} F_{11}(z^{-1}) & \cdots & F_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ F_{n1}(z^{-1}) & \cdots & F_{nn}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} + \\
& + \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \begin{bmatrix} a_{01} \\ \vdots \\ a_{0n} \end{bmatrix} + \begin{bmatrix} e_1[(k + p_1 m_1)T_0] \\ \vdots \\ e_n[(k + p_n m_n)T_0] \end{bmatrix} = \\
& = \begin{bmatrix} y_1[(r_1 + p_1)h_1 | r_1 h_1] \\ \vdots \\ y_n[(r_n + p_n)h_n | r_n h_n] \end{bmatrix} + \begin{bmatrix} e_1[(k + p_1 m_1)T_0] \\ \vdots \\ e_n[(k + p_n m_n)T_0] \end{bmatrix},
\end{aligned}$$

де прогнозовані значення вихідні координат будуть дорівнювати:

$$\begin{bmatrix} y_1[(r_1 + p_1)h_1 | r_1 h_1] \\ \vdots \\ y_n[(r_n + p_n)h_n | r_n h_n] \end{bmatrix} = \begin{bmatrix} \frac{1}{A_{11}(z_1^{-1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \times$$

$$\begin{aligned}
& \times \begin{bmatrix} F_{11}(z^{-1}) & \cdots & F_{1n}(z^{-1}) \\ \vdots & \ddots & \vdots \\ F_{n1}(z^{-1}) & \cdots & F_{nn}(z^{-1}) \end{bmatrix} \times \begin{bmatrix} \xi_1(kT_0) \\ \vdots \\ \xi_n(kT_0) \end{bmatrix} + \\
& + \begin{bmatrix} 1 & \cdots & 0 \\ \frac{1}{A_{11}(z_1^{-1})} & \ddots & \vdots \\ \vdots & \ddots & 1 \\ 0 & \cdots & \frac{1}{A_{nn}(z_n^{-1})} \end{bmatrix} \begin{bmatrix} a_{01} \\ \vdots \\ a_{0n} \end{bmatrix}. \tag{9.4.11}
\end{aligned}$$

У цьому виразі кожну складову $y_j^*[(r_j + p_j)h_j | r_j h_j]$ на основі (9.4.3), (9.4.4) необхідно обчислювати за такою рекурентною процедурою:

$$\begin{aligned}
y_j^*[(r_j + p_j)h_j | r_j h_j] = & -a_1 y_j^*[(r_j + p_j)h_j | r_j h_j] + \\
& + F_{j1}(z^{-1})\xi_1(kT_0) + \dots + F_{jn}(z^{-1})\xi_n(kT_0) + a_{0j}.
\end{aligned}$$

Розглянемо визначення матричного поліному $\begin{bmatrix} F_{11}(z^{-1}) & \cdots & F_{1n}(z^{-1}) \\ \vdots & & \vdots \\ F_{n1}(z^{-1}) & \cdots & F_{nn}(z^{-1}) \end{bmatrix}$.

Запишемо матричне діофантове рівняння (9.4.9) у формі одновимірних діофантових рівнянь:

$$C_{ij}(z^{-1}) = A_{ii}(z^{-1})L_{ij}(z^{-1}) + z^{-p_i m_i} F_{ij}(z^{-1})$$

при $i = 1, \dots, n; j = 1, \dots, n$.

Ці рівняння розв'язуються відносно поліномів $F_{ij}(z^{-1})$ при заданих p_i і поліномах $C_{ij}(z^{-1}), A_{ii}(z^{-1})$ відповідно до методики, викладеної в підрозд. 9.3.2, тобто

$$F_{ij}(z^{-1}) = (-1)^{p_i} a_{1ii}^{p_i} C_{ij}(z^{-1}). \tag{9.4.12}$$

Приклад 9.2. Однотемпова математична модель АРСС двовимірного динамічного процесу має вигляд:

$$\begin{bmatrix} A_{11}(z^{-1}) & 0 \\ 0 & A_{22}(z^{-1}) \end{bmatrix} \begin{bmatrix} y_1(kT_0) \\ y_2(kT_0) \end{bmatrix} = \begin{bmatrix} B_{11}(z^{-1}) & B_{12}(z^{-1}) \\ B_{21}(z^{-1}) & B_{22}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \xi_2(kT_0) \end{bmatrix}, \tag{9.4.13}$$

де $A_{11}(z^{-1}) = (1 + \alpha_1 z^{-1}) = (1 - 0,9802 z^{-1}); A_{22}(z^{-1}) = (1 + \alpha_2 z^{-1}) = (1 - 0,99 z^{-1}); B_{11}(z^{-1}) = (1 + b_{11} z^{-1}) = (1 - 0,5 z^{-1}); B_{12}(z^{-1}) = (1 + b_{12} z^{-1}) = (1 + 0,6 z^{-1}); B_{21}(z^{-1}) = (1 + b_{21} z^{-1}) = (1 - 0,7 z^{-1}); B_{22}(z^{-1}) = (1 + b_{22} z^{-1}) =$

$$= (1 - 0,45z^{-1}).$$

При розробці даної моделі застосований базовий період квантування $T_0 = 16$ с.

Постановка задачі. 1). Розробити різнатемпову модель при дискретизації координати y_1 з періодом квантування $h_1 = m_1 T_0 = 80$ с, а координати y_2 з періодом квантування $h_2 = m_2 T_0 = 160$ с, де $m_1 = 5$, $m_2 = 10$.

2). Виконати процедуру прогнозування координати y_1 на період квантування h_1 , а координати y_2 — на період h_2 вперед.

Розв'язання. На основі рекуррентної процедури, приведеної в підрозд. 9.2.1, з вихідної моделі (9.4.13) розроблено різнатемпова модель процесу

$$\begin{aligned} & \begin{bmatrix} A_{11}(z^{-1}) & 0 \\ 0 & A_{22}(z^{-1}) \end{bmatrix} \begin{bmatrix} y_1 \left[\left[\frac{k}{5} \right] h_1 \right] \\ y_2 \left[\left[\frac{k}{10} \right] h_2 \right] \end{bmatrix} = \\ & = \begin{bmatrix} C_{11}(z^{-1}) & C_{12}(z^{-1}) \\ C_{21}(z^{-1}) & C_{22}(z^{-1}) \end{bmatrix} \begin{bmatrix} \xi_1(kT_0) \\ \xi_2(kT_0) \end{bmatrix}, \end{aligned} \quad (9.4.14)$$

$$\text{де } A_{11}(z^{-1}) = 1 + \alpha_1^5 z^{-5}; \quad A_{22}(z^{-1}) = 1 - \alpha_2^{10} z^{-10};$$

$$C_{11}(z^{-1}) = 1 + (b_{11} - \alpha_1)z^{-1} - \alpha_1(b_{11} - \alpha_1)z^{-2} +$$

$$+ \alpha_1^2(b_{11} - \alpha_1)z^{-3} - \alpha_1^3(b_{11} - \alpha_1)z^{-4} + \alpha_1^4 b_{11} z^{-5};$$

$$C_{12}(z^{-1}) = 1 + (b_{12} - \alpha_1)z^{-1} - \alpha_1(b_{12} - \alpha_1)z^{-2} +$$

$$+ \alpha_1^2(b_{12} - \alpha_1)z^{-3} - \alpha_1^3(b_{12} - \alpha_1)z^{-4} + \alpha_1^4 b_{12} z^{-5};$$

$$C_{21}(z^{-1}) = 1 + (b_{21} - \alpha_2)z^{-1} - \alpha_2(b_{21} - \alpha_2)z^{-2} +$$

$$+ \alpha_2^2(b_{21} - \alpha_2)z^{-3} - \alpha_2^3(b_{21} - \alpha_2)z^{-4} + \alpha_2^4(b_{21} - \alpha_2)z^{-5} - \alpha_2^5(b_{21} - \alpha_2)z^{-6} +$$

$$\alpha_2^6(b_{21} - \alpha_2)z^{-7} - \alpha_2^7(b_{21} - \alpha_2)z^{-8} + \alpha_2^8(b_{21} - \alpha_2)z^{-9} - \alpha_2^9 b_{21} z^{-10};$$

$$C_{22}(z^{-1}) = 1 + (b_{22} - \alpha_2)z^{-1} - \alpha_2(b_{22} - \alpha_2)z^{-2} +$$

$$+ \alpha_2^2(b_{22} - \alpha_2)z^{-3} - \alpha_2^3(b_{22} - \alpha_2)z^{-4} + \alpha_2^4(b_{22} - \alpha_2)z^{-5} - \alpha_2^5(b_{22} - \alpha_2)z^{-6} +$$

$$+ \alpha_2^6(b_{22} - \alpha_2)z^{-7} - \alpha_2^7(b_{22} - \alpha_2)z^{-8} + \alpha_2^8(b_{22} - \alpha_2)z^{-9} - \alpha_2^9 b_{22} z^{-10},$$

де z^{-1} — оператор зворотного зсуву на один період квантування T_0 .

При підстановці значень коефіцієнтів з урахуванням рівностей (9.4.2), (9.4.3) запишемо модель (9.4.14) в різнатемпovій формі за двома каналами:

$$\begin{aligned}
y_1 \left[\left[\frac{k}{5} \right] h_1 \right] = & 0,9048 \cdot y_1 \left[\left(\left[\frac{k}{5} \right] - 1 \right) h_1 \right] + \xi_1(kT_0) + \\
& + 0,4802 \xi_1[(k-1)T_0] + 0,47069 \xi_1[(k-2)T_0] + \\
& + 0,46137 \xi_1[(k-3)T_0] + 0,45223 \xi_1[(k-4)T_0] - \\
& - 0,46156 \xi_1[(k-5)T_0] + \xi_2(kT_0) + 1,5802 \xi_2[(k-1)T_0] + \\
& + 1,5489 \xi_2[(k-2)T_0] + 1,51824 \xi_2[(k-3)T_0] + \\
& + 1,48818 \xi_2[(k-4)T_0] + 0,55387 \xi_2[(k-5)T_0]. \tag{9.4.15}
\end{aligned}$$

$$\begin{aligned}
y_2 \left[\left[\frac{k}{10} \right] h_2 \right] = & 0,90438 \cdot y_2 \left[\left(\left[\frac{k}{10} \right] - 1 \right) h_2 \right] + \xi_1(kT_0) + \\
& + 0,29 \xi_1[(k-1)T_0] + 0,2871 \xi_1[(k-2)T_0] + 0,2842 \xi_1[(k-3)T_0] + \\
& + 0,28138 \xi_1[(k-4)T_0] + 0,27857 \xi_1[(k-5)T_0] + 0,2558 \xi_1[(k-6)T_0] + \\
& + 0,273 \xi_1[(k-7)T_0] + 0,2703 \xi_1[(k-8)T_0] + 0,26759 \xi_1[(k-9)T_0] - \\
& 0,2649 \xi_1[(k-10)T_0] + \xi_2(kT_0) + 0,54 \xi_2[(k-1)T_0] + 0,5346 \xi_2[(k-2)T_0] + \\
& + 0,52925 \xi_2[(k-3)T_0] + 0,52396 \xi_2[(k-4)T_0] + \\
& + 0,51872 \xi_2[(k-5)T_0] - 0,51353 \xi_2[(k-6)T_0] + \\
& + 0,50839 \xi_2[(k-7)T_0] + 0,5033 \xi_2[(k-8)T_0] + \\
& + 0,49828 \xi_2[(k-9)T_0] - 0,41107 \xi_2[(k-10)T_0]. \tag{9.4.16}
\end{aligned}$$

Складемо діофантове рівняння у матричній формі:

$$\begin{aligned}
\begin{bmatrix} C_{11}(z^{-1}) & C_{12}(z^{-1}) \\ C_{21}(z^{-1}) & C_{22}(z^{-1}) \end{bmatrix} = & \begin{bmatrix} A_{11}(z^{-1}) & 0 \\ 0 & A_{22}(z^{-1}) \end{bmatrix} \times \\
\times & \begin{bmatrix} L_{11}(z^{-1}) & L_{12}(z^{-1}) \\ L_{21}(z^{-1}) & L_{22}(z^{-1}) \end{bmatrix} + \begin{bmatrix} z^{-5} & 0 \\ 0 & z^{-10} \end{bmatrix} \begin{bmatrix} F_{11}(z^{-1}) & F_{12}(z^{-1}) \\ F_{21}(z^{-1}) & F_{22}(z^{-1}) \end{bmatrix},
\end{aligned}$$

з якого при $L_{11}(z^{-1}) = C_{11}(z^{-1})$, $L_{12}(z^{-1}) = C_{12}(z^{-1})$, $L_{21}(z^{-1}) = C_{21}(z^{-1})$, $L_{22}(z^{-1}) = C_{22}(z^{-1})$ визначаємо поліноми:

$$F_{11}(z^{-1}) = -\alpha_1^5 - \alpha_1^5(b_{11} - \alpha_1)z^{-1} + \alpha_1^6(b_{11} - \alpha_1)z^{-2} -$$

$$-\alpha_1^7(b_{11} - \alpha_1)z^{-3} + \alpha_1^8(b_{11} - \alpha_1)z^{-4} - \alpha_1^9 b_{11} z^{-5};$$

$$F_{12}(z^{-1}) = -\alpha_1^5 - \alpha_1^5(b_{12} - \alpha_1)z^{-1} + \alpha_1^6(b_{12} - \alpha_1)z^{-2} -$$

$$\begin{aligned}
& -\alpha_1^7(b_{12}-\alpha_1)z^{-3} + \alpha_1^8(b_{12}-\alpha_1)z^{-4} - \alpha_1^9b_{12}z^{-5}; \\
F_{21}(z^{-1}) &= \alpha_2^{10} + \alpha_2^{10}(b_{21}-\alpha_2)z^{-1} - \alpha_2^{11}(b_{21}-\alpha_2)z^{-2} + \\
& + \alpha_2^{12}(b_{21}-\alpha_2)z^{-3} - \alpha_2^{13}(b_{21}-\alpha_2)z^{-4} + \alpha_2^{14}(b_{21}-\alpha_2)z^{-5} - \alpha_2^{15}(b_{21}-\alpha_2)z^{-6} + \\
& + \alpha_2^{16}(b_{21}-\alpha_2)z^{-7} - \alpha_2^{17}(b_{21}-\alpha_2)z^{-8} + \alpha_2^{18}(b_{21}-\alpha_2)z^{-9} - \alpha_2^{19}b_{21}z^{-10}; \\
F_{22}(z^{-1}) &= \alpha_2^{10} + \alpha_2^{10}(b_{22}-\alpha_2)z^{-1} - \alpha_2^{11}(b_{22}-\alpha_2)z^{-2} + \\
& \alpha_2^{12}(b_{22}-\alpha_2)z^{-3} - \alpha_2^{13}(b_{22}-\alpha_2)z^{-4} + \alpha_2^{14}(b_{22}-\alpha_2)z^{-5} - \alpha_2^{15}(b_{22}-\alpha_2)z^{-6} + \\
& + \alpha_2^{16}(b_{22}-\alpha_2)z^{-7} - \alpha_2^{17}(b_{22}-\alpha_2)z^{-8} + \alpha_2^{18}(b_{22}-\alpha_2)z^{-9} - \alpha_2^{19}b_{22}z^{-10}.
\end{aligned}$$

Тоді на основі (9.4.11) можна виконати прогнозування вихідних координат при різнометровій дискретизації координат:

$$\begin{aligned}
y_1^* \left[\left(\left[\frac{k}{5} \right] + 1 \right) h_1 \left| \left[\frac{k}{5} \right] h_1 \right. \right] &= 0,9048 \cdot y_1 \left[\left[\frac{k}{5} \right] h_1 \right] + 0,9048 \xi_1(kT_0) + \\
& + 0,43488 \xi_1[(k-1)T_0] + 0,42587 \xi_1[(k-2)T_0] + \\
& + 0,41744 \xi_1[(k-3)T_0] + 0,40917 \xi_1[(k-4)T_0] - \\
& - 0,41762 \xi_1[(k-5)T_0] + 0,9048 \xi_2(kT_0) + 1,42976 \xi_2[(k-1)T_0] + \\
& + 1,40145 \xi_2[(k-2)T_0] + 1,3737 \xi_2[(k-3)T_0] + \\
& + 1,3465 \xi_2[(k-4)T_0] + 0,50114 \xi_2[(k-5)T_0]. \tag{9.4.17}
\end{aligned}$$

$$\begin{aligned}
y_2^* \left[\left(\left[\frac{k}{10} \right] + 1 \right) h_2 \left| \left[\frac{k}{10} \right] h_2 \right. \right] &= 0,90438 y_2 \left[\left[\frac{k}{10} \right] h_2 \right] + \\
& + 0,90438 \xi_1(kT_0) + 0,26227 \xi_1[(k-1)T_0] + 0,25964 \cdot \\
& \xi_1[(k-2)T_0] + 0,25705 \xi_1[(k-3)T_0] + 0,25448 \xi_1[(k-4)T_0] + \\
& + 0,25193 \xi_1[(k-5)T_0] + \\
& 0,24941 \xi_1[(k-6)T_0] + 0,24692 \xi_1[(k-7)T_0] + \\
& + 0,24445 \xi_1[(k-8)T_0] + 0,242 \xi_1[(k-9)T_0] - \\
& - 0,57831 \xi_1[(k-10)T_0] + 0,90438 \xi_2(kT_0) + \\
& + 0,48836 \xi_2[(k-1)T_0] + 0,48348 \xi_2[(k-2)T_0] + \\
& + 0,47864 \xi_2[(k-3)T_0] + 0,47386 \xi_2[(k-4)T_0] + \\
& + 0,46912 \xi_2[(k-5)T_0] + 0,46443 \xi_2[(k-6)T_0] +
\end{aligned}$$

$$\begin{aligned}
 & + 0,46443\xi_2[(k-7)T_0] + 0,45978\xi_2[(k-8)T_0] + \\
 & + 0,45063\xi_2[(k-9)T_0] - 0,37177\xi_2[(k-10)T_0]. \quad (9.4.18)
 \end{aligned}$$

Результати цифрового моделювання різnotемпової моделі (9.4.15), (9.4.16) та прогнозованих значень вихідних координат (9.4.17), (9.4.18) при додатковій подачі на входи одиничних ступінчастих збурень приведені на рис. 9.2.

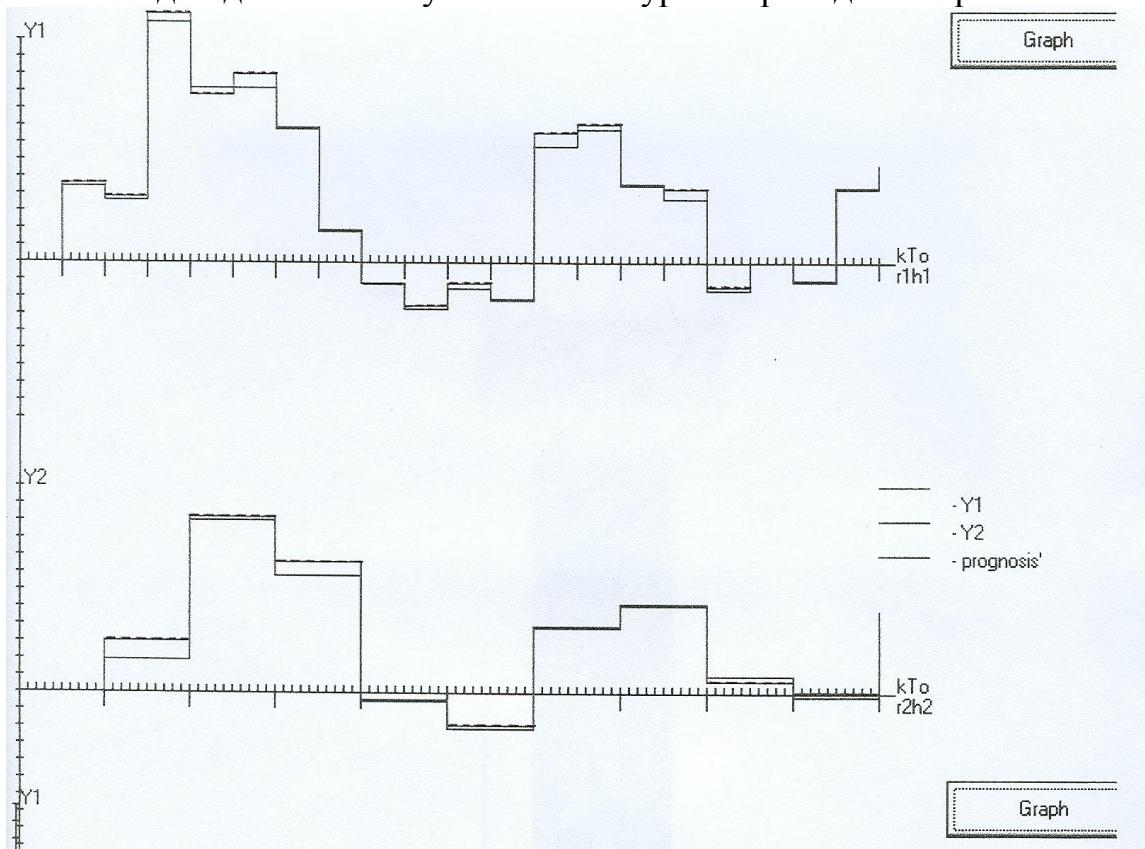


Рис. 9.2. Графіки моделювання вихідних змінних ($\cdots y_1$, $-- y_2$) та їхніх прогнозованих значень (—) різnotемпової двовимірної системи

9.5 Прогнозування нестационарних процесів на основі математичних моделей з різнометровою дискретизацією

Динаміка нестационарного процесу представлена математичною моделлю APICC з різнометровою дискретизацією

$$A(z_1^{-1})\nabla y(rh) = C(z^{-1})\xi(kT_0) + a_0, \quad (9.5.1)$$

де перша різниця $\nabla y(rh) = y(rh) - y[(r-1)h]$. Поліном $A(z_1^{-1})$ має вигляд $A(z_1^{-1}) = 1 + a_1 z_1^{-1}$,

$$C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_{n_c} z^{-n_c}, \quad (9.5.2)$$

а співвідношення періодів квантування в (9.5.1) представлено виразом (9.2.2), при якому оператори зворотного зсуву в поліномах $A(z_1^{-1})$, $C(z^{-1})$ пов'язані співвідношенням (9.2.7). В результаті поліном $A(z_1^{-1})$ можна записати через оператор z , який відповідає базовому періоду квантування T_0 на основі виразу (9.2.8). На основі рівностей (9.2.2), (9.2.5), (9.2.7), (9.2.8) запишемо рівняння APICC (9.5.1) у такій формі:

$$(1 - z^{-m})A(z^{-1})y(kT_0) = C(z^{-1})\xi(kT_0) + a_0.$$

Представимо цей вираз відносно координати $y(kT_0)$:

$$y(rh) = \frac{C(z^{-1})}{(1 - z^{-m})A(z^{-1})}\xi(kT_0) + \frac{a_0}{(1 - z^{-m})A(z^{-1})}.$$

Виконаємо зсув за часом вперед на один період квантування h при умові, що $a_0 = const$:

$$\begin{aligned} y\left[\left(\left[\frac{k}{m}\right] + 1\right)h\right] &= z^m \frac{C(z^{-1})}{(1 - z^{-m})A(z^{-1})} + \frac{a_0}{(1 - z^{-m})A(z^{-1})} = \\ &= y[(k+m)T_0]. \end{aligned} \quad (9.5.3)$$

Складемо діофантове рівняння

$$C(z^{-1}) = (1 - z^{-m})A(z^{-1})L(z^{-1}) + z^{-m}F(z^{-1}), \quad (9.5.4)$$

на основі якого вираз (9.5.3) можна записати таким чином:

$$\begin{aligned} y[(k+m)T_0] &= y\left[\left(\left[\frac{k}{m}\right] + 1\right)h\right] = \\ &= z^m L(z^{-1})\xi(kT_0) + \frac{F(z^{-1})}{(1 - z^{-m})A(z^{-1})} + \end{aligned}$$

$$+ \frac{a_0}{(1 - z^{-m})A(z^{-1})}. \quad (9.5.5)$$

Перша складова в правій частині (9.5.5) представляє собою помилку прогнозування вихідного сигналу y на m періодів квантування T_0 або на один період квантування h , тобто:

$$e[(k+m)T_0] = e\left[\left(\left[\frac{k}{m}\right] + 1\right)h\right] = z^m L(z^{-1})\xi(kT_0).$$

Тоді прогнозоване значення вихідного сигналу буде:

$$\begin{aligned} y^*[(k+m)T_0 | kT_0] &= y^*\left[\left(\left[\frac{k}{m}\right] + 1\right)h \left|\left[\frac{k}{m}\right]h\right.\right] = \\ &= \frac{F(z^{-1})}{(1 - z^{-m})A(z^{-1})} + \frac{a_0}{(1 - z^{-m})A(z^{-1})}. \end{aligned} \quad (9.5.6)$$

При цьому поліном $F(z^{-1})$ однозначно визначається з діофанового рівняння (9.5.4) при $L(z^{-1}) = C(z^{-1})$ на основі $(1 - a_1 + a_1 z^{-m})C(z^{-1})$.

9.6 Прогнозування динаміки процесів у просторах змінних для моделей з різnotемповою дискретизацією

Розглянемо класичну модель авторегресії і ковзного середнього з додатковим вхідним сигналом управління (ARMAX) при однотемповій дискретизації і наявності запізнення по управлінню:

$$A(z^{-1})y(kT_0) = z^{-d}B(z^{-1})u(kT_0) + C(z^{-1})\xi(kT_0) + \delta, \quad (9.6.1)$$

де $y(kT_0)$, $u(kT_0)$, $\xi(kT_0)$ — відповідно вихідна координата, керуюче діяння і збурення в моменти часу kT_0 , $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$ — поліноми відносно оператора зворотного зсуву z^{-1} на один період квантування T_0 . Сигнал $\xi(kT_0)$ представляє збурення у вигляді послідовності дискретного білого шуму з нульовим середнім, а δ — величина зміщення. Запізнення $q = \left[\frac{\tau}{T_0}\right]$, де r — час запізнення, а $\left[\frac{\tau}{T_0}\right]$ — ціла частина частки від ділення τ на період T_0 .

При використанні моделей ARMAX для опису динаміки процесів у дискретному часі основною проблемою є зміна зміщення, середнього значення вихідної координати. Зміна зміщення виникає при зміні загального режиму

процесу, а також статичного коефіцієнта передачі $\frac{y}{u}$ внаслідок нелінійності реального процесу, яка не врахована в лінеаризованій моделі (9.6.1). При цьому зміщення δ буде невідомою величиною, яка змінюється за часом.

Другою проблемою при описі динаміки реальних процесів за допомогою моделі (9.6.1) є наявність невимірюваних збурень з ненульовим середнім на проміжках часу, порівнюваних з інерційністю процесів.

Зміна за часом зміщення δ і наявність невимірюваного збурення $\xi(kT_0)$ з ненульовим середнім приводять до значного зміщення при прогнозуванні вихідної координати $y(kT_0)$ відомими методами [2, 25, 48].

У роботі [41] для однотемпової моделі ARMAX (9.6.1) запропонований k -інкрементний прогнозатор, який забезпечує компенсацію зміни за часом зміщення δ в моделі (9.6.1) при невимірюваних збуреннях $\xi(kT_0)$.

Мета даного підрозділу полягає в розробці методики прогнозування стаціонарних часових рядів при дискретизації вихідних координат і керуючих діянь із збільшеними періодами квантування h , а вхідних збурень — з малим базовим періодом квантування T_0 в умовах наявності невимірюваних збурень з ненульовим середнім та зміни за часом зміщення δ , а також адаптивної настройки функції прогнозування.

9.6.1 Розробка функції прогнозування вихідної координати на d періодів квантування h

Вихідна динамічна модель (9.6.1) для другого порядку поліномів $A(z^{-1})$, $B(z^{-1})$ при різнометровій дискретизації на основі (9.2.12) може бути представлена в різницевій формі:

$$\begin{aligned} y\left[\left[\frac{k}{m}\right]h\right] = & -a_1 y\left[\left(\left[\frac{k}{m}\right]-1\right)h\right] - a_2 y\left[\left(\left[\frac{k}{m}\right]-2\right)h\right] + \\ & + b_1 u\left[\left(\left[\frac{k}{m}\right]-d\right)h\right] + b_2 u\left[\left(\left[\frac{k}{m}\right]-d-1\right)h\right] + \\ & + v\left[\left[\frac{k}{m}\right]h\right] + c_1 v\left[\left[\frac{k}{m}\right]h-T_0\right] + \dots + \\ & + c_{2m} v\left[\left[\frac{k}{m}\right]h-2mT_0\right] + a_0, \end{aligned} \quad (9.6.2)$$

Її можна записати в класичній формі:

$$A(z_1^{-1})y(rh) = z_1^{-d}B(z_1^{-1})u(rh) + C(z^{-1})v(kT_0) + a_0, \quad (9.6.3)$$

де збільшений період квантування при дискретизації вихідної координати визначається співвідношенням

$$h = mT_0. \quad (9.6.4)$$

При цьому m — ціле число, які більше одиниці, а T_0 — бізний період квантування. В моделі (9.6.2) $\left[\frac{k}{m}\right]$ — ціла частина частки від ділення номера дискретного відліку k на коефіцієнт m . Поліноми різnotемпової моделі (9.6.3) мають вигляд

$$A(z_1^{-1}) = 1 + a_1z_1^{-1} + a_2z_1^{-2}, \quad (9.6.5)$$

$$B(z_1^{-1}) = b_0 + b_1z_1^{-1}, \quad (9.6.6)$$

$$C(z^{-1}) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{2m}z^{-2m}. \quad (9.6.7)$$

На основі (9.6.4) співвідношення операторів зворотного зсуву в поліномах (9.6.5)–(9.6.7) буде

$$z_1^{-1} = z^{-m}. \quad (9.6.8)$$

Таким чином оператор z^{-m} визначає зворотний зсув на період h . Тоді рівності (9.6.5), (9.6.6) можна записати:

$$A(z^{-1}) = 1 + a_1z^{-m} + a_2z^{-2m}; \quad (9.6.9)$$

$$B(z^{-1}) = b_0 + b_1z^{-m}, \quad (9.6.10)$$

а вихідну модель (9.6.3) можна представити таким чином:

$$A(z^{-1})y(rh) = z^{-dm}B(z^{-1})u(rh) + C(z^{-1})v(kT_0) + a_0. \quad (9.6.11)$$

Випадкові збурення $v(kT_0)$ в моделі (9.6.11) мають постійний дрейф низької частоти і на проміжках часу, які можна порівняти з постійною часу процесу, мають ненульове середнє. Модель динаміки такого збурення можна представити у вигляді марківського процесу:

$$v[(k+1)T_0] = g v(kT_0) + \xi(kT_0), \quad (9.6.12)$$

де $\{\xi(kT_0)\}$ — послідовність збурень у вигляді дискретного білого шуму з нульовим середнім. Відповідно до [7] у випадку $g = [0,8\dots0,95]$ у спектральному складі вихідного сигналу на виході фільтра (9.6.12) вирішальне значення будуть мати низькочастотні складові. Параметр a_0 в моделі (9.6.11) визначає рівень процесу і може змінюватися за часом внаслідок зміни середнього значення координати у або наявності нелінійностей у процесі, які не враховані у лінійній моделі (9.6.11).

При врахуванні рівностей (9.6.4), (9.6.8), (9.6.9), (9.6.10) розв'яжемо рівняння (9.6.11) відносно $y(rh)$ і виконаємо зсув за часом вперед на dh :

$$\begin{aligned} y[(r+d)h] &= y[(k+dm)T_0] = \frac{B(z^{-1})}{A(z^{-1})} u(rh) + \\ &+ z^{dm} \frac{C(z^{-1})}{A(z^{-1})} v(kT_0) + \frac{a_0}{A(z^{-1})}. \end{aligned} \quad (9.6.13)$$

Складемо таке діофантове рівняння:

$$C(z^{-1}) = A(z^{-1})L(z^{-1}) + z^{-dm}F(z^{-1}), \quad (9.6.14)$$

на основі якого вираз (9.6.13) можна перетворити до такого вигляду:

$$\begin{aligned} y[(r+d)h] &= y[(k+dm)T_0] = \frac{B(z^{-1})}{A(z^{-1})} u(rh) + \\ &+ z^{dm}L(z^{-1})v(kT_0) + \frac{F(z^{-1})}{A(z^{-1})} v(kT_0) + \frac{a_0}{A(z^{-1})}. \end{aligned}$$

Тоді функція прогнозування на d періодів квантування h буде визначатися таким чином:

$$\begin{aligned} y^*[r+d]h | rh] &= y^*[[(k+dm)T_0 | kT_0] = \\ &= \frac{B(z^{-1})}{A(z^{-1})} u(rh) + \frac{F(z^{-1})}{A(z^{-1})} v(kT_0) + \frac{a_0}{A(z^{-1})}, \end{aligned} \quad (9.6.15)$$

а помилка прогнозування $e[(r+d)] = e[(k+dm)T_0] = z^{dm}L(z^{-1})v(kT_0)$.

Розглянемо методику визначення полінома $F(z^{-1})$ з діофантового рівняння (9.6.14) з врахуванням (9.6.7), (9.6.9), (9.6.11).

1. При прогнозуванні на $1h = mT_0$ діофантове рівняння (9.6.14) запишеться у вигляді

$$C(z^{-1}) = (1 + a_1z^{-m} + a_2z^{-2m})L(z^{-1}) + z^{-m}F(z^{-1}).$$

Вибираємо поліном $L(z^{-1}) = C(z^{-1})$. Тоді з діофантового рівняння визначаємо

$$F(z^{-1}) = -(a_1 + a_2z^{-m})C(z^{-1}).$$

2. При прогнозуванні на $2h = 2mT_0$ діофантове рівняння запишеться так:

$$C(z^{-1}) = (1 + a_1z^{-m} + a_2z^{-2m})L(z^{-1}) + z^{-4m}F(z^{-1}).$$

При виборі полінома $L(z^{-1}) = (1 - a_1z^{-m})C(z^{-1})$ поліном $F(z^{-1})$ однозначно визначається:

$$F(z^{-1}) = [(a_1^2 - a_2) + a_1a_2z^{-m}]C(z^{-1}).$$

3. При прогнозуванні на $3h = 3mT_0$ запишемо діофантове рівняння

$$C(z^{-1}) = (1 + a_1 z^{-m} + a_2 z^{-2m})L(z^{-1}) + z^{-3m}F(z^{-1}).$$

Виберемо поліном $L(z^{-1}) = (1 - a_1 z^{-m} + (a_1^2 - a_2)z^{-2m})C(z^{-1})$. Тоді однозначно визначається:

$$F(z^{-1}) = -[(a_1^3 - 2a_1 a_2) + a_2(a_1^2 - a_2)z^{-m}]C(z^{-1}).$$

4. При прогнозуванні на $4h = 4mT_0$ діофантове рівняння буде

$$C(z^{-1}) = (1 + a_1 z^{-m} + a_2 z^{-2m})L(z^{-1}) + z^{-4m}F(z^{-1}).$$

Шляхом вибору поліному $L(z^{-1}) = [1 - a_1 z^{-m} + (a_1^2 - a_2)z^{-2m} - (a_1^3 - 2a_1 a_2)z^{-3m}] \times C(z^{-1})$ з діофантового рівняння визначаємо:

$$F(z^{-1}) = [(a_1^4 - 3a_1^2 a_2 + a_2^2) + a_2(a_1^3 - 2a_1 a_2)z^{-m}]C(z^{-1}).$$

Таким чином, при прогнозуванні вихідної координати процесу (9.6.11) на d періодів h при $L(z^{-1}) = (1 + l_1 z^{-m} + l_2 z^{-2m} + \dots + l_{(d-1)} z^{-(d-1)m})C(z^{-1})$ структура полінома $F(z^{-1})$ в загальному вигляді запишеться

$$F(z^{-1}) = (f_0 + f_1 z^{-m})C(z^{-1}), \quad (9.6.16)$$

а функція прогнозування (9.6.15)

$$\begin{aligned} y^*[(r+d)h | rh] &= y^*[(k+dm)T_0 | kT_0] = \frac{B(z^{-1})}{A(z^{-1})}u(rh) + \\ &+ \frac{(f_0 + f_1 z^{-m})}{A(z^{-1})}C(z^{-1})v(kT_0) + \frac{a_0}{A(z^{-1})}. \end{aligned} \quad (9.6.17)$$

У зв'язку з тим, що у вираз для $F(z^{-1})$ відповідно до (9.6.16) входить поліном $C(z^{-1})$, відома методика прогнозування [7] не підходить при різнометровій дискретизації, тому що при цьому поліном $C(z^{-1})$ скорочується. Внаслідок цього витрачається інформація про передісторію дії збурень. Тому у випадку, коли збурення $v(kT_0)$ не вимірюються, в цьому розділі розробка функції прогнозування буде відрізнятися від методики [23, 48].

Визначимо з моделі (9.6.11):

$$C(z^{-1})v(kT_0) = A(z^{-1})y(rh) - z^{-dm}B(z^{-1})u(rh) - a_0.$$

Піставимо цей вираз у рівняння (9.6.15). В результаті будемо мати:

$$\begin{aligned} y^*[(r+d)h | rh] &= y^*[(k+dm)T_0 | kT_0] = \\ &= \frac{[1 - z^{-dm}(f_0 + f_1 z^{-m})]B(z^{-1})}{A(z^{-1})}u(rh) + \end{aligned}$$

$$+(f_0 + f_1 z^{-m})y(rh) + \frac{a_0(1 - f_0 - f_1 z^{-m})}{A(z^{-1})}.$$

Враховуючи (9.6.9) прогнозоване значення координати y можна представити у вигляді

$$\begin{aligned} y^*[(r+d)h | rh] = & -a_1 y^* \{ [r + (d-1)]h | (r-1)h \} - \\ & -a_2 y^* \{ [r + (d-2)]h | (r-2) \} + (1 + a_1 z^{-m} + \\ & + a_2 z^{-2m})(f_0 + f_1 z^{-m})y(rh) + [1 - z^{-dm}(f_0 + \\ & + f_1 z^{-m})]B(z^{-1})u(rh) + a_0(1 - f_0 - f_1 z^{-m}). \end{aligned} \quad (9.6.18)$$

Запишемо це рівняння із зсувом за часом назад на dh :

$$\begin{aligned} y^*[rh | (r-d)h] = & -a_1 y^*[(r-1)h | (r-d-1)h] - \\ & -a_2 y^*[(r-2)h | (r-d-2) + (1 + a_1 z^{-m} + a_2 z^{-2m}) \times \\ & \times (f_0 + f_1 z^{-m})y[(r-d)h] + [1 - z^{-dm}(f_0 + f_1 z^{-m})] \times \\ & \times B(z^{-1})u[(r-d)h] + a_0(1 - f_0 - f_1 z^{-m})z^{-dm}. \end{aligned}$$

Введемо такі різницеві оператори:

$$\begin{aligned} \Delta_{dh} y^* \{ [r + (d-1)]h | (r-1)h \} = & y^* \{ [r + (d-1)]h | (r-1)h \} - \\ & - y^*[(r-1)h | (r-d-1)h]; \\ \Delta_{dh} y^* \{ [r + (d-2)]h | (r-2)h \} = & y^* \{ [r + (d-2)]h | (r-2)h \} - \\ & - y^*[(r-2)h | (r-d-2)h]; \\ \Delta_{dh} y(rh) = & y(rh) - y[r-d]h; \\ \Delta_{dh} u(rh) = & u(rh) - u[(r-d)h]; \\ \Delta_{dh} a_0(1 - f_0 - f_1 z^{-m}) = & a_0(1 - f_0 - f_1 z^{-m}) - \\ & - a_0(1 - f_0 - f_1 z^{-m})z^{-dm}. \end{aligned} \quad (9.6.19)$$

Вирахуємо почленно вираз (9.6.18) із (9.6.13). Тоді прогнозоване значення y буде представлене у приростах змінних:

$$\begin{aligned} y^*[(r+d)h | rh] = & y(rh) - a_1 \Delta_{dh} y^* \{ [r + (d-1)]h | \\ & | (r-1)h \} - a_2 \Delta_{dh} y^* \{ [r + (d-2)]h | (r-2)h \} + \\ & + (1 + a_1 z^{-m} + a_2 z^{-2m})(f_0 + f_1 z^{-m})\Delta_{dh} y(rh) + \end{aligned}$$

$$\begin{aligned}
& + [1 - z^{-dm} (f_0 + f_1 z^{-m})] B(z^{-1}) \Delta_{dh} u(rh) + \\
& + \Delta_{dh} a_0 (1 - f_0 - f_1 z^{-m}). \tag{9.6.20}
\end{aligned}$$

Прогнозовані координати $y^*[r + (d-1)h | (r-1)h]$, $y^*[r + (d-2)h | (r-2)h]$, які входять в різницеві оператори (9.6.19), визначаються на основі процедури (9.6.18) відповідно при прогнозуванні на $(d-1)h$ і $(d-2)h$. При цьому необхідно окремо обчислити коефіцієнти f_0 , f_1 , відповідно визначенням $F(z^{-1})$ з діофантового рівняння (9.6.14) для прогнозування на $(d-1)h$ і $(d-2)h$.

9.6.2 Адаптивна настройка параметрів функції прогнозування

Динамічні параметри прогнозованих фінансово-економічних і соціальних процесів змінюються за часом внаслідок зміни зовнішніх умов, законодавчих актів, конфліктних та кризових ситуацій і т.п. Тому коефіцієнти математичної моделі (9.6.11) і розробленої функції прогнозування (9.6.20) також будуть змінюватися за часом. Для адаптивної настройки параметрів функції (9.6.20) запишемо її із зсувом назад за часом на dh

$$\begin{aligned}
y(rh) = & y[(r-d)h] - a_1 \Delta_{dh} y[(r-1)h] - a_2 \Delta_{dh} y[(r-2)h] + \\
& + (1 + a_1 z^{-m} + a_2 z^{-2m}) (f_0 + f_1 z^{-m}) \Delta_{dh} y[(r-d)h] + \\
& + [1 + z^{-dm} (f_0 + f_1 z^{-m})] B(z^{-1}) \Delta_{dh} u[(r-d)h] + \\
& + \Delta_{dh} a_0 (1 - f_0 - f_1 z^{-m}).
\end{aligned}$$

Після перемноження і врахування (9.6.6) і (9.6.8) попередній вираз можна записати:

$$\begin{aligned}
y(rh) = & y[(r-d)h] - a_1 \Delta_{dh} y[(r-1)h] - a_2 \Delta_{dh} y[(r-2)h] + \\
& + f_0 \Delta_{dh} y[(r-d)h] + (a_1 f_0 + f_1) \Delta_{dh} y[(r-d-1)h] + \\
& + (a_2 f_0 + a_1 f_1) \Delta_{dh} y[(r-d-2)h] + a_2 f_1 \Delta_{dh} y[(r-d-3)h] + \\
& + b_0 \Delta_{dh} u(rh) + b_1 \Delta_{dh} u[(r-1)h] + b_0 f_0 \Delta_{dh} u[(r-d)h] + \\
& + (b_0 f_1 + b_1 f_0) \Delta_{dh} u[(r-d-1)h] + \\
& + b_1 f_1 \Delta_{dh} u[(r-d-2)h] + \Delta_{dh} a_0 (1 - f_0 - f_1 z^{-m}). \tag{9.6.21}
\end{aligned}$$

При уведенні коефіцієнтів $f'_1 = a_1 f_0 + f_1$; $f_2 = a_2 f_0 + a_1 f_1$; $f_3 = a_2 f_1$; $b_2 = b_0 f_0$; $b_3 = b_0 f_1 + b_1 f_0$; $b_4 = b_1 f_1$; $a'_0 = \Delta_{dh} a_0 (1 - f_0 - f_1 z^{-m})$ вираз (9.6.21) можна представити у формі перемноження векторів

$$y(rh) = \bar{X}^T(rh) \bar{\theta} + \xi(rh), \tag{9.6.22}$$

де $\bar{X}(rh)$ — вектор вимірюваних дискретних відліків таких координат процесу:

$$\begin{aligned}\bar{X}^T(rh) = & \{y[(r-d)h], \Delta_{dh}y[(r-1)h], \Delta_{dh}y[(r-2)h] \\ & \Delta_{dh}y[(r-d)h], \Delta_{dh}y[(r-d-1)h], \Delta_{dh}y[(r-d-2)h] \\ & \Delta_{dh}y[(r-d-3)h], \Delta_{dh}u(rh), \Delta_{dh}u(r-1)h] \\ & \Delta_{dh}u[(r-d)h], \Delta_{dh}u[(r-d-1)h], \Delta_{dh}u[(r-d-2)h], 1\}\end{aligned}$$

$\bar{\theta}^T$ — вектор невідомих параметрів процесу, який необхідно періодично оцінювати мірою надходження дискретних вимірювань вхідних і вихідних координат процесу

$$\bar{\theta}^T = \{1, -a_1, -a_2, f_0, f'_1, f_2, f_3, b_0, b_1, b_2, b_3, b_4, a'_0\};$$

$\xi(rh)$ — дискретний білий шум.

Для оцінки вектора параметрів в реальному часі застосуємо рекурентний метод найменших квадратів:

$$\begin{aligned}\hat{\theta}(rh) = & \hat{\theta}[(r-1)h] + K(rh)\{y(rh) - \bar{X}^T(rh)\hat{\theta}[(r-1)h]\}; \\ K(rh) = & P[(r-1)h] \cdot \bar{X}(rh)\{\beta + \bar{X}^T(rh)\bar{P}[(r-1)h]\bar{X}(rh)\}^{-1},\end{aligned}$$

де фактор експоненціального забування доцільно вибирати в межах $0,9 < \beta < 1$.

Матриця коваріації помилки оцінки визначається таким чином:

$$\begin{aligned}P(rh) = & \frac{1}{\beta}\{P[(r-1)h] - P[(r-1)h]\bar{X}(rh) \cdot \\ & \cdot \{\beta + \bar{X}^T(rh)P[(r-1)h]\bar{X}(rh)\}^{-1}\bar{X}^T(rh) \cdot P[(r-1)h]\}.\end{aligned}$$

Початковими значеннями матриці коваріації можна прийняти $P(0) = \alpha I$, α — велике додатне число, а I — одинична матриця.

Приклад 9.3.. Динаміка процесу зображена різнометровою моделлю

$$\begin{aligned}(1+a_1z^{-2}+a_2z^{-4})y(rh) = & (1+c_1z^{-1}+ \\ & +c_2z^{-2}+c_3z^{-3}+c_4z^{-4})v(kT_0)+a_0,\end{aligned}$$

де $a_1 = -0,9539$; $a_2 = 0,1107$; $c_1 = 0,3726$; $c_2 = -0,6125$; $c_3 = -0,045$; $c_4 = 0,0665$; $a_0 = 0,1568$.

Періоди квантування зв'язані співвідношенням $h = 2T_0$.

Необхідно розробити процедуру прогнозування координат у на час $2h$, якщо на вхід поступають збурення $v(kT_0)$ з низькочастотним дрейфом відповідно до (9.6.12).

Запишемо вищенаведену модель у різницевій формі при умові $rh = \left[\frac{k}{2}\right]h$:

$$\begin{aligned}
y \left[\left[\frac{k}{2} \right] h \right] = & -a_1 y \left[\left(\left[\frac{k}{2} \right] - 1 \right) h \right] - a_2 y \left[\left(\left[\frac{k}{2} \right] - 2 \right) h \right] + \\
& + v(kT_0) + c_1 v[(k-1)T_0] + c_2 v[(k-2)T_0] + \\
& + c_3 v[(k-3)T_0] + c_4 v[(k-4)T_0] + a_0,
\end{aligned}$$

де збурення $v[(k-i)T_0]$ ($i=0,1,\dots,4$) мають низькочастотний дрейф і генеруються відповідно до (9.6.12).

Визначимо коефіцієнти f_0 , f_1 функції прогнозування (9.6.20). Для діофантового рівняння (9.6.14) при $d=2$ виберемо поліном

$$L(z^{-1}) = (1 - a_1 z^{-2})(1 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3} + c_4 z^{-4}).$$

Тоді дістанемо:

$$F(z^{-1}) = [(a_1^2 - a_2) + a_1 a_2 z^{-2}](1 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3} + c_4 z^{-4}),$$

з якого визначимо $f_0 = (a_1^2 - a_2) = 0,7992$; $f_1 = a_1 a_2 = -0,10559$. У результаті функція прогнозування (9.6.20) запишеться таким чином:

$$\begin{aligned}
y^* \left[\left(\left[\frac{k}{2} \right] + 2 \right) h \left| \left[\frac{k}{2} \right] h \right. \right] = & y \left[\left[\frac{k}{2} \right] h \right] - a_1 \cdot \Delta_{2h} y^* \left[\left(\left[\frac{k}{2} \right] + 1 \right) h \right| \\
& \left. \left| \left(\left[\frac{k}{2} \right] - 1 \right) h \right. \right] - a_2 \cdot \Delta_{2h} y^* \left[\left[\frac{k}{2} \right] h \right| \left. \left| \left(\left[\frac{k}{2} \right] - 2 \right) h \right. \right] + \\
& + (1 + a_1 z^{-2} + a_2 z^{-4})(f_0 + f_1 z^{-2}) \cdot \Delta_{2h} y \left[\left[\frac{k}{2} \right] h \right| + \\
& + \Delta_{2h} a_0 (1 - f_0 - f_1 z^{-2}) = \\
y \left[\left[\frac{k}{2} \right] h \right] + & 0,9539 \cdot \Delta_{2h} y^* \left[\left(\left[\frac{k}{2} \right] + 1 \right) h \right| \left. \left| \left(\left[\frac{k}{2} \right] - 1 \right) h \right. \right] - \\
- 0,1107 \cdot \Delta_{2h} y \left[\left[\frac{k}{2} \right] h \right| \left. \left| \left(\left[\frac{k}{2} \right] - 2 \right) h \right. \right] + & 0,7922 \Delta_{2h} y \left[\left[\frac{k}{2} \right] h \right| - \\
- 0,8678 \cdot \Delta_{2h} y \left[\left(\left[\frac{k}{2} \right] - 1 \right) h \right| + & 0,1891 \cdot \Delta_{2h} y \left[\left(\left[\frac{k}{2} \right] - 2 \right) h \right| - \\
- 0,0116 \cdot \Delta_{2h} y \left[\left(\left[\frac{k}{2} \right] - 3 \right) h \right]. & \quad (9.6.23)
\end{aligned}$$

На рис. 9.3 показано графік зміни збурення $v(kT_0)$ відповідно до (9.6.12) при $g=0,85$, а на рис. 9.4 — результати прогнозування координати y на основі

функції (9.6.18) без приростів змінних. Із графіка видно, що якість прогнозування є низькою. Це пояснюється тим, що в збуренні $v(kT_0)$ (рис. 9.3) на окремих проміжках середнє значення відрізняється від нуля. На рис. 9.5 показані результати прогнозування у на основі функції (9.6.23), зображеній у приростах вихідної координати. При цьому підтверджується висока якість прогнозування.

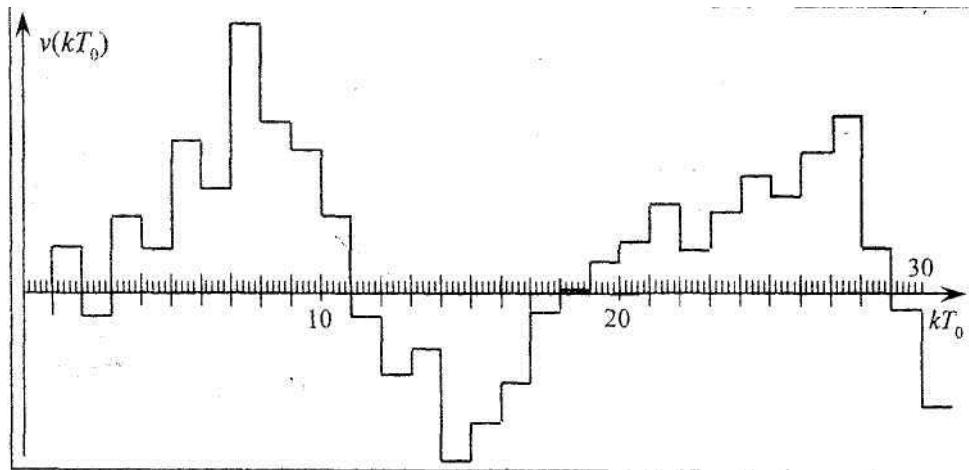


Рис. 9.3. Графік марківського процесу $v[(k+1)T_0] = gv(kT_0) + \xi(kT_0)$ при $g = 0,85$

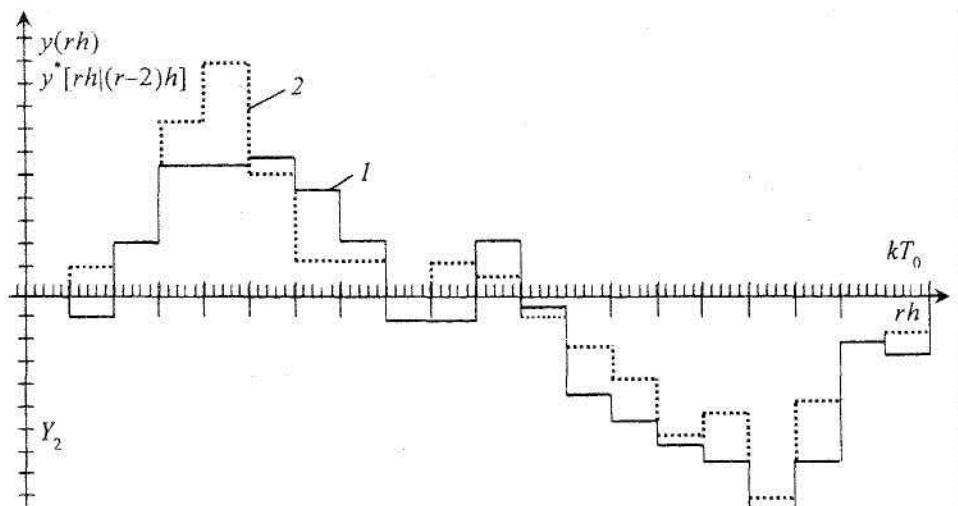


Рис. 9.4. Графіки моделювання вихідної змінної (1) та її прогнозованого значення (2) на основі функції прогнозування без приростів змінних

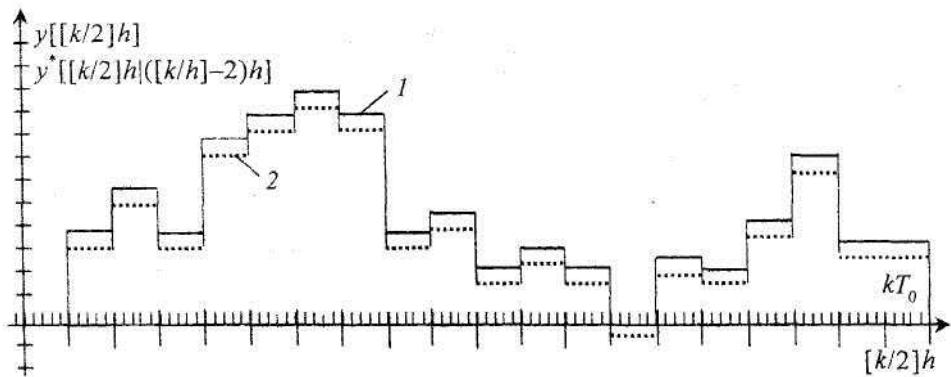


Рис. 9.5. Графіки моделювання вихідної змінної (1) $y\left[\left[\frac{k}{2}\right]h\right]$ та її прогнозованого значення (2) на основі функції прогнозування в приростах змінних

Основна література

1. *Бард Й.* Нелинейное оценивание параметров. – Москва: Статистика, 1979. – 346 с.
2. *Бідюк П.І.* Системний підхід до прогнозування на основі моделей часових рядів // Системні дослідження та інформаційні технології, 2003, № 3, с. 88-110.
3. *Бідюк П.І.* Моделювання і прогнозування гетероскедастичних процесів // Системні дослідження та інформаційні технології, 2004, № 1, с. 115-134.
4. *Бідюк П.І., Савенков О.І., Баклан І.В.* Часові ряди: моделювання і прогнозування. – Київ: ЕКМО, 2003. – 144 с.
5. *Бідюк П.І., Литвиненко В.І.* Моделювання і прогнозування гетероскедастичних процесів // Збірник наукових праць Миколаївського гуманітарного університету ім. Петра Могили, 2004, Вип. 22, с.24-39.
6. *Бідюк П.І., Литвиненко В.І., Помазан Л.В.* Алгоритм ідентифікації стохастичних нелінійних систем на основі генетичного алгоритму // Адаптивні САУ, 2006, № 9(29), с. 3-15.
7. *Бокс Дж., Джсенкінс Г.* Анализ временных рядов. – Москва: Мир, 1974. – 406 с.
8. *Грін В. Г.* Економетричний аналіз. – Київ: Основи, 2005. – 1198 с.
9. *Грооп Д.* Методы идентификации систем. – Москва: Мир, 1979. – 304 с.
10. *Дэннис Дж., Шнабель Р.* Численные методы безусловной оптимизации и решения нелинейных уравнений. – Москва: Мир, 1988. – 440 с.
11. *Зайченко Ю.П.* Основи проектування інтелектуальних систем. – Київ: Слово, 2003. – 351 с.
12. *Згуровский М.З., Подладчиков В.Н.* Аналитические методы калмановской фильтрации. – Київ: Наукова думка, 1995. – 285 с.
13. *Згуровский М.З., Бідюк П.І.* Анализ и управление большими космическими конструкциями. – Київ: Наукова думка, 1997. – 450 с.
14. *Зельнер А.* Байесовские методы в эконометрии. – Москва: Статистика, 1980. – 438 с.
15. *Івахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. – Київ: Наукова думка, 1982. – 296 с.
16. *Кашьян Р.Л., Рао А.Р.* Построение динамических стохастических моделей по экспериментальным данным. – Москва: Наука, 1983. – 385с.
17. *Киричков В.Н.* Идентификация объектов систем управления технологическими процессами. – Київ: Вища школа, 1990. – 263 с.
18. *Кроптя А.В., Бідюк П.І.* Аналіз і методи оцінювання екстремальних значень // Наукові вісті НТУУ «КПІ», 2005, № 4, с. 34 – 47.
19. *Кузовков Н.Т.* Модальное управление и наблюдающие устройства. – М.: Машиностроение, 1976. – 184 с.

20. Лук'яненко І.Г., Городніченко Ю.О. Сучасні економетричні методи у фінансах. – Київ: Літера, 2002. – 350 с.
21. Марпл С.Л. Цифровой спектральный анализ.– Москва: Мир, 1990. – 584 с.
22. Моррис У. Наука об управлении – байесовский подход. – Москва: Мир, 1971. – 304 с.
23. Острем К. Введение в стохастическую теорию управления. — Мир, 1973. — 319 с.
24. Райбман Н.С., Капитоненко В.В., Овсепян Ф.А., Варлаки П.М. Дисперсионная идентификация. – Москва: Наука, 1981. – 336 с.
25. Романенко В.Д. Прогнозирование динамических процессов на основе математических моделей временных рядов с разнотемповой дискретизацией // Системні дослідження та інформаційні технології. — 2005. — № 2. — С. 23—41.
26. Саркисян С.А., Каспин В.И., Лисичкин В.А., Минаев Э.С., Пасечник Г.С. Теория прогнозирования и принятия решений. – Москва: Высшая школа, 1977. – 351 с.
27. Себер Дж. Линейный регрессионный анализ. – Москва: Мир, 1980. – 256 с.
28. Терентьев А.Н., Бидюк П.И. Эвристический метод построения байесовских сетей // Математичні машини і системи, 2006, № 3, с. 12-23.
29. Феллер В. Введение в теорию вероятностей и ее приложения. – Москва: Мир, 1984. – 527 с. (т.1); 751 с. (т.2).
30. Хей Дж. Введение в методы байесовского статистического вывода. – Москва: Финансы и статистика, 1987. – 335 с.
31. Химмельбау Д. Анализ процессов статистическими методами. – Москва: Мир, 1973. – 957 с.
32. Четыркин Е.М., Калихман И.Л. Вероятность и статистика. – Москва: Финансы и статистика, 1982. – 319 с.
33. Эллиотт Р. Стохастический анализ и его приложения. – Москва: Мир, 1986. – 352 с.
34. Ющенко В., Лисицький В. Гроші: розвиток попиту та пропозиції в Україні. – Київ: Скарби, 1998. – 285 с.
35. Andreasen S., Jensen F.V., Olesen K.G. Medical expert systems based on causal probabilistic networks // Int. J. of Biomedical Computing, 1991, v. 28, pp. 1-30.
36. Andrews D.W.K. Asymptotics for semiparametric econometric models via stochastic equicontinuity // Econometrica, 1994, v. 62, № 1, pp. 43-72.
37. Bernardo J.M., Smith A.F.M. Bayesian theory. – New York: John Wiley & Sons, Ltd., 2000. – 586 p.
38. Beveridge S., Nelson C. A new approach to decomposition of economic time series into permanent and transitory components with particular

- attention to measurement of business cycle // J. of Monetary Economics, 1981, v. 7, pp. 151-174.
39. *Casella G., Berger R.L.* Statistical inference. – Pacific Grove, CA: Duxbury, 2002. – 660 p.
 40. *Chow C.K., Liu C.N.* Approximating discrete probability distributions with dependence trees // IEEE Transactions on Information Theory, 1968, v.14. – pp. 462-467.
 41. *Clarke D.W., Hadgson A.J.F., Tufts P.S.* Offset problem and K-incremental predictors in self-turning control // Proceeding IEE. 1983. — 130. — P. 217—226.
 42. *Dahl C.M., Levine M.* Nonparametric estimation of volatility models with serially dependent innovations / TR-09-04, Purdue University, 2004, 31 p.
 43. *Dechter R.* Constraint processing. – San Francisco, CA: Morgan Kaufmann Publishers, 2003. – 479 p.
 44. *Enders W.* Applied Econometric Time Series. – New York: John Wiley & Sons, Inc., 1995. – 450 p.
 45. *Gilks W.R., Richardson S., Spiegelhalter D.J.* Markov Chain Monte Carlo in Practice. – London: Chapman & Hall, 1996. – 486 p.
 46. *Grunwald P.* A tutorial introduction to the minimum description length principle / Advances in minimum description length: Theory and applications MIT Press, Cambridge, MA, USA, 2005. – 80 p.
 47. *Johnston J., DiNardo J.* Econometric methods. - New York: McGraw-Hill, Inc., 1997. – 525 p.
 48. *Klarke D.W., Phil M.A.D., Gawthrop.* Self-tuning controller // Proc. the IEE: Cont. scie. — 1975. — 122, N 9. — P. 929—935.
 49. *Lewis D.C.* Industrial and business forecasting methods. – London: Butterworth Scientific, 1982. – 145 p.
 50. *Silvestrini A.* Temporal Aggregation of Univariate and Multivariate Time Series Models: a Survey / A.Silvestrini, D.Veredas // Bank of Italy Temi de Discussione. — 2008. — Working Paper N 685.
 51. *Tsay R.S.* Analysis of financial time series. – New York: J. Wiley & Sons, Inc., 2002. – 455 p.
 52. *West J.C.* Nonlinear Signal Distortion Correlation // Int. J. Control, 1965, No.2, pp. 529-538.
 53. Vapnik V. Statistical Learning Theory. – New York: Wiley, 1998.
 54. Vapnik V. An Overview of Statistical Learning Theory // IEEE Transactions on Neural Networks, September 1999, 10(5): pp. 988-999.

Додаткова література

1. Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. – М.: Статистика, 1974. – 240 с.
2. Андерсон Т. Статистический анализ временных рядов. – М.: Мир, 1976. – 755 с.
3. Бідюк П.І. Проектування інформаційних СППР. – Київ: НТУУ «КПІ», 2005. – 121 с.
4. Бриллинджер Д. Временные ряды: обработка данных и теория. – М.: Мир, 1980. – 536 с.
5. Баженов В.А., Тимошук О.Л., Палиенко Е.В. Математические модели прогнозирования потребления электроэнергии с помощью временных рядов / Сучасні проблеми електротехніки та автоматики: Доповіді за матеріалами Міжнародної науково-технічної конференції молодих учених , аспірантів і студентів, (Київ, жовтень 2009 р.) — Київ: «Політехніка», 2009, Т.1 — С.276-280.
6. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в науке и технике. – М.: Мир, 1980. – 610 с.
7. Дорохина Е.Ю., Преснякова Л.Ф., Тихомиров Н.П. Сборник задач по эконометрике. – М.: Экзамен, 2003. – 224 с.
8. Дрейпер Н.Р., Смит Г. Прикладной регрессионный анализ. – Киев: Диалектика, 2007. – 912 с.
9. Закс Л. Статистическое оценивание. – М.: Статистика, 1976. – 598 с.
10. Кельберт М.Я., Сухов Ю.М. Основные понятия теории вероятностей и математической статистики. – М.: МЦНМО, 2007. – 456 с.
11. Кендалл М.Дж., Стьюарт А. Теория распределений. – М.: Наука, 1966. – 588 с.
12. Кобзарь А.И. Прикладная математическая статистика. – М.: Физматлит, 2006. – 816 с.
13. Кокс Д., Хинкли Д. Теоретическая статистика. – Москва: Мир, 1978. – 560 с.
14. Ликеш И., Ляга Й. Основные таблицы математической статистики. – М.: Финансы и статистика, 1985. – 358 с.
15. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. – М.: Финансы и статистика, 2003. – 414 с.
16. Терентьев А.Н., Бідюк П.І. Евристический метод построения байесовских сетей / Міжнародна НТК „Інтелектуальні системи підтримки прийняття рішень та прикладі аспекти сучасних інформаційних технологій. – Євпаторія, травень 2006, т. 1, с. 401-403.
17. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика. – М.: Экзамен, 2007. – 512 с.
18. Уилкс С. Математическая статистика. – Москва: Наука, 1967. – 632 с.

19. Хей Дж. Введение в методы байесовского статистического вывода. – М.: Финансы и статистика, 1987. – 335 с.
20. Эддоус М., Стэнсфилд Р. Методы принятия решений. – М.: ЮНИТИ, 1997. – 590 с.
21. Casella G., Berger R.L. Statistical Inference. – Pacific Grove, CA: DUXBURY, 2002. – 660 p.
22. Chow C.K., Liu C.N. Approximating discrete probability distributions with dependence trees. // IEE Transactions on information theory, Vol. IT-14, No. 3, May 1968, 6 pp.
23. Chung K.L., AitSahlia F. Elementary Probability Theory. – New York: Springer, 2003. – 470 p.
24. Cuthbert D., Wood F.S. Fitting Equations to Data. – New York: John Wiley & Sons, 1980. – 458 p.
25. Dean T., Kanazawa K. A model for reasoning about persistence and causation // Computational Intelligence, 1989, v. 5, No. 3, pp. 142-150.
26. Garg S. Controls and health management technologies for intelligent aerospace propulsion systems / NASA-TM, 2004 – 212915. – 28 p.
27. Ghahramani Z., On structural variational inference / Technical Report: CRG-TR-97-1, University of Toronto, CSD, 1997. – 67 p.
28. Grunwald P. A Tutorial Introduction to the Minimum Description Length Principle // Advances in Minimum Description Length: Theory and Applications MIT Press, Cambridge, MA, USA, 2005, – 80 p.
29. Handbook of Human-Computer Interaction / Helander M. (ed.). – Amsterdam: Elsevier Science Publishers B.V. (North Holland), 1988. – 1150 p.
30. Kalman R.E., Bucy R.S. New results in linear filtering and prediction // Journal of Basic Engineering (ASME), 1961, D83, pp. 95-108.
31. Magnus J.R., Neudecker H. Matrix Differential Calculus with Applications. – New York: John Wiley & Sons, 1999. – 496 p.
32. Rabiner L.R. A tutorial on hidden Markov models and selected applications in speech recognition // IEEE, Trans. ASSP, 1989, v. 77, No.2, pp. 257-285.
33. Rauch H.E. Solutions to the linear smoothing problem // IEEE Trans. Automatic Control, 1963, AC-8(4), pp. 371-372.
34. Rossi P.E., Allenby G.M. Bayesian statistics and marketing // Marketing Science, 2003, v. 22, № 3, pp. 304-328.
35. Roweis S., Ghahramani Z. A unifying review of linear Gaussian models // Neural Computation, 1999, v. 11, N0.2, pp. 305-345.
36. Stair R.M. Principles of Information Systems. – Washington: ITPC, 1996. – 656 p.
37. Strait P.T. Probability and Statistics with Applications. – New York: Harcourt Brace Jovanovich, Inc., 1983. – 581 p.
38. Tymoshchuk O.L., K.Paliyenko. Eventual forecast functions for different time-series models / The nonlinear analysis and application 2009: Materials

- of the International scientific conference (April 02-04th 2009, Kyiv). — Kyiv: NTUU “KPI”, 2009. — PP.55.
- 39. Welch G., Bishop G. An introduction to the Kalman filter / Technical report: TR95-041. — University of North Carolina, Chapel Hill, 1995. — 16 p.
 - 40. Whitten J.L., Bentley L.D., Barlow V.M. Systems Analysis & Design Methods. — Boston: Irwin, 1989. — 797 p.

Додаток А

МЕТОДИ ОЦІНЮВАННЯ ПАРАМЕТРІВ МАТЕМАТИЧНИХ І СТАТИСТИЧНИХ МОДЕЛЕЙ

A0.1 Вступ

Однією із важливих складових процесу побудови математичних і статистичних моделей є коректний вибір методів оцінювання параметрів моделей. Коректний вибір методу оцінювання вимагає знання статистичних характеристик часових рядів (або даних іншого типу) і знання структури моделі. Якщо модель лінійна, то частіше всього використовують звичайний метод найменших квадратів (МНК), застосування якого пов'язане з обмеженнями на випадкові процеси, що мають місце при описанні об'єктів та процесів будь-якої природи. Зокрема, це центрованість випадкового процесу та його некорельованість із залежною змінною, а також відсутність автокореляції. Цей розділ присвячено методам лінійного оцінювання – МНК та його рекурсивній версії – РМНК.

Корисною модифікацією МНК, методу максимальної правдоподібності (ММП) та інших методів лінійного оцінювання є їх рекурсивні версії, які дозволяють оцінювати параметри моделей в реальному часі на основі нескінченних вибірок даних. В цьому розділі розглянемо рекурсивну модифікацію МНК, як самого популярного методу оцінювання лінійних моделей (особливо в реальному часі) та виконаємо аналіз збіжності його оцінок.

Дані, які ми отримуємо в результаті виконання експериментів чи від організацій, які забезпечують статистичні дані, – це випадкові величини з детермінованою складовою. Тому необхідно пам'ятати, що оцінки параметрів моделі, обчислені за допомогою випадкових вибірок даних, – це також випадкові величини, які необхідно аналізувати методами, що застосовуються до випадкових процесів. Загалом оцінки параметрів завжди знаходяться в деякому інтервалі значень, який ми називаємо довірчим. Для визначення довірчого інтервалу використовують відомі статистичні методи.

A0.2 Звичайний метод найменших квадратів

Розглянемо лінійний процес, на який діють вхідні впливи $x_1(k), \dots, x_p(k)$, а вихідною змінною є $y(k)$ (рис. A0.1). В регресійному аналізі $y(k)$ розглядається як випадкова і, частіше всього, нормальну розподілена величина, тобто: $\{y(k)\} \sim N(\mu_y, \sigma_y^2)$. Можливість розглядати $y(k)$ як процес з нормальним розподілом дає центральна гранична теорема (ЦГТ), оскільки на вихідну змінну діють велике число впливів $x_1(k), \dots, x_{p-1}(k)$ та невимірювані випадкові збурення, які часто позначають через $\varepsilon(k)$ або $w(k)$.

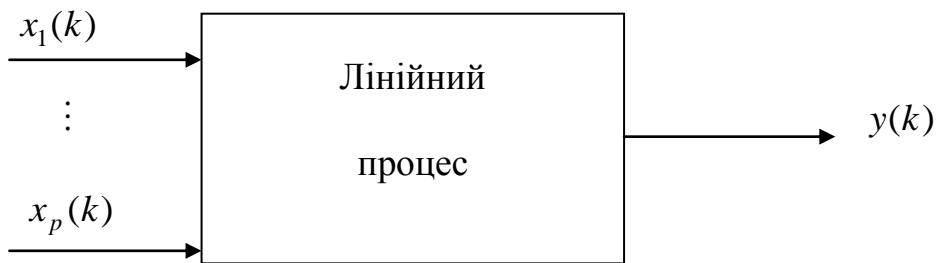


Рис. A0.1. Лінійний процес з p входами і одним виходом

Для описання лінійного процесу, представленого на рис. A0.1., можна скористатись математичною моделлю (множинною регресією) вигляду:

$$y(k) = b_0 + b_1 x_1(k) + b_2 x_2(k) + \dots + b_p x_p(k) + \varepsilon(k), \quad (\text{A0.2.1})$$

де $\varepsilon(k)$ – випадковий процес, природа якого розглянута в розділі 2; b_0, b_1, \dots, b_p – невідомі коефіцієнти моделі, які можна об'єднати у p -вимірний вектор параметрів моделі $\beta = [b_0 \ b_1 \ b_2 \dots b_p]^T$. Якщо зібрати експериментальні дані щодо функціонування даного процесу, тобто, вибірки вимірів довжиною N для кожної змінної, то можна записати систему рівнянь:

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} 1 & x_1(1) & x_2(1) & \cdots & x_p(1) \\ 1 & x_1(2) & x_2(2) & \cdots & x_p(2) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_1(N) & x_2(N) & \cdots & x_p(N) \end{bmatrix} \times \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} \varepsilon(1) \\ \varepsilon(2) \\ \vdots \\ \varepsilon(N) \end{bmatrix} \quad (\text{A0.2.2})$$

або в матричному вигляді:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon, \quad (\text{A0.2.3})$$

де \mathbf{y} – вектор вимірів залежної змінної $y(k)$ вимірності N ; $\mathbf{X}[N \times (p+1)]$ – матриця вимірів незалежних змінних (регресорів або екзогенних змінних); β – вектор невідомих параметрів моделі вимірності $p+1$; ε – випадковий процес з нормальним розподілом, який характеризується нульовим середнім і скінченою постійною дисперсією, тобто: $\{\varepsilon(k)\} \sim N(0, \sigma_\varepsilon^2)$. Для того щоб мінімізувати дисперсії оцінок оцінок параметрів моделі необхідно забезпечити лінійну незалежність стовпчиків матриці \mathbf{X} , тобто, забезпечити максимальний ранг $\text{rank}[\mathbf{X}] = p+1$.

Оцінювання поліноміальної моделі

Якщо процес має один вхід, але між входом і виходом існує нелінійна залежність поліноміального типу, то модель (A0.2.1) приймає вигляд:

$$y(k) = b_0 + b_1 x(k) + b_2 x^2(k) + \dots + b_p x^p(k) + \varepsilon(k), \quad (\text{A0.2.4})$$

тобто, (A0.2.4) – поліном порядку $p-1$. Такі моделі називають ще *псевдолінійними*, оскільки їх параметри можна оцінювати за допомогою звичайного МНК. Модель (A0.2.4) у векторно-матричній формі для N вимірів змінних має вигляд:

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} 1 & x(1) & x^2(1) & \cdots & x^p(1) \\ 1 & x(2) & x^2(2) & \cdots & x^p(2) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x(N) & x^2(N) & \cdots & x^p(N) \end{bmatrix} \times \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} \varepsilon(1) \\ \varepsilon(2) \\ \vdots \\ \varepsilon(N) \end{bmatrix},$$

або

$$\mathbf{y} = \mathbf{X} \theta + \varepsilon. \quad (\text{A0.2.5})$$

Застосуємо ідею методу найменших квадратів до оцінювання моделі (A0.2.4). Для визначення оцінок вектора коефіцієнтів (параметрів) $\theta = [b_0, b_1, b_2, \dots, b_p]^T$ полінома (A0.2.4), які відповідають найкращому наближенню з точки зору мінімума суми квадратів похибок моделі, необхідно знайти мінімум суми (квадратичного функціоналу):

$$\min_{\theta} \sum_{k=1}^N [y(k) - b_0 - b_1 x(k) - \dots - b_p x^p(k)]^2,$$

де $y(k)$ – фактичні виміри залежної змінної; N – число вимірів; $x(k)$ – виміри незалежної змінної.

Це досягається диференціюванням наведеного функціоналу по коефіцієнтах моделі та прирівнюванням похідних нулю. Далі оцінки коефіцієнтів знаходять як розв'язки отриманої системи так званих *нормальних* рівнянь. Похідна по b_0 :

$$-2 \left\{ \sum_{k=1}^N [y(k) - b_0 - b_1 x(k) - \dots - b_p x^p] \right\} = 0$$

або

$$Nb_0 + b_1 \sum_{k=1}^N x(k) + b_2 \sum_{k=1}^N x^2(k) + \dots + b_p \sum_{k=1}^N x^p(k) = \sum_{k=1}^N y(k). \quad (\text{A0.2.6})$$

Похідна по b_1 :

$$-2 \left\{ \sum_{k=1}^N [y(k) - b_0 - b_1 x(k) - \dots - b_p x^p] x(k) \right\} = 0,$$

або

$$b_0 \sum_{k=1}^N x(k) + b_1 \sum_{k=1}^N x^2(k) + b_2 \sum_{k=1}^N x^3(k) + \dots + b_p \sum_{k=1}^N x^{p+1}(k) = \sum_{k=1}^N x(k) y(k). \quad (\text{A0.2.7})$$

Похідні по інших коефіцієнтах:

$$b_0 \sum_{k=1}^N x^2(k) + b_1 \sum_{k=1}^N x^3(k) + b_2 \sum_{k=1}^N x^4(k) + \dots + b_p \sum_{k=1}^N x^{p+2}(k) = \sum_{k=1}^N x^2(k) y(k). \quad (\text{A0.2.8})$$

.

.

$$b_0 \sum_{k=1}^N x^p(k) + b_1 \sum_{k=1}^N x^{p+1}(k) + b_2 \sum_{k=1}^N x^{p+2}(k) + \dots + b_p \sum_{k=1}^N x^{2p}(k) = \sum_{k=1}^N x^p(k) y(k). \quad (\text{A0.2.9})$$

Оцінки вектора параметрів $\theta = [b_0, b_1, b_2, \dots, b_p]^T$ полінома (A0.2.4) знайдемо в результаті розв'язку системи нормальних рівнянь (A0.2.6) – (A0.2.9). Таким чином, матриця вимірів незалежних змінних для системи нормальних рівнянь має вигляд:

$$\mathbf{X}' = \begin{bmatrix} N & \sum_{k=1}^N x(k) & \sum_{k=1}^N x^2(k) & \dots & \sum_{k=1}^N x^p(k) \\ \sum_{k=1}^N x(k) & \sum_{k=1}^N x^2(k) & \sum_{k=1}^N x^3(k) & \dots & \sum_{k=1}^N x^{p+1}(k) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^N x^p(k) & \sum_{k=1}^N x^{p+1}(k) & \sum_{k=1}^N x^{p+2}(k) & \dots & \sum_{k=1}^N x^{2p}(k) \end{bmatrix}.$$

Тепер систему нормальних рівнянь для N вимірів змінних можна записати у векторно-матричній формі так:

$$\mathbf{y}' = \mathbf{X}' \theta. \quad (\text{A0.2.10})$$

Звідси знайдемо вираз для вектора оцінок параметрів моделі:

$$\hat{\theta} = [\mathbf{X}']^{-1} \mathbf{y}', \quad (\text{A0.2.11})$$

де $\mathbf{y}' = \left[\sum_1^N y(k) \quad \sum_1^N x(k)y(k) \quad \dots \quad \sum_1^N x^p(k)y(k) \right]^T$; $\hat{\theta}$ – вектор оцінок параметрів моделі. Тобто оцінку вектора параметрів за методом найменших квадратів можна знайти шляхом розв'язку системи лінійних (нормальних) рівнянь.

Модель типу

$$y(k) = b_0 + b_1 e^{-b_2 x(k)} + \varepsilon(k)$$

називають *нелінійною відносно параметрів*, оскільки вона містить адитивний член, в якому є добуток одного параметра на інший. Для оцінювання моделей такого типу необхідно застосувати нелінійний МНК або метод максимальної правдоподібності, хоча в багатьох випадках такі моделі можна лінеаризувати шляхом логарифмування.

Застосування МНК до оцінювання регресійної моделі загального вигляду
 $\mathbf{y} = \mathbf{X}\theta + \varepsilon$

Отримаємо загальний вираз для обчислення оцінок параметрів регресійної моделі $\mathbf{y} = \mathbf{X}\theta + \varepsilon$ із умови мінімізації квадратичного критерію – суми квадратів похибок моделі. Оскільки змінна $\varepsilon(k)$ – невимірювана величина (її оцінку, $\hat{\varepsilon}(k) = e(k)$), ми знайдемо після оцінювання параметрів

моделі), то мінімізація суми квадратів похибок моделі означає мінімізацію скаляра:

$$\varepsilon^T \varepsilon = \| \mathbf{y} - \hat{\mathbf{y}} \|^2, \quad (\text{A0.2.12})$$

де $\hat{\mathbf{y}} = \mathbf{X}\hat{\theta}$ – оцінка залежності змінної, отримана за допомогою вимірювань незалежних змінних та оцінок параметрів моделі $\hat{\theta}$.

Для того щоб отримати вираз для обчислення вектора оцінок параметрів моделі скористаємося виразом: $\varepsilon = \mathbf{y} - \mathbf{X}\theta$. Таким чином, маємо:

$$\begin{aligned} \varepsilon^T \varepsilon &= (\mathbf{y} - \mathbf{X}\theta)^T (\mathbf{y} - \mathbf{X}\theta) = \\ &= \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\theta - \theta^T \mathbf{X}^T \mathbf{y} + \theta^T \mathbf{X}^T \mathbf{X}\theta = \mathbf{y}^T \mathbf{y} - 2\theta^T \mathbf{X}^T \mathbf{y} + \theta^T \mathbf{X}^T \mathbf{X}\theta. \end{aligned}$$

Тепер необхідно взяти часткову похідну цього виразу по θ і прирівняти її нулю:

$$\frac{\partial \varepsilon^T \varepsilon}{\partial \theta} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X}\theta = 0,$$

а звідси

$$\mathbf{X}^T \mathbf{X}\theta = \mathbf{X}^T \mathbf{y}$$

або

$$\theta = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y}. \quad (\text{A0.2.13})$$

Мінімальне значення скаляра $\varepsilon^T \varepsilon$ називають залишковою сумою квадратів похибок (*RSS = residual square sum*).

Після обчислення оцінок вектора параметрів моделі можна знайти фактичну похибку рівняння:

$$\begin{aligned} \mathbf{e} &= \min_{\beta} [\varepsilon^T \varepsilon] = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\theta = \mathbf{y} - \mathbf{X}[\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y} = \\ &= \{\mathbf{I}_N - \mathbf{X}[\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T\} \mathbf{y} = (\mathbf{I}_N - \mathbf{P}) \mathbf{y}, \end{aligned} \quad (\text{A0.2.14})$$

де $\mathbf{P} = \mathbf{X}[\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T$; $(\mathbf{I}_N - \mathbf{P})$ – матриця ортогональної проекції (вектори \mathbf{e} і \mathbf{y} – ортогональні внаслідок того, що θ обчислюється через похідну); $\mathbf{P} [N \times N]$ – матриця лінійного перетворення експериментальних даних, оскільки

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\theta} = \mathbf{X}[\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y} = \mathbf{P}\mathbf{y}.$$

Таким чином, вектор вимірів залежної змінної можна представити у вигляді ортогонального розкладу на дві складові:

$$\mathbf{y} = \mathbf{P}\mathbf{y} + (\mathbf{I}_N - \mathbf{P})\mathbf{y}. \quad (\text{A0.2.15})$$

A0.3 Властивості матриць \mathbf{P} та $\mathbf{I}_N - \mathbf{P}$

Сформулюємо властивості матриць \mathbf{P} та $\mathbf{I}_N - \mathbf{P}$ у вигляді теореми:

Теорема A0.1.

- (1) Матриці \mathbf{P} та $\mathbf{I}_N - \mathbf{P}$ симетричні та ідемпотентні (рівнопотужні, тобто, $\mathbf{A}^2 = \mathbf{A}$). (Ідемпотента матриця вимірності $N \times N$ має $k < N$ одиничних власних значень, а інші $N - k$ дорівнюють нулю; лише одинична матриця є ідемпотентною матрицею повного рангу.)
- (2) $\text{rank } [\mathbf{I}_N - \mathbf{P}] = \text{tr}[\mathbf{I}_N - \mathbf{P}] = N - p$.
- (3) $[\mathbf{I}_N - \mathbf{P}]\mathbf{X} = 0$.

Доведення

- (1) Матриця \mathbf{P} симетрична за визначенням, а

$$[\mathbf{I}_N - \mathbf{P}]^T = \mathbf{I}_N - \mathbf{P}^T = \mathbf{I}_N - \mathbf{P}.$$

Крім того, $\mathbf{P}^2 = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \cdot \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{X} \mathbf{I}_p (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{P}$.

$$(\mathbf{I}_N - \mathbf{P})^2 = \mathbf{I}_N - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{I}_N - \mathbf{P}.$$

- (2) Оскільки матриця $\mathbf{I}_N - \mathbf{P}$ симетрична та ідемпотентна, то

$$\text{rank } [\mathbf{I}_N - \mathbf{P}] = \text{tr}[\mathbf{I}_N - \mathbf{P}] = N - \text{tr}\mathbf{P} = N - p,$$

оскільки

$$\text{tr}\mathbf{P} = \text{tr}[\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T] = \text{tr}[\mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}] = \text{tr}\mathbf{I}_p = p.$$

(3) $(\mathbf{I}_N - \mathbf{P})\mathbf{X} = \mathbf{X} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} = \mathbf{X} - \mathbf{X} = 0$, а це свідчить про те, що похибки моделі не залежать від \mathbf{X} .

A0.4 Властивості оцінок, отриманих за МНК

1. Незміщеність. Якщо математичне сподівання випадкового процесу $E[\varepsilon(k)] = 0$, то оцінки параметрів будуть незміщені, тобто, не будуть містити методичної похибки, яка зменшує або збільшує оцінки параметрів на всьому інтервалі існування вибірки даних.

Якщо $E[\varepsilon(k)] = 0$, то $E[\mathbf{y}] = E[\mathbf{X}\theta + \varepsilon] = \mathbf{X}\theta$, а тому

$$E[\hat{\theta}] = E[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}] = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \cdot E[\mathbf{y}] = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \cdot \mathbf{X}\theta = \theta. \quad (\text{A0.4.1})$$

Таким чином, $\hat{\theta}$ є незміщеною оцінкою θ .

2. Дисперсія оцінок. Якщо похибки моделі $\varepsilon(k)$ некорельовані і мають незмінну в часі скінченну дисперсію, тобто,

$$E[\varepsilon(k)\varepsilon(l)] = \begin{cases} \sigma_\varepsilon^2, & k = l, \\ 0, & k \neq l; \end{cases} \quad \sigma_\varepsilon^2 \leq c, \quad c = \text{const} > 0, \quad (\text{A0.4.2})$$

то дисперсія вектора ε визначається так: $\text{var}[\varepsilon] = \sigma_\varepsilon^2 \mathbf{I}_N$.

Дисперсія вектора оцінок параметрів моделі:

$$\begin{aligned} \text{cov}[\hat{\theta}] &= \text{cov}[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}] = E[\theta \theta^T] = E\{[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \cdot \mathbf{y}^T \mathbf{X}^T (\mathbf{X}^T \mathbf{X})^{-1}]\} = \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}^T (\mathbf{X}^T \mathbf{X})^{-1} \cdot E[\mathbf{y} \mathbf{y}^T] = (\mathbf{X}^T \mathbf{X})^{-1} \sigma_\varepsilon^2, \end{aligned} \quad (\text{A0.4.3})$$

тобто, діагональні елементи матриці $(\mathbf{X}^T \mathbf{X})$, помножені на σ_ε^2 будуть дисперсіями вектора параметрів моделі.

Примітка. Для отримання виразу (A0.4.3) ми скористалися наступною теоремою: якщо \mathbf{x} і \mathbf{y} – випадкові вектори, а \mathbf{A} і \mathbf{B} – постійні матриці відповідної вимірності, то математичне сподівання

$$E[\mathbf{A}\mathbf{x}, \mathbf{B}\mathbf{y}] = \mathbf{A} E[\mathbf{x}, \mathbf{y}] \mathbf{B}^T = \mathbf{A} E[\mathbf{x}\mathbf{y}^T] \mathbf{B}^T.$$

3. В своєму класі методів оцінювання оцінка $\hat{\theta}$, отримана за МНК, має найменшу дисперсію. В зв'язку з цим оцінки МНК називають найкращими лінійними незміщеними оцінками (НЛНО). Доведення цього твердження можна знайти, наприклад, в книзі Себера Дж. [...].

4. Розподіл оцінок МНК. Наведені вище статистичні властивості оцінок, отриманих за МНК, дають можливість сформулювати визначення їх розподілу: оскільки $\{y(k)\} \sim N(\mathbf{X}\theta, \sigma_e^2 \mathbf{I}_N)$, то

$$\hat{\theta} \sim N_p(\theta, \sigma_e^2 (\mathbf{X}^T \mathbf{X})^{-1}). \quad (\text{A0.4.4})$$

Аналіз якості оцінок, обчислених за іншими методами, можна виконати подібним підходом.

A0.5 Рекурсивний метод найменших квадратів

Розглянемо різницеве рівняння (лінійна регресія) типу

$$\begin{aligned} y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) &= b_1 v(k-1) + b_2 v(k-2) + \dots \\ &\quad b_m v(k-m) + v(k), \end{aligned} \quad (\text{A0.5.1})$$

де $y(k)$ – залежна змінна; $v(k)$ – збурення (шум) випадкового характеру, яке у більшості випадків розглядається як білий шум. Відзначимо, що випадкова змінна $v(k)$ включає в себе ті похибки, які не пояснюються прийнятою моделлю (A0.5.1). Модель (A0.5.1) можна представити у векторному вигляді як

$$y(k) = \theta^T \psi(k) + v(k), \quad (\text{A0.5.2})$$

де $\theta^T = [a_1 \dots a_n \ b_1 \dots b_m]$ – вектор параметрів моделі; вектор вимірів визначено як

$$\psi^T(k) = [-y(k-1) - y(k-2) \dots - y(k-n) \ v(k-1) \ v(k-2) \dots v(k-m)]. \quad (\text{A0.5.3})$$

Вектор параметрів θ необхідно оцінити за допомогою вимірів вектора $\psi(k)$, які зареєстровані в моменти часу $k = 1, 2, 3, \dots, N$. Природно визначити оцінки параметрів моделі, виходячи із умови мінімізації «помилки

рівняння». Тобто, можна записати функціонал для мінімізації у наступному вигляді

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N \alpha_k [y(k) - \theta^T \psi(k)]^2, \quad (\text{A0.5.4})$$

який мінімізується по відношенню до вектора параметрів θ , тобто необхідно знайти оцінку цього вектора параметрів з наступної умови:

$$\min_{\theta} J_N(\theta).$$

В функціоналі (A0.5.4) α_k – послідовність додатніх чисел, які відіграють роль вагових коефіцієнтів для вимірів; $y(k)$ – скалярний вимір в момент k . Досить часто це послідовність одиниць, або вибір значень цих коефіцієнтів зв'язують з дисперсією шумової складової $v(k)$.

Величину $\hat{y}(k | \theta) = \theta^T \psi(k)$ можна розглядати як «прогнозоване» значення для $y(k)$, що ґрунтуються на векторі параметрів θ . Таким чином, критерій (A0.5.4) можна розглядати як спробу вибрати таку модель, яка дасть змогу визначити найкращий (на один крок) прогноз вихідного сигналу. Критерій $J_N(\theta)$ є квадратичним по відношенню до вектора θ , а тому його можна мінімізувати аналітично.

Запишемо похідну критерію оптимізації по вектору параметрів

$$\begin{aligned} \frac{\partial J(\theta)}{\partial \theta} &= \frac{2}{N} \sum_{k=1}^N \alpha_k [y(k) - \theta^T \psi(k)] \left(-\frac{\partial \theta^T \psi(k)}{\partial \theta} \right) = \\ &= \frac{2}{N} \sum_{k=1}^N \alpha_k [y(k) - \theta^T \psi(k)] \left(-\frac{\partial \psi^T(k) \theta}{\partial \theta} \right) = \\ &= \frac{2}{N} \sum_{k=1}^N \alpha_k [y(k) - \theta^T \psi(k)] \left(-\psi^T(k) \frac{\partial \theta}{\partial \theta} \right). \end{aligned}$$

Оскільки в правій частині похідна вектора параметрів по векторному аргументу буде матрицею, то для приведення у відповідність вимірностей множників в круглих дужках в правій частині ми скористалися рівністю $\theta^T \psi(k) = \psi^T(k) \theta$.

Прирівняємо похідну по вектору параметрів нулю

$$\sum_{k=1}^N \alpha_k [y(k) - \theta^T \psi(k)] [-\psi^T(k)] = 0$$

або

$$\sum_{k=1}^N \alpha_k \theta^T \psi(k) \psi^T(k) = \sum_{k=1}^N \alpha_k y(k) \psi^T(k),$$

де $\psi(k) \psi^T(k)$ – симетрична інформаційна матриця Фішера, тобто $[\psi(k) \psi^T(k)] = [\psi(k) \psi^T(k)]^T$.

Якщо транспонувати обидві частини останнього рівняння, то отримаємо

$$\sum_{k=1}^N \alpha_k |\psi(k) \psi^T(k)|^T \theta = \sum_{k=1}^N \alpha_k \psi(k) y(k),$$

що еквівалентно

$$\sum_{k=1}^N \alpha_k |\psi(k) \psi^T(k)| \theta = \sum_{k=1}^N \alpha_k \psi(k) y(k).$$

Введемо позначення $\hat{\theta}$ для оцінки вектора параметрів θ і визначимо його за допомогою наступного рівняння:

$$\hat{\theta}(N) = \left[\sum_{k=1}^N \alpha_k \psi(k) \psi^T(k) \right]^{-1} \sum_{k=1}^N \alpha_k \psi(k) y(k). \quad (\text{A0.5.5})$$

при умові, що існує обернена матриця.

Перейдемо до рекурсивної форми запису останнього рівняння. Введемо наступне позначення:

$$\bar{\mathbf{R}}(k) = \sum_{k=1}^N \alpha_k \psi(k) \psi^T(k).$$

Із визначення величини $\bar{\mathbf{R}}(k)$ витікає, що

$$\bar{\mathbf{R}}(k) = \bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k) \quad \text{або} \quad \bar{\mathbf{R}}(k-1) = \bar{\mathbf{R}}(k) - \alpha_k \psi(k) \psi^T(k).$$

Таким чином, для вектора оцінок параметрів можна записати наступний вираз:

$$\begin{aligned}
\hat{\theta}(k) &= \bar{\mathbf{R}}^{-1}(k) \left[\sum_{i=1}^{k-1} \alpha_i \psi(i) y(i) + \alpha_k \psi(k) y(k) \right] = \\
&= \bar{\mathbf{R}}^{-1}(k) \left[\sum_{i=1}^{k-1} \alpha_i \psi(i) (\psi^T(i) \hat{\theta}(i-1)) + \alpha_k \psi(k) y(k) \right] = \\
&= \bar{\mathbf{R}}^{-1}(k) \left[\bar{\mathbf{R}}(k-1) \hat{\theta}(k-1) + \alpha_k \psi(k) y(k) \right] = \\
&= \bar{\mathbf{R}}^{-1}(k) \left[\bar{\mathbf{R}}(k) \hat{\theta}(k-1) + \alpha_k \psi(k) [y(k) - \psi^T(k) \hat{\theta}(k-1)] \right] \quad (\text{A0.5.6})
\end{aligned}$$

Останній вираз отримано із врахуванням того, що допоміжна змінна $\bar{\mathbf{R}}(k-1) = \bar{\mathbf{R}}(k) - \alpha_k \psi(k) \psi^T(k)$.

Тепер виконаємо множення в правій частині (A0.5.6) і в результаті отримаємо рекурсивні рівняння для вектора оцінок параметрів моделі (A0.5.1)

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \alpha_k \bar{\mathbf{R}}^{-1}(k) \psi(k) [y(k) - \hat{\theta}^T(k-1) \psi(k)], \quad (\text{A0.5.7})$$

$$\bar{\mathbf{R}}(k) = \bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k). \quad (\text{A0.5.8})$$

Для того, щоб уникнути переповнення при обчисленні оцінок параметрів за допомогою системи (A0.5.7), (A0.5.8), матрицю $\bar{\mathbf{R}}(k)$ (її називають ще інформаційною матрицею Фішера) усереднюють наступним чином:

$$\mathbf{R}(k) = \frac{1}{k} \bar{\mathbf{R}}(k) = \frac{1}{k} [\bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k)] \quad (\text{A0.5.9})$$

Оскільки $\bar{\mathbf{R}}(k-1) = (k-1)\mathbf{R}(k-1)$, то (A0.5.9) можна записати як

$$\begin{aligned}
\mathbf{R}(k) &= \frac{k-1}{k} \mathbf{R}(k-1) + \frac{1}{k} \alpha_k \psi(k) \psi^T(k) = \\
&= \mathbf{R}(k-1) + \frac{1}{k} [\alpha_k \psi(k) \psi^T(k) - \mathbf{R}(k-1)]. \quad (\text{A0.5.10})
\end{aligned}$$

Таким чином друга форма для РМНК має наступний вигляд:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{1}{k} \alpha_k \mathbf{R}^{-1}(k) \psi(k) \left[y(k) - \hat{\theta}^T(k-1) \psi(k) \right], \quad (\text{A0.5.11})$$

$$\mathbf{R}(k) = \mathbf{R}(k-1) + \frac{1}{k} \left[\alpha_k \psi(k) \psi^T(k) - \mathbf{R}(k-1) \right]. \quad (\text{A0.5.12})$$

Недоліком обох наведених форм є необхідність обчислення оберненої матриці, що неможливо у випадку, коли матриця $\mathbf{R}(k)$ вироджена. Тому розглянемо ще одну форму РМНК, яка не потребує обернення матриці.

Алгоритм РМНК без операції обернення матриці

Введемо матрицю $\mathbf{P}(k) = \bar{\mathbf{R}}^{-1}(k) = \frac{1}{k} \mathbf{R}^{-1}(k)$ і скористаємося нею замість $\mathbf{R}(k)$. При цьому для отримання рекурсивних спiввiдношень скористаємося також наступною лемою про обернення матриць:

$$[\mathbf{A} + \mathbf{BCD}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} [\mathbf{DA}^{-1} \mathbf{B} + \mathbf{C}^{-1}]^{-1} \mathbf{DA}^{-1}. \quad (\text{A0.5.13})$$

Доведення леми виконується шляхом домноження обох частин (A0.5.13) на матрицю $[\mathbf{A} + \mathbf{BCD}]$.

Введемо наступні позначення:

$$\mathbf{A} = \mathbf{P}^{-1}(k-1), \quad \mathbf{B} = \psi(k), \quad \mathbf{C} = \alpha_k, \quad \mathbf{D} = \psi^T(k)$$

і перепишемо вираз $\bar{\mathbf{R}}(k) = \bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k)$ при $\bar{\mathbf{R}}(k) = \mathbf{P}^{-1}(k)$ у вигляді:

$$\mathbf{P}^{-1}(k) = \mathbf{P}^{-1}(k-1) + \alpha_k \psi(k) \psi^T(k) \text{ або } \mathbf{P}(k) = [\mathbf{P}^{-1}(k-1) + \psi(k) \alpha_k \psi^T(k)]^{-1}.$$

Праву частину останнього виразу можна формально поставити у вiдповiднiсть лiвiй частинi (A0.5.13). В результатi отримаємо:

$$\mathbf{P}(k) = \mathbf{P}(k-1) - \mathbf{P}(k-1) \psi(k) \left[\psi^T(k) \mathbf{P}(k-1) \psi(k) + \frac{1}{\alpha_k} \right]^{-1} \psi^T(k) \mathbf{P}(k-1) =$$

$$= \mathbf{P}(k-1) - \frac{\mathbf{P}(k-1)\psi(k)\psi^T(k)\mathbf{P}(k-1)}{1/\alpha_k + \psi^T(k)\mathbf{P}(k-1)\psi(k)}. \quad (\text{A0.5.14})$$

Використовуючи такий же підхід, знайдемо вираз для (A0.5.11) без операції обернення матриці:

$$\begin{aligned} \gamma(k) &= \frac{1}{k} \mathbf{R}^{-1}(k) \psi(k) \alpha_k = \alpha_k \mathbf{P}(k) \psi(k) = \\ &= \alpha_k \mathbf{P}(k-1) \psi(k) - \frac{\alpha_k \mathbf{P}(k-1) \psi(k) \psi^T(k) \mathbf{P}(k-1) \psi(k)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)} = \\ &= \frac{\mathbf{P}(k-1) \psi(k)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)}. \end{aligned} \quad (\text{A0.5.15})$$

Таким чином РМНК має наступну форму:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \gamma(k)[y(k) - \hat{\theta}^T(k-1)\psi(k)]; \quad (\text{A0.5.16})$$

$$\gamma(k) = \frac{\mathbf{P}(k-1) \psi(k)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)}; \quad (\text{A0.5.17})$$

$$\mathbf{P}(k) = \mathbf{P}(k-1) - \frac{\mathbf{P}(k-1) \psi(k) \psi^T(k) \mathbf{P}(k-1)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)}. \quad (\text{A0.5.18})$$

Алгоритм (A0.5.16) – (A0.5.18) часто використовується для оцінювання параметрів об'єктів та процесів в реальному часі. Кожний новий вимір використовується для уточнення оцінок параметрів математичної моделі.

Приклад оцінювання параметрів моделі першого порядку.
Розглянемо послідовність операцій, що виконуються в процесі ідентифікації параметрів об'єкта першого порядку за допомогою РМНК.

Скористаємося різницевим рівнянням першого порядку

$$y(k) + a_1 y(k-1) = b_1 v(k-1) + v(k).$$

Для побудови алгоритму ідентифікації модель представляють у вигляді

$$\hat{y}(k-1) = \psi^T(k-1) \hat{\theta}(k-1).$$

де $\psi^T(k-1) = [-y(k-1) \ u(k-1)]$, $\hat{\theta}(k-1) = [\hat{a}_1(k-1) \ \hat{b}_1(k-1)]^T$.

Алгоритм оцінювання вектора параметрів $\hat{\theta}(k)$ складається із наступних кроків:

1. Реєстрація (введення) нових значень вимірів $y(k)$, $v(k)$, які відповідають моменту k . В даному випадку значення $v(k)$ – більш, а значення $y(k)$ обчислюються по моделі.
2. Знайти поточну похибку моделі

$$e(k) = y(k) - \hat{y}(k-1) = y(k) - [-y(k-1) \ v(k-1)] \begin{bmatrix} \hat{a}_1(k-1) \\ \hat{b}_1(k-1) \end{bmatrix}.$$

3. Обчислити нові значення оцінок параметрів

$$\begin{bmatrix} \hat{a}_1(k) \\ \hat{b}_1(k) \end{bmatrix} = \begin{bmatrix} \hat{a}_1(k-1) \\ \hat{b}_1(k-1) \end{bmatrix} + \begin{bmatrix} \gamma_1(k-1) \\ \gamma_2(k-1) \end{bmatrix} e(k).$$

4. Сформувати новий вектор вимірів

$$\psi^T(k) = [-y(k) \ u(k)].$$

5. Обчислити проміжний вектор

$$\begin{aligned} \mathbf{P}(k) \psi(k) &= \begin{bmatrix} p_{11}(k) & p_{12}(k) \\ p_{21}(k) & p_{22}(k) \end{bmatrix} \begin{bmatrix} -y(k) \\ v(k) \end{bmatrix} = \\ &= \begin{bmatrix} -p_{11}(k)y(k) + p_{12}(k)v(k) \\ -p_{21}(k)y(k) + p_{22}(k)v(k) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \mathbf{r}. \end{aligned}$$

6. Обчислити проміжний скаляр

$$\psi^T(k) \mathbf{P}(k) \psi(k) = [-y(k) \ v(k)] \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = -r_1 y(k) + r_2 v(k) = s.$$

7. Нове значення вектора гамма

$$\begin{bmatrix} \gamma_1(k) \\ \gamma_2(k) \end{bmatrix} = \frac{1}{s + \lambda} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}, \quad \lambda = \alpha_k^{-1}.$$

8. Обчислити нову інформаційну матрицю

$$\begin{aligned} \mathbf{P}(k+1) &= \frac{1}{\lambda} [\mathbf{P}(k) - \gamma(k) \psi^T(k) \mathbf{P}(k)] = \\ &= \frac{1}{\lambda} [\mathbf{P}(k) - \gamma(k) [\mathbf{P}(k) \psi(k)]^T] = \\ &= \frac{1}{\lambda} [\mathbf{P}(k) - \gamma(k) \mathbf{r}^T] = \frac{1}{\lambda} \begin{bmatrix} p_{11}(k) - \gamma_1 r_1 & p_{12}(k) - \gamma_1 r_2 \\ p_{21}(k) - \gamma_2 r_1 & p_{22}(k) - \gamma_2 r_2 \end{bmatrix}. \end{aligned}$$

9. Збільшити k на одиницю: $k = k+1$ і перейти на п.1.

Для того щоб можна було почати обчислення в момент $k = 0$, необхідно присвоїти змінним наступні початкові значення:

$$\hat{\theta}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \quad \mathbf{P}(0) = \begin{bmatrix} p_{11}(0) & 0 \\ 0 & p_{22}(0) \end{bmatrix},$$

де $p_{11}(0) \gg 0$, $p_{22}(0) \gg 0$. Тобто початкові значення діагональних елементів матриці $\mathbf{P}(0)$ задаються досить великими додатніми числами (наприклад, кілька сотень). Якщо є інформація щодо значень коефіцієнтів моделі, то нею можна скористатись для того щоб скоротити тривалість переходного процесу. Наприклад, якщо ми встановили, що кореляція між сусідніми значеннями ряду додатна, тобто, $\text{cov}[y(k) y(k-1)] > 0$, то коефіцієнту a_1 можна присвоїти невелике додатне значення: $a_1(0) = 0,01$. Такий підхід, як правило, дозволяє скоротити тривалість переходного процесу. Значення параметра λ вибирають наступним чином: якщо коефіцієнти рівняння, що оцінюються, не змінюються в часі, то $\lambda = 0,95 \div 0,98$ (саме такий випадок розглядається в лабораторній роботі); якщо коефіцієнти рівняння змінюються в часі, то $\lambda = 0,6 \div 0,7$.

Загалом, тривалість переходного процесу при оцінюванні параметрів математичної моделі залежить від числа невідомих параметрів (коефіцієнтів) моделі. Якщо в моделі невідомі 1-2 коефіцієнти, то необхідно не більше 20-30 кроків (значень вимірювань) для обчислення їх точних значень.

Якщо необхідно оцінити 3-4 коефіцієнти, то знадобиться 60-80 кроків. Найбільше число кроків для оцінювання параметрів моделі необхідно виконати у випадку, коли оцінюються параметри багатовимірної моделі, тобто, коли число рівнянь моделі більше одиниці. В такому випадку число кроків (число разів виконання алгоритму РМНК) може сягати кількох сотень і навіть тисяч.

Метод найменших квадратів забезпечує виконання основних вимог щодо якості оцінок – *незміщеність, консистентність і ефективність*, якщо випадкова величина у правій частині рівняння задовольняє наступним вимогам: $\{\varepsilon(k)\}$ – некорельзований процес з нульовим середнім і некорельзований із залежною змінною в лівій частині рівняння. Небхідно також пам'ятати, що якість оцінок, отриманих за допомогою РМНК, наближається до якості оцінок, отриманих за допомогою МНК, після закінчення переходного процесу. Оцінити статистичні характеристики процесу $\{\varepsilon(k)\}$ можна після оцінювання коефіцієнтів моделі, тобто $\hat{\varepsilon}(k) = e(k)$, де $e(k)$ – похибка моделі. Це означає, що перед оцінюванням моделі ми робимо припущення, що $\{\varepsilon(k)\}$ задовольняє висунутим вимогам, а після оцінювання коефіцієнтів необхідно зробити перевірку цього припущення.