

Лабораторна робота №1

Методичні вказівки до виконання

Оцінювання параметрів різницевого рівняння авторегресії із ковзним середнім на основі експериментальних даних за допомогою МНК і РМНК

Мета роботи: навчитися оцінювати параметри (коефіцієнти) різницевого рівняння (РР) типу авторегресії з заданим порядком за допомогою (а) методу найменших квадратів (МНК) та (б) рекурсивного методу найменших квадратів (РМНК). Проаналізувавши отримані результати необхідно (1) визначити найкращий метод оцінювання параметрів РР, (2) обрати «найкращу» модель для конкретного процесу за допомогою множини статистичних параметрів, що характеризують якість моделі.

ЗМІСТ

	Стор.
<u>Завдання на виконання роботи</u>	1
<u>1. Загальний краткий опис методів МНК та РМНК</u>	6
<u>2 Теоретичні положення</u>	8
<u>2.1 Рекурсивний метод найменших квадратів</u>	8
<u>2.2 Алгоритм РМНК без операції обернення матриці</u>	12
<u>3 Приклад оцінювання параметрів моделі першого порядку</u>	14
<u>Контрольні запитання</u>	18
<u>Література</u>	19
<u>Післямова</u>	20

Завдання на виконання роботи

1. Ознайомитися з теоретичними положеннями та прикладом оцінювання параметрів рівняння першого порядку за допомогою РМНК. Вибрати одну з книг додаткової літератури та прочитати розділ про рекурсивне оцінювання параметрів. Оцінювання параметрів математичних і статистичних моделей називають ще **параметричною ідентифікацією**.

2. Вибрати мову програмування (Delphi, C++) або систему Matlab, MathCad і написати програму, яка реалізує алгоритми МНК та РМНК. При цьому в програмі передбачити:

- можливість «вставки» різницевого рівняння з відомими коефіцієнтами, яке імітує (описує) динаміку об'єкта; при цьому порядок відомого рівняння може коливатися від авторегресії з ковзним середнім АРКС(1,1) до АРКС(3,3); (тобто мається на увазі що користувач задає 1. порядок авторегресії, 2. порядок ковзного середнього, 3. коефіцієнти $a_0, a_1, a_2, a_3, b_1, b_2, b_3$ що наведені в табл. 1 згідно з номером бригади).
- можливість завантаження часового ряду з текстового файлу (для перевірки роботоспроможності програми по тестовим даним викладача).

Таблиця 1

№ бригади	a_0	a_1	a_2	a_3	b_1	b_2	b_3
1	0	0,22	-0,18	0,08	0,5	0,25	0,25
2	0,05	0,45	0,09	-0,38	0,6	0,4	0
3	0,08	0,25	-0,25	-0,19	0,7	0,2	0,1
4	0	0,22	-0,18	0,08	0,5	0,5	0
5	0,1	0,135	-0,207	0,315	1	0	0
6	0	-0,045	-0,079	0,525	0,3	0,4	0,3
7	0	0,33	0,15	0,11	0,5	0	0,5
8	0	0,15	-0,33	0,25	0,3	0,7	0
9	0	-0,35	0,12	0,47	0,1	0,8	0,1
10	0,3	0,15	0,13	0	0,22	0,78	0
11	0,4	0,05	-0,05	0,5	0,4	0,5	0,1
12	0,33	0,3	0,35	0,37	0,15	0,85	0
13	0,1	0,21	0,32	0,43	0,4	0,4	0,2

– обчислення сумми квадратів похибок моделі, параметри якої оцінюються за допомогою МНК та РМНК, тобто $S = \sum_{k=1}^N e^2(k)$, де N – об’єм вибірки даних ($N \geq 100$);

– обчислення коефіцієнта детермінації $R^2 = \frac{\text{var}(\hat{y})}{\text{var}(y)}$, де $\text{var}(\hat{y})$ – дисперсія залежної змінної, оціненої за допомогою побудованої математичної моделі; $\text{var}(y)$ – дисперсія залежної змінної, оцінена за допомогою її фактичних значень (для хорошої моделі $R^2 \rightarrow 1$). Як обчислити значення $e(k)$ та $\hat{y}(k)$ показано в прикладі оцінювання, який наведено в кінці описання цієї роботи;

– обчислення критерію Акайке для визначення адекватності моделі процесу

$$IKA = N \ln \left(\sum_{k=1}^N e^2(k) \right) + 2n,$$

де $n = p + q + 1$ – число параметрів моделі, що оцінюються за алгоритмом РМНК (в системі *Eviews* цей критерій позначається через *AIC*); N – об’єм вибірки даних. Наприклад, якщо оцінюються параметри рівняння $AR(2)$:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + v(k),$$

$$\text{то } n = 2 + 0 + 1.$$

Тобто число параметрів авторегресії дорівнює 2, число параметрів ковзного середнього дорівнює 0, а +1 враховує оцінювання константи a_0 . Для кращої моделі IKA приймає мінімальне значення.

3. Напишіть процедуру для генерування некорельованих нормально розподілених псевдовипадкових чисел (типу «білого» шуму) для подачі на вход об’єкта. Тобто, ці числа відіграватимуть роль змінної $u(k)$ (або $v(k)$). Якщо ви використовуєте рівняння з ковзним середнім, то немає необхідності ще раз додавати шумову складову $v(k)$, тому що шум наявний у змінній

$u(k)$. Таким чином, **випадкова змінна вводиться в рівняння тільки один раз!**

Нормально розподілена некорельована послідовність псевдовипадкових чисел має широкий частотний спектр (якщо обчислити для неї перетворення Фур'є), а тому її називають “білим” шумом по аналогії з сонячним світлом, який також має широкий частотний спектр складових.

Якщо інструментальна система, в якій ви програмуєте алгоритм РМНК, дозволяє генерувати тільки рівномірно розподілені числа, то для отримання нормально розподілених чисел можна скористатись, наприклад, простою формулою:

$$\varepsilon_n = \sum_{i=1}^{12} v_i - 6,$$

де v_i , $i = 1, \dots, 12$ – рівномірно розподілені числа в діапазоні від -1 до +1; ε_n – нормально розподілене число з нульовим середнім та одиничною дисперсією. Тобто, на одне нормально розподілене число необхідно згенерувати 12 рівномірно розподілених чисел.

Крім *білого шуму* для достатнього збудження на вхід процесу можна подавати інші широкополосні сигнали, тобто, сигнали, які містять великий набір частотних складових. До них відносяться: **одиничний імпульс** та **псевдовипадкова двійкова послідовність**. Перетворення Фур'є цих сигналів також свідчить про те, що вони мають широкий частотний спектр, а тому є високоінформативними з точки зору наповнення гармонічними складовими.

4. Знайдіть оцінки параметрів (коефіцієнтів) рівняння для значень порядку моделі по авторегресії від 1 до 3, і по ковзному середньому від 1 до 3. Всього досить розглянути 9 різних моделей (АРКС(1,1), АРКС(1,2),...,АРКС(3,3)) відповідно кожним методом (МНК та РМНК). Виберіть «об'єкт» із наведеної вище таблиці 1 у відповідності з номером бригади.

Необхідно зауважити що коефіцієнти моделей наведені в табл.1 задовільняють умовам стійкості моделей.

Нагадаємо, що необхідна умова стійкості моделі:

$$\sum_i a_i < 1,$$

а достатньою умовою стійкості моделі є наступна:

$$\sum_i |a_i| < 1.$$

Обидві суми не повинні включати коефіцієнт a_0 .

5. Побудуйте графік зміни у часі оцінки для кожного коефіцієнта математичної моделі (*графік перехідного процесу алгоритму оцінювання*). Графік буде мати коливальний характер при невеликих значеннях дискретного часу k , а при збільшенні k оцінка коефіцієнта повинна прямувати до його точного значення. Тривалість перехідного процесу оцінювання параметрів рівняння залежить від числа коефіцієнтів, що оцінюються за допомогою алгоритмів МНК та РМНК.

6. Нарисуйте графіки зміни коефіцієнта детермінації, суми квадратів похибок рівняння (залишків) та значення критерію Акайке в залежності від порядку рівняння (кількості параметрів), що оцінюється.

7. За допомогою розрахованих характеристик виберіть «найкращу» модель. Обґрунтуйте вибір «найкращої» моделі. Поясніть чому для вибору кращої моделі використовують декілька статистичних параметрів, а не один параметр?

8. По отриманим результатам зробить висновок який метод МНК або РМНК дає кращі результати? Поясніть чому в техніці використовується саме метод РМНК, в чому головний недолік МНК?

7. Оформіть протокол лабораторної роботи (один на бригаду, теоретичні положення роздруковувати не потрібно) і дайте письмові

висновки за виконаною роботою. В файлі *03_report_example.doc* в якості прикладу наведена бажана структура протокола.

8. Дайте (усно) відповіді на контрольні запитання.

1. Загальний краткий опис методів МНК та РМНК

Нехай ми маємо АРКС(3,3) записане у загальному вигляді

$$y(k) = a_0 + a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + a_3 \cdot y(k-3) + \\ + v(k) + b_1 \cdot v(k-1) + b_2 \cdot v(k-2) + b_3 \cdot v(k-3) + \varepsilon$$

В матричному вигляді це рівняння можна записати як

$$Y = X \cdot \theta + \varepsilon$$

де Y – вектор спостережень,

$X = (1, y(k-1), y(k-2), y(k-3), v(k), v(k-1), v(k-2), v(k-3))$ – матриця

вимірів, $\theta = (a_0, a_1, a_2, a_3, 1, b_1, b_2, b_3)^T$ – параметри (коефіцієнти) АРКС(3,3), а ε – вектор похибок.

Метод МНК оцінювання параметрів АРКС(3,3) записується як

$$\theta = (X^T X)^{-1} X^T Y.$$

Вивід методу МНК можна знайти в будь-якій книзі з економетрики або регресійного аналізу, наприклад в [7] глава 6.

Для тих студентів, що будуть робити програму на *Delphi* в папці *DelphiUnit* лежить файл *matrices.pas*, в якому реалізовані процедури та функції для роботи з матрицями.

Приклад

Var X, X_tr: Matrix;

Begin

```

// створення масиву X
X.M:=кількість строк;
X.N:=кількість стовпчиків;
SetSize(X, X.M, X.N);

// створення масиву X_tr
X_tr.M:=X.N;
X_tr.N:=X.M;
SetSize(X_tr, X_tr.M, X_tr.N);

// транспонування масиву X. Результат операції зберігається в X_tr
X_tr:=Transpose(X);
End;

```

З іншими функціями і процедурами можна ознайомитися в самому файлі *matrices.pas*

Метод РМНК

Початкові значення алгоритму $\theta_0 = \alpha, P_0 = \beta \cdot I, \beta \gg 0$

i – й крок рекурсії

$$P_i = P_{i-1} - \frac{P_{i-1} \cdot x^T(i) \cdot x(i) \cdot P_{i-1}}{1 + x(i) \cdot P_{i-1} \cdot x^T(i)}$$

$$\theta_i = \theta_{i-1} + P_i \cdot x^T(i) \cdot [y^T(i) - x(i) \cdot \theta_{i-1}]$$

де $x(i)$ – це i – та строка матриці вимірів.

Математичне обґрунтування методу РМНК описується в другому розділі цього документа, а також в файлі *03_WenY_RMNK.pdf* [8].

В файлі *02_info_rmnk.doc* наведені приклади використання методів МНК та РМНК, а в папці *rmnk_matlab7*, для ознайомлення в якості прикладу, знаходяться матлабовські файли для методу РМНК.

2 Теоретичні положення

2.1 Рекурсивний метод найменших квадратів

Розглянемо різницеве рівняння (лінійна регресія) типу

$$\begin{aligned} y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) = \\ = b_1 v(k-1) + b_2 v(k-2) + \dots + b_m v(k-m) + v(k), \end{aligned} \quad (2.1)$$

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 (2.1). Модель (2.1) можна представити у векторному вигляді як

$$y(k) = \theta^T \psi(k) + v(k), \quad (2.2)$$

де $\theta^T = [a_1 \dots a_n \ b_1 \dots b_m]$ – вектор параметрів моделі; вектор вимірів визначено як

$$\psi^T(k) = [-y(k-1) - y(k-2) \dots - y(k-n) \ v(k-1) \ v(k-2) \dots v(k-m)]. \quad (2.3)$$

Вектор параметрів θ необхідно оцінити за допомогою вимірів вектора $\psi(k)$, які зареєстровані в моменти часу $k = 1, 2, 3, \dots, N$. Природньо визначити оцінки параметрів моделі, виходячи із умови мінімізації «помилки рівняння». Тобто, можна записати функціонал для мінімізації у наступному вигляді

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N \alpha_k [y(k) - \theta^T \psi(k)]^2, \quad (2.4)$$

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 (2.4) 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 базується на векторі параметрів θ . Таким чином, критерій (2.4) можна розглядати як спробу вибрати таку модель, яка дасть змогу визначити найкращий прогноз (на один крок) 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 скористались 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

$$\sum_{k=1}^N \alpha_k \left| \psi(k) \psi^T(k) \right|^T \theta = \sum_{k=1}^N \alpha_k \psi(k) y(k),$$

що еквівалентно

$$\sum_{k=1}^N \alpha_k \left| \psi(k) \psi^T(k) \right| \theta = \sum_{k=1}^N \alpha_k \psi(k) y(k).$$

Введемо позначення $\hat{\theta}$ для оцінки вектора параметрів θ і визначимо його за допомогою наступного рівняння:

$$\hat{\theta}(N) = \left[\sum_{k=1}^N \alpha_k \psi(k) \psi^T(k) \right]^{-1} \sum_{k=1}^N \alpha_k \psi(k) y(k). \quad (2.5)$$

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

$$\bar{\mathbf{R}}(k) = \bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k) \text{ або } \bar{\mathbf{R}}(k-1) = \bar{\mathbf{R}}(k) - \alpha_k \psi(k) \psi^T(k).$$

Таким чином, для вектора оцінок параметрів можна записати наступний вираз:

$$\begin{aligned}
\hat{\theta}(k) &= \bar{\mathbf{R}}^{-1}(k) \left[\sum_{i=1}^{k-1} \alpha_i \psi(i) y(i) + \alpha_k \psi(k) y(k) \right] = \\
&= \bar{\mathbf{R}}^{-1}(k) \left[\sum_{i=1}^{k-1} \alpha_i \psi(i) (\psi^T(i) \hat{\theta}(i-1)) + \alpha_k \psi(k) y(k) \right] = \\
&= \bar{\mathbf{R}}^{-1}(k) [\bar{\mathbf{R}}(k-1) \hat{\theta}(k-1) + \alpha_k \psi(k) y(k)] = \\
&= \bar{\mathbf{R}}^{-1}(k) [\bar{\mathbf{R}}(k) \hat{\theta}(k-1) + \alpha_k \psi(k) [y(k) - \psi^T(k) \hat{\theta}(k-1)]] . \quad (2.6)
\end{aligned}$$

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 (2.6) і в результаті отримаємо рекурсивні рівняння для вектора оцінок параметрів моделі (2.1)

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \alpha_k \bar{\mathbf{R}}^{-1}(k) \psi(k) [y(k) - \hat{\theta}^T(k-1) \psi(k)], \quad (2.7)$$

$$\bar{\mathbf{R}}(k) = \bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k). \quad (2.8)$$

Для того, щоб уникнути переповнення при обчисленні оцінок параметрів за допомогою системи (2.7), (2.8), матрицю $\bar{\mathbf{R}}(k)$ (її називають ще інформаційною матрицею Фішера) усереднюють наступним чином:

$$\mathbf{R}(k) = \frac{1}{k} \bar{\mathbf{R}}(k) = \frac{1}{k} [\bar{\mathbf{R}}(k-1) + \alpha_k \psi(k) \psi^T(k)] \quad (2.9)$$

Оскільки $\bar{\mathbf{R}}(k-1) = (k-1)\mathbf{R}(k-1)$, то (2.9) можна записати як

$$\begin{aligned}\mathbf{R}(k) &= \frac{k-1}{k}\mathbf{R}(k-1) + \frac{1}{k}\alpha_k\psi(k)\psi^T(k) = \\ &= \mathbf{R}(k-1) + \frac{1}{k}[\alpha_k\psi(k)\psi^T(k) - \mathbf{R}(k-1)].\end{aligned}\quad (2.10)$$

Таким чином друга форма для РМНК має наступний вигляд:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{1}{k}\alpha_k\mathbf{R}^{-1}(k)\psi(k)[y(k) - \hat{\theta}^T(k-1)\psi(k)],\quad (2.11)$$

$$\mathbf{R}(k) = \mathbf{R}(k-1) + \frac{1}{k}[\alpha_k\psi(k)\psi^T(k) - \mathbf{R}(k-1)].\quad (2.12)$$

Недоліком обох наведених форм є необхідність обчислення оберненої матриці, що неможливо, якщо матриця $\mathbf{R}(k)$ вироджена. Тому розглянемо ще одну форму РМНК, яка не потребує обернення матриці.

2.2 Алгоритм РМНК без операції обернення матриці

Введемо матрицю $\mathbf{P}(k) = \bar{\mathbf{R}}^{-1}(k) = \frac{1}{k}\mathbf{R}^{-1}(k)$ і скористаємось нею замість $\mathbf{R}(k)$. При цьому для отримання рекурсивних співвідношень скористаємось також наступною лемою про обернення матриць:

$$[\mathbf{A} + \mathbf{BCD}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}[\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1}]^{-1}\mathbf{DA}^{-1}.\quad (2.13)$$

Доведення леми виконується шляхом домноження обох частин (2.13) 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 відповідність лівій частині (2.13). В результаті отримаємо:

$$\begin{aligned} \mathbf{P}(k) &= \mathbf{P}(k-1) - \mathbf{P}(k-1) \psi(k) \left[\psi^T(k) \mathbf{P}(k-1) \psi(k) + \frac{1}{\alpha_k} \right]^{-1} \psi^T(k) \mathbf{P}(k-1) = \\ &= \mathbf{P}(k-1) - \frac{\mathbf{P}(k-1) \psi(k) \psi^T(k) \mathbf{P}(k-1)}{1/\alpha_k + \psi^T(k) \mathbf{P}(k-1) \psi(k)}. \end{aligned} \quad (2.14)$$

Використовуючи такий же підхід, знайдемо вираз для (2.11) без операції обернення матриці:

$$\begin{aligned} \gamma(k) &= \frac{1}{k} \mathbf{R}^{-1}(k) \psi(k) \alpha_k = \alpha_k \mathbf{P}(k) \psi(k) = \\ &= \alpha_k \mathbf{P}(k-1) \psi(k) - \frac{\alpha_k \mathbf{P}(k-1) \psi(k) \psi^T(k) \mathbf{P}(k-1) \psi(k)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)} = \\ &= \frac{\mathbf{P}(k-1) \psi(k)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)}. \end{aligned} \quad (2.15)$$

Таким чином РМНК має наступну форму:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \gamma(k) [y(k) - \hat{\theta}^T(k-1) \psi(k)]; \quad (2.16)$$

$$\gamma(k) = \frac{\mathbf{P}(k-1) \psi(k)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)}; \quad (2.17)$$

$$\mathbf{P}(k) = \mathbf{P}(k-1) - \frac{\mathbf{P}(k-1) \psi(k) \psi^T(k) \mathbf{P}(k-1)}{\alpha_k^{-1} + \psi^T(k) \mathbf{P}(k-1) \psi(k)}. \quad (2.18)$$

Алгоритм (2.16) – (2.18) часто використовується для оцінювання параметрів об'єктів та процесів в реальному часі. Кожний новий вимір використовується для уточнення оцінок параметрів математичної моделі.

3 Приклад оцінювання параметрів моделі першого порядку

Розглянемо послідовність операцій, що виконуються в процесі ідентифікації параметрів об'єкта першого порядку за допомогою РМНК.

Скористаємось різницеvim 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

$$\hat{y}(k-1) = \psi^T(k-1) \hat{\theta}(k-1).$$

$$\text{де } \psi^T(k-1) = [-y(k-1) \ u(k-1)], \quad \hat{\theta}(k-1) = [\hat{a}_1(k-1) \ \hat{b}_1(k-1)]^T.$$

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 $v(k)$ – білий шум, 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

$$\mathbf{P}(k+1) = \frac{1}{\lambda} [\mathbf{P}(k) - \gamma(k) \psi^T(k) \mathbf{P}(k)] =$$

$$= \frac{1}{\lambda} [\mathbf{P}(k) - \gamma(k) [\mathbf{P}(k) \psi(k)]^T] =$$

$$= \frac{1}{\lambda} [\mathbf{P}(k) - \gamma(k) \mathbf{r}^T] = \frac{1}{\lambda} \begin{bmatrix} p_{11}(k) - \gamma_1 r_1 & p_{12}(k) - \gamma_1 r_2 \\ p_{21}(k) - \gamma_2 r_1 & p_{22}(k) - \gamma_2 r_2 \end{bmatrix}.$$

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 перехідного 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 перехідного процесу. Значення параметра λ вибирають наступним чином: якщо коефіцієнти рівняння, що оцінюється, 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 перехідного 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 вимірів) 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: $\{\varepsilon(k)\}$ – некорельований процес з нульовим середнім і некорельований із залежною змінною в лівій частині рівняння. Необхідно також пам'ятати, що якість оцінок, отриманих за допомогою РМНК, наближається до якості оцінок, отриманих за допомогою МНК, після закінчення перехідного 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.

Контрольні запитання

1. В чому полягає принципова відмінність рекурсивної процедури оцінювання параметрів від нерекурсивної?
2. Які умови отримання ефективних оцінок параметрів математичних моделей за допомогою РМНК, (порівняйте із звичайним МНК) ?
3. Порівняйте якість оцінок, отриманих за допомогою МНК і РМНК.
4. Які статистичні характеристики використовують для вибору кращої моделі?

5. Чому для виконання параметричної ідентифікації використовують *білий* шум і чому його називають *білим*?
6. Що означає термін «кольоровий шум»?
7. Які інші інформативні сигнали можна подавати на вхід об'єкту для того щоб отримати інформативний сигнал на виході і чому?
8. Поясніть вибір початкових умов для алгоритму РМНК?
9. Як змінюється тривалість перехідного процесу оцінювання параметрів при зміні порядку моделі?
10. Які параметри можна використати для контролю правильності процесу функціонування алгоритму РМНК?
11. Доведіть лему про обернення матриць (в тексті документа це [формула 2.13](#)).
12. Назвіть інші, відомі вам алгоритми рекурсивної ідентифікації?
13. Наведіть приклади використання РМНК в реальному часі.

Література

1. Бідюк П.І., Савенков О.І. Часові ряди: моделювання і прогнозування. – Київ: ЕКМО, 2003. – 144 с.
2. Дрейпер Н. Р., Смит. Г. Прикладной регрессионный анализ, 3-е изд.: Пер. с англ. – М.: Издательский дом “Вильямс”, 2007. – 912 с.
3. Wen Y. Recersive Least square // Control Automático, CINVESTAV-IPN. – 2003. – 9 p.

Післямова

У найбільш кмітливих студентів при виконанні даної лабораторної роботи можуть з'явитися наступні питання.

1. Чи дійсно v являється ковзним середнім по відношенню до y ?

Насправді, скоріш за все v не являється дійсним ковзним середнім по відношенню до y , тому що в даній роботі спочатку генерується v , а вже по ньому створюється часовий ряд.

Якщо робити за наукою, то v повинен обчислюватися на основі вихідного сигналу y або на основі залишків авторегресійної моделі. (Це буде розглядатися в наступних лабораторних роботах.)

Але неможна в одній роботі розглядати одразу багато нових незрозумілих тем. Тому при виконанні роботи вважайте що v це ковзне середнє або сприймайте його просто як регресор.

2. Чому методи МНК та РМНК дають погані оцінки коефіцієнтів?

Скоріш за все проблема полягає в тому що коли ви генеруєте y то використовуєте формулу

$$y(k) = a_0 + a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + a_3 \cdot y(k-3) + v(k) + b_1 \cdot v(k-1) + b_2 \cdot v(k-2) + b_3 \cdot v(k-3) + \varepsilon$$

А коли оцінюєте коефіцієнти $a_0, a_1, a_2, a_3, b_1, b_2, b_3$ то забуваєте про $v(k)$. Тому необхідно зробити одне з двох:

1. Або ввести ще один коефіцієнт, наприклад $coef$ при $v(k)$ і відповідно оцінювати вже коефіцієнти $a_0, a_1, a_2, a_3, coef, b_1, b_2, b_3$

$$y(k) = a_0 + a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + a_3 \cdot y(k-3) + \\ + coef \cdot v(k) + b_1 \cdot v(k-1) + b_2 \cdot v(k-2) + b_3 \cdot v(k-3) + \varepsilon$$

Якщо все зроблено коректно то після оцінювання $coef \rightarrow 1$.

2. Або зробити попереднє допоміжне перетворення. Це перетворення полягає в позбутті $v(k)$. Перенісіть $v(k)$ в ліву частину, та введіть новий вектор $y1(k) = y(k) - v(k)$. Після чого вже оцінюйте $a_0, a_1, a_2, a_3, b_1, b_2, b_3$ відносно нового рівняння:

$$y1(k) = y(k) - v(k) = a_0 + a_1 \cdot y(k-1) + a_2 \cdot y(k-2) + a_3 \cdot y(k-3) + \\ + b_1 \cdot v(k-1) + b_2 \cdot v(k-2) + b_3 \cdot v(k-3) + \varepsilon$$