

МІНІСТЕРСТВО ОСВІТИ І НАУКИ  
МОЛОДІ ТА СПОРТУ УКРАЇНИ  
МІЖНАРОДНИЙ ЕКОНОМІКО-ГУМАНІТАРНИЙ  
УНІВЕРСИТЕТ ІМЕНІ АКАДЕМІКА  
СТЕПАНА ДЕМ'ЯНЧУКА

# Р.М.ЛІТНАРОВИЧ

ПОБУДОВА І ДОСЛІДЖЕННЯ МАТЕМАТИЧНОЇ  
МОДЕЛІ ЗА ДЖЕРЕЛАМИ ЕКСПЕРИМЕНТАЛЬНИХ  
ДАНИХ МЕТОДАМИ РЕГРЕСІЙНОГО АНАЛІЗУ

НАВЧАЛЬНИЙ ПОСІБНИК



Рівне, 2011

УДК 51-7:519.87

Літнарів Р.М. Побудова і дослідження математичної моделі за джерелами експериментальних даних методами регресійного аналізу. Навчальний посібник, МЕНУ, Рівне, 2011.-140 с.

Litnarovich R. M. Construction and research of mathematical model on the sources of experimental data by the methods of regressive analysis. Train aid. IEGU, Rivne, 2011.-140 p

Рецензенти: В.Г.Бурачек, доктор технічних наук, професор  
С.С. Парняков, доктор технічних наук, професор  
В.О.Боровий, доктор технічних наук, професор  
Відповідальний за випуск: Й.В. Джунь, доктор фізико-математичних наук, професор

Регресійний аналіз є основним статистичним методом побудови математичних моделей об'єктів або явищ по експериментальним даним. Основні результати в даний час отримані стосовно до лінійних регресійних моделей.

**Ключові слова:** математична модель регресійний аналіз, експеримент, опрацювання матеріалів.

Регрессионный анализ является основным статистическим методом построения математических моделей объектов или явлений по экспериментальным данным. Основные результаты в настоящее время получены относительно линейных регрессионных моделей.

**Ключевые слова:** математическая модель регрессионный анализ, экспериментобработка материалов.

A regressive analysis is the basic statistical method of construction of mathematical models of objects or phenomena for to experimental information. It is presently got basic results in relation to linear regressive models.

**Keywords:** a mathematical model is a regressive analysis, experiment, working of materials.

© Літнарів Р.М.

ЗМІСТ	Стор.
Передмова.....	5
Розділ 1. Задачі регресійного аналізу.....	6
1.1.Регресійні моделі.....	6
1.2. Приклади приведення задачі опрацювання експериментальних даних до задачі регресійного аналізу.....	13
1.2.1. Задача оцінки імпульсної характеристики системи.....	13
1.2.2 Задача оцінки коефіцієнтів різницевого рівняння дискретної моделі системи.....	16
Розділ 2. Основні положення класичного регресійного аналізу.....	20
2.1. Базові (фундаментальні) положення класичного лінійного регресійного аналізу.....	20
2.2. Класичні оцінки параметрів регресії методом найменших квадратів і їх властивості.....	22
2.3. Рекурентний алгоритм методу найменших квадратів.....	27
2.4. Статистичний аналіз якості регресійної моделі.....	31
Розділ 3. Обчислювальні алгоритми методу найменших квадратів.....	40
3.1. Загальна характеристика чисельних методів.....	40
3.2. Вплив помилок округлення і похибок вихідних даних.....	44
3.3. Прямі методи рішення лінійних рівнянь.....	48
3.4. Методи LU – розкладання. Метод Гауса.....	51
3.5. Метод квадратних коренів.....	58
3.6. Методи QR-розкладення.....	60
3.7. Метод сингулярного розкладення.....	63
3. 8. Порівняння методів.....	66
Розділ 4. Регресійний аналіз найпростіших поліноміальних моделей .....	67
4.1. Поліноміальна модель нульового порядку .....	67
4.2. Поліноміальна модель першого порядку.....	70
4.3. Поліноміальна модель другого порядку.....	73
Розділ 5. Особливості регресійного аналізу при	

порушенні базових положень.....	74
5.1. Регресійний аналіз при неоднорідних і корельованих збуреннях.....	74
5.2.Регресійний аналіз в умовах мультиколінеарності.....	78
5.2.1. Метод псевдо-обертання.....	80
5.2.2 Метод регуляризації.....	84
5. 3. Вибір найкращої структури регресійної моделі.....	89
5.3.1. Побудова і перевірка всіх можливих регресій.....	90
5.3.2. Метод виключення.....	91
5.3.3 Метод крокової регресії.....	92
5.4. Регресійний аналіз в умовах похибок в регресорах.....	93
Розділ 6. Результати експериментальних досліджень.....	96
6.1. Побудова математичної моделі класичним МНК.....	96
6.2. Оцінка точності зрівноваженої моделі.....	103
6.3. Побудова математичної моделі узагальненим методом найменших квадратів (УМНК).....	119
Висновки по .....	132
Літературні джерела.....	133

## ПЕРЕДМОВА

Регресійний аналіз є основним статистичним методом побудови математичних моделей об'єктів або явищ по експериментальним даним

Основним завданням регресійного аналізу являється отримання оцінок параметрів регресії  $(\hat{\beta}_1, \dots, \hat{\beta}_m)$ , які були б оптимальними в певному сенсі. Отримані оцінки, які будемо представляти у вигляді компонентів вектора  $\hat{\beta} = [\hat{\beta}_1, \dots, \hat{\beta}_m]^T$ , дозволяють вирішувати задачу оцінки (відновлення) регресії і її прогнозу.

Для застосування узагальненого методу найменших квадратів (УМНК) необхідно знати ковариаційну матрицю вектора збурень  $\Omega$ , що зустрічається вкрай рідко на практиці. Якщо ж вважати всі  $n(n+1)/2$  елементів симетричної ковариаційної матриці  $\Omega$ , невідомими параметрами узагальненої моделі (в доповнення до  $(m+1)$  параметрам  $\beta_i$ ), то загальне число параметрів значно перевищить число спостережень  $n$ , що зробить оцінку цих параметрів нерозрешимою задачею.

Встановлено, що при застосуванні кореляційної матриці узагальненого методу найменших квадратів, середня квадратична похибка одиниці ваги зменшилась у шість раз.

. Проведена оцінка параметрів регресійної моделі «зваженим» методом найменших квадратів, що приводить в кінцевому рахунку до ефективних оцінок параметрів моделі.

. Для узагальненої регресійної моделі, на відміну від класичної, коефіцієнт детермінації не являється задовільною мірою якості моделі. В загальному випадку  $R^2$  може виходити навіть за межі інтервалу  $[0;1]$ , а додавання (видалення) пояснюючої змінної не обов'язково приводить до його збільшення (зменшення).

Для магістрантів ВНЗ, аспірантів і здобувачів наукового ступеня

## Розділ 1. Задачі регресійного аналізу.

### 1.1. Регресійні моделі.

Регресійний аналіз є основним статистичним методом побудови математичних моделей об'єктів або явищ по експериментальним даним. Ці моделі зв'язують кількісні змінні –*результуючу і пояснючі* (в різних областях використовуються різні назви цих змінних: вихід, цільова, вихідна або результуюча ознака, відгук, результативна, ендогенна або залежна змінна; факторні, прогнозні, екзогенні змінні, фактори, фактор-аргументи, предикат ори та інш. [1, 3, 46, 78].

Відмітимо до речі, що визначувана в ході аналізу функція регресії лише формально встановлює відповідність між змінними цих двох груп, хоча вони в дійсності можуть і не бути в причинно-наслідкових стосунках. Тому встановлювані в ході регресійного аналізу зв'язки можуть іноді помилково тлумачитися як причинно-наслідкові. Таким чином, можуть виникнути так звані нонсенс регресії (помилкові, абсурдні), які не мають практичного сенсу [1,78]. З цієї причини перед застосуванням статистичного апарату на основі професійно-логічного аналізу проблеми необхідно вирішити, яку із змінних розглядати як результуючу, а які з реєстрованих величин - як пояснючі.

Розглянемо загальну схему регресійного аналізу [72].. Нехай результуюча змінна  $Y$  пов'язана з деякими пояснючими змінними  $x_1, \dots, x_k$ , які зручно представляти у вигляді компонент вектора  $x = [x_1, \dots, x_k]^T$ , (т- транспонування). Зв'язок являється стохастичним: значення у змінної  $Y$ , отримані в різних експериментах

при фіксованих значеннях вектора  $x$ , випадковим чином флюктує навколо деякого невідомого рівня  $\eta(x)$ :

$$Y \equiv Y(x) = \eta(x) + \varepsilon, \quad (1.1)$$

де друга складова визначає випадкове відхилення результуючої змінної від величини  $\eta(x)$ . Випадкові відхилення  $\varepsilon$  можуть слугувати проявом впливу не врахованих у векторі  $x$  (і може бути, випадкових) факторів, випадковими похибками вимірів результуючої змінної та іншими причинами, які більш детально будуть обмірковуватися нижче. Середнє значення відхилень приймається рівним нулю, тому математичне сподівання результуючої змінної співпадає зі значенням функції  $\eta(x)$ .

$$M\{Y(x)\} = \eta(x). \quad (1.2)$$

Це рівняння називається регресією (рівнянням регресії), а функція  $\eta(x)$  - функцією регресії (ФР).

Існує велике число типів регресійних моделей, визначених видом функції регресії (ФР)  $\eta(x)$ , які, як правило, залежать не тільки від пояснюючих змінних, але і від деяких параметрів  $\beta_1, \dots, \beta_m$ , які також зручно представляти у вигляді векторів  $\beta = [\beta_1, \dots, \beta_m]^T$ :

$$\eta(x) = \eta(x, \beta). \quad (1.3)$$

Для таких функцій регресії (ФР) задача їх визначення зводиться до задачі оцінки вектора параметрів  $\beta$  по експериментальним даним. В залежності від того, як ці параметри входять у функцію регресії (ФР), моделі діляться на лінійні і нелінійні (по параметрам).

Основні результати в даний час отримані стосовно до лінійних регресійних моделей, які в загальному вигляді можна записати слідуючим чином:

$$Y(x) = \eta(x, \beta) + \varepsilon = \sum_{j=1}^m \beta_j f_j(x) + \varepsilon, \quad (1.4)$$

де  $f_j(x) \equiv f_j(x_1, \dots, x_k)$  - деякі відомі функції пояснюючих змінних, що не включають в себе невідомі коефіцієнти  $\beta_j$ . Функції  $f_j(x)$  називають регресорами.

Цю модель зручно представити у векторній формі:

$$Y \equiv Y(x) = f^T \beta + \varepsilon, \quad (1.5)$$

де  $f^T = [f_1(x), \dots, f_m(x)]^T$  - вектор регресорів.

Особливо часто використовуються поліноміальні регресори, зокрема, так звана лінійна (по факторам) поліноміальна регресійна модель:

$$Y = \sum_{j=1}^m \beta_j x_j + \varepsilon = x^T \beta + \varepsilon. \quad (1.6)$$

Якщо  $M\{Y\} = 0$ , то розмірності векторів  $x$  і  $\beta$  рівні числу пояснюючих змінних ( $k = m$ ). В протилежному випадку розмірності векторів приймають на одиницю більше ( $m = k + 1$ ): в число пояснюючих змінних вводять фіктивну змінну, рівну одиниці ( $x = [1 \ x_1, \dots, x_k]^T$ ). Тоді  $\beta_1$  дорівнює значенню функції регресії (ФР) при нульових значеннях дійсних факторів і модель (1.6) приймає вигляд



- матриця регресорів;  $\varepsilon = [\varepsilon_1, \dots, \varepsilon_n]^T$  - вектор відхилень;

$\eta = [\eta(x_1), \dots, \eta(x_n)]^T$  - вектор регресії.

Наприклад, для лінійної моделі спостережень (1.7) матриця регресорів, яку також називають матрицею плану експеримента, має вигляд

$$F = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1\{m-1\}} \\ 1 & x_{21} & x_{22} & \dots & x_{2\{m-1\}} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{n\{m-1\}} \end{bmatrix}, \quad (1.12)$$

де  $x_{ij}$  - значення  $j$ -го фактора в  $i$ -му експерименті.

Для поліноміальної моделі (1.8)

$$F = \begin{bmatrix} 1 & x_1 & x_{12} & \dots & x_1^{m-1} \\ 1 & x_2 & x_2 & \dots & x_2^{m-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n & \dots & x_n^{m-1} \end{bmatrix}, \quad (1.13)$$

де  $x_i$  - значення єдиної пояснюючої змінної в  $i$ -му експерименті.

Зокрема, якщо пояснюча змінна – час, відліки якого беруться з постійним кроком  $T$ , тобто  $x_i = (i-1)T$ , то

$$F = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & T & T^2 & \dots & T^{m-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & (n-1)T & [(n-1)T]^2 & \dots & [(n-1)T]^{m-1} \end{bmatrix}. \quad (1.14)$$

Основним завданням регресійного аналізу являється отримання оцінок параметрів регресії ( $\hat{\beta}_1, \dots, \hat{\beta}_m$ ), які були б оптимальними в певному сенсі. Отримані оцінки, які будемо представляти у вигляді компонентів вектора  $\hat{\beta} = [\hat{\beta}_1, \dots, \hat{\beta}_m]^T$ , дозволяють вирішувати задачу оцінки (відновлення) регресії і її прогнозу. У першому випадку як оцінка «дійсного» значення результуючої змінної, відповідної вектору регресорів  $f_i (i \in [1, n])$  вже використаному при формуванні оцінки вектора параметрів, застосовується величина

$$\hat{y}_i \equiv \hat{\eta}(x_i) = f_i^T \hat{\beta}, \quad (1.15)$$

а як оцінка вектора регресії  $\eta = [\eta(x_1), \dots, \eta(x_n)]^T$  - вектор

$$\hat{y} \equiv \hat{\eta} = F \hat{\beta}. \quad (1.16)$$

У другому випадку для оцінки значення функції регресії при векторі  $x_j$  пояснюючих змінних, не співпадаючим ні з одним із раніше зафіксованих значень  $x_i (i = \overline{1, n})$ , використовується величина

$$\hat{y}_j = \hat{\eta}(x_j) = f_j^T \hat{\beta}. \quad (1.17)$$

## 1.2. Приклади приведення задачі опрацювання експериментальних даних до задачі регресійного аналізу

### 1.2.1. Задача оцінки імпульсної характеристики системи

Приведемо приклади, які показують, що багато завдань оцінки параметрів моделі джерела експериментальних даних можуть бути сформульовані в термінах лінійної регресії.

*Відновлення сигналу на вході лінійної динамічної системи.*

Сигнал  $Y(t)$ , спостережуваний на виході лінійної динамічної системи з відомою імпульсною характеристикою  $h(t)$ , представляється у вигляді суми сигналу  $y_u(t)$  - результату перетворення системою невідомого сигналу  $u(t)$ , і похибки вимірювань  $\varepsilon(t)$ :

$$Y(t) = y_u(t) + \varepsilon(t), \quad (1.18)$$

де  $y_u(t) = \int_{-\infty}^t h(t-\tau)u(\tau)d\tau$ ;  $\varepsilon(t)$  - випадковий процес з

нульовим математичним сподіванням і відомою кореляційною функцією  $R\varepsilon(\tau)$ .

Якщо виконано  $n$  вимірів  $y[i]$  відліків реалізації  $y(t)$  сигналу  $Y(t)$  ( $y[i] \equiv y(t)I_{t=(i-1)T}$ ) і визначена дискретна імпульсна характеристика системи  $h[i] \equiv h(t)I_{t=(i-1)T}$ , то з врахуванням виразу (1.18) може бути отримана наступна система рівнянь [5]:

$$y = F\beta + \varepsilon, \quad (1.19)$$

де  $y = [y[1], \dots, y[n]]^T$ ;

$$F = \begin{bmatrix} Th[1] & 0 & 0 & \dots & 0 \\ Th[2] & Th[1] & 0 & \dots & 0 \\ Th[3] & Th[2] & Th[1] & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ Th[m] & Th[m-1] & Th[m-2] & \dots & 0 \\ 0 & Th[m] & Th[m-1] & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & Th[1] \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & Th[m] \end{bmatrix};$$

$\varepsilon = [[1], \dots, \varepsilon[n]]^T$  - вектор з нульовим математичним сподіванням ( $M\{\varepsilon\} = 0$ ) і відомою кореляційною матрицею

$$D\{\varepsilon\} = [R_\varepsilon((i-j)T)] \quad (i, j = \overline{1, n}); \quad \beta = [u[1], \dots, u[m]]^T.$$

### Ідентифікація лінійної динамічної системи

Існує декілька постановок задачі ідентифікації лінійних систем [5, 8, 24], тобто задачі визначення динамічних характеристик систем по експериментальним даним. Ми розглянемо лише два варіанти, які часто використовуються на практиці [5, 21, 24, 44, 45].

Задача оцінки імпульсної характеристики системи  $h(t)$  по вхідному  $U(t)$  і вихідному  $Y(t)$  сигналам, вимірюваним з похибками  $U(t) = u_0(t) + \delta(t)$ ;  $Y(t) = y_u(t) + \varepsilon(t)$ . При

цьому  $y_u(t)$  - результат перетворення сигналу  $u(t)$  досліджуваною системою. Для вихідного сигналу можна записати

$$Y(t) = \int_0^t u(t-\tau)h(\tau)d\tau + \varepsilon(t). \quad (1.20)$$

Шуми  $\delta(\tau)$  і  $\varepsilon(\tau)$  мають нульове математичне сподівання і відомі кореляційні функції  $R_\delta(\tau)$  і  $R_\varepsilon(\tau)$ .

При дискретному спостереженні реалізацій сигналів  $u(t)$  і  $y(t)$  з кроком дискретизації  $T$  і отриманні наборів відліків  $\{y[i]\}, \{u[i]\}$  ( $i = \overline{1, n}$ ), можна отримати систему лінійних рівнянь, якій повинна задовольняти дискретна імпульсна характеристика  $h[i]$  ( $i = \overline{1, m}$ ):

$$y = F\beta + \varepsilon, \quad (1.21)$$

де

$$y = [y[1], \dots, y[n]]^T;$$

$$F = \begin{bmatrix} Tu[1] & 0 & 0 & \dots & 0 \\ Tu[2] & Tu[1] & 0 & \dots & 0 \\ Tu[3] & Tu[2] & Th[1] & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ Tu[m] & Tu[m-1] & Tu[m-2] & \dots & 0 \\ 0 & Tu[m] & Tu[m-1] & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & Tu[1] \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & Tu[m] \end{bmatrix};$$

$$\varepsilon = [\varepsilon[1], \dots, \varepsilon[n]]^T; \beta = [h[1], \dots, h[m]]^T.$$

### 1.2.2 Задача оцінки коефіцієнтів різницевого рівняння дискретної моделі системи

Задача оцінки коефіцієнтів різницевого рівняння дискретної моделі системи по вхідному і вихідному дискретним сигналам, спостережуваним з похибками (див. вище).

Нехай різничеве рівняння, яке описує систему, має вигляд

$$y[i] + a_{N-1}y[i-1] + \dots + a_0y[i-N] = b_p u[i-N+P] + \dots + b_0u[i-N], \quad (1.22)$$

де  $N$ - порядок рівняння;  $P < N$ ;  $a_i, b_j$  - невідомі коефіцієнти ( $i = \overline{0, N-1}; j = \overline{0, P}$ ).

Це рівняння можна переписати наступним чином:

$$-y[i] = a_0y[i-N] + \dots + a_{N-1}y[i-1] - b_0u[i-N] - \dots - b_p u[i-N+P]. \quad (1.23)$$

Припустимо, що в результаті спостережень отримано  $M$  відліків реалізацій вхідного і вихідного сигналів ( $M \geq 2N$ ). Тоді з врахуванням виразу (1.23) можна записати систему рівнянь, із яких визначаються коефіцієнти рівняння (1.22):

$$y = F\beta + \varepsilon, \quad (1.24)$$

де

$$y = [-y[N+1] \dots - y[M]]^T;$$



$$F = \begin{bmatrix} y[1] & y[2] & \dots & y[N] & -u[1] & \dots & -u[P+1] \\ y[2] & y[3] & \dots & y[N+1] & -u[2] & \dots & -u[P+1] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ y[M-N] & y[M-N+1] & \dots & y[M-1] & -u[M-N] & \dots & -u[M-N+P] \end{bmatrix};$$

$$\varepsilon = -\varepsilon[N+1], \dots, -\varepsilon[M]]^T; \quad \beta = [a_0, \dots, a_{N-1} \quad b_0, \dots, b_p]^T.$$

### Ідентифікація моделі часового ряду (дискретного випадкового процесу)

Для опису реальних випадкових процесів широко використовуються їх моделі у вигляді деяких гіпотетичних лінійних дискретних систем, що формують процеси з аналогічними кореляційними властивостями із некорельованого шуму. Такі моделі називаються формуючими фільтрами [23,49]. Зокрема, як в економетрії, так і в техніці широко застосовуються моделі авто регресії – ковзаючого середнього. Розглянемо авторегресійну модель. Вона визначається різницеvim рівнянням:

$$Y[i] + \sum_{j=1}^m a_j Y[i-j] = N[i], \quad (1.25)$$

де  $Y[i]$  – відліки випадкового процесу;  $a_j$  - коефіцієнти моделі ( $j = \overline{1, m}$ );  $N[i]$  - відліки (неспостерігаємі) некорельованого породжуваного шуму

$$(M\{N[i]\} = 0, M\{N[i]N[j]\} = \sigma^2 \delta_{ij}).$$

Авторегресійна модель може описувати як стаціонарний так і нестаціонарний випадковий процес. Необхідною і достатньою умовою стаціонарності являється розташування коренів  $p_i (i = \overline{1, m})$  характеристичного рівняння  $p^m + a_1 p^{m-1} + \dots + a_m = 0$  всередині круга одиничного радіуса, тобто умови  $|p_i| < 1$ .

Рішення рівняння (1.25) можна записати в рекурентній формі:

$$Y[i] = \sum_{j=1}^m (-a_j) Y[i-j] + N[i] \quad (i \geq m+1), \quad (1.26)$$

Нехай виконуються виміри відліків реалізації випадкового процесу  $Y[i]: y[1], \dots, y[k]$  ( $k \geq 2m$ ). Тоді на основі виразу (1.25) може бути складена система рівнянь, якій повинні задовольняти шукані коефіцієнти моделі:

$$y = F\beta + \varepsilon, \quad (1.27)$$

де

$$y = [y[m+1], \dots, y[k]]^T;$$

$$F = \begin{bmatrix} -y[m] & -y[m-1] & \dots & -y[1] \\ -y[m+1] & -y[m] & \dots & -y[2] \\ \dots & \dots & \dots & \dots \\ -y[k-1] & -y[k-2] & \dots & -y[k-m] \end{bmatrix};$$

$$\beta = [a_1, \dots, a_m]^T; \quad \varepsilon = [n[m+1], \dots, n[k]]^T.$$

### Оцінка параметрів поліноміальної функції, вимірюваною з похибками

Нехай  $f(t)$  – детермінована (корисна) складова випадкового спостережуваного процесу, що є поліномом степені  $m-1$ :

$$f(t) = a_0 + a_1 t + \dots + a_m t^{m-1}. \quad (1.28)$$

Зокрема, така апроксимація часто використовується для представлення незбурених траєкторій руху різних об'єктів [49]. Нехай в результаті експерименту отримано  $n \geq m$  відліків спостерігаємого процесу

$$Y(t_i) : \{y[i]\} \quad (i = \overline{1, n}), \text{ де } y[i] \equiv y[t_i].$$

Тоді, враховуючи, що

$$y(t_i) = f(t_i) + \varepsilon(t_i), \quad (1.29)$$

нескладно записати систему рівнянь для визначення коефіцієнтів поліному (1.28):

$$y = F\beta + \varepsilon, \quad (1.30)$$

де

$$y = [y[1], \dots, y[n]]^T;$$

$$F = \begin{bmatrix} 1 & t_1 & t_1^2 & \dots & t_1^{m-1} \\ 1 & t_2 & t_2^2 & \dots & t_2^{m-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_n & t_n^2 & \dots & t_n^{m-1} \end{bmatrix};$$

$$\varepsilon = [\varepsilon[1], \dots, \varepsilon[n]]^T; \quad \beta = [a_0, \dots, a_m]^T.$$

При цьому  $\varepsilon$  - випадковий вектор з нульовим математичним сподіванням і відомою кореляційною матрицею  $D\{\varepsilon\}$ .

## Розділ 2. Основні положення класичного регресійного аналізу

### 2.1. Базові (фундаментальні) положення класичного лінійного регресійного аналізу

Класичний лінійний регресійний аналіз опирається на систему положень про властивості регресійної моделі, виконання яких гарантує отримання оптимальних оцінок параметрів і функції регресії [1, 3, 15, 27, 78]. Вони стосуються перш за все випадкової змінної  $\varepsilon$ , що враховує вплив неврахованих факторів, випадкових перешкод і помилок вимірювань. Крім того, ці положення пред'являють певні вимоги до матриці регресорів  $F$ . Перелічимо найбільш важливі із них.

1. На вектор невідомих параметрів не накладено ніяких обмежень:  $\beta \in \mathbb{R}^m$ .
2. Вектор  $\varepsilon$  є  $n$ -мірна випадкова величина.
3. Математичне сподівання вектора відхилень рівне нулю  $M\{\varepsilon\} = 0$ .
4. Компоненти вектора  $\varepsilon$  не корельовані між собою і мають однакові дисперсії  $\sigma_\varepsilon^2$ :

$$\text{cov}\{\varepsilon_i, \varepsilon_j\} \equiv M\{\varepsilon_i, \varepsilon_j\} = \sigma_\varepsilon^2 \delta_{ij}$$

або

$$D\{\varepsilon\} \equiv [\text{cov}\{\varepsilon_i, \varepsilon_j\}] \equiv M\{\varepsilon\varepsilon^T\} = \sigma_\varepsilon^2 E,$$

де  $\delta_{ij}$  - символ Кронекера;  $E$  - одинична матриця.

Це положення часто називають умовою однорідності (гомоскедастичності) і некорельованості вимірів. Якщо вона не виконується, виміри неоднорідні (гетероскедастичні) і (або) корельовані.

Із неї слідують співвідношення для математичного сподівання і кореляційної матриці вектора значень залежної змінної:

$$\bar{Y} = F\beta \equiv \eta;$$

$$D\{Y\} \equiv M\{(Y - \bar{Y})(Y - \bar{Y})^T\} = D\{\varepsilon\} = \sigma_\varepsilon^2 E.$$

5. Випадкові величини  $\varepsilon_i (i = \overline{1, n})$  і вектор  $\varepsilon$  в цілому мають гаусів розподіл. При виконанні положення 4

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2); \varepsilon \sim N(N(0, \sigma_\varepsilon^2 E)).$$

6. Матриця регресорів  $F$  детермінована, тобто її елементи  $f_{ij} = f_j(x_j)$  не являються випадковими величинами.

7. Ранг матриці регресорів рівний числу параметрів функції регресії:  $\text{rank } F = m$ .

Класичним регресійним аналізом називають процедуру оцінювання регресійних параметрів і статистичний аналіз моделі при виконанні цих положень. Його основні положення розглядаються нижче. В подальшому перелічені положення будуть послаблюватися, так як

їх порушення являються скоріше правилом, ніж виключенням.

## 2.2. Класичні оцінки параметрів регресії методом найменших квадратів і їх властивості

Розглянемо задачу оцінювання параметрів регресії  $\beta_i (i = \overline{1, m})$  за результатами експериментів  $y_i (i = \overline{1, n})$ . Для оцінювання можуть використовуватися різні методи. В класичному регресійному аналізі використовується метод найменших квадратів (МНК). В цьому випадку оцінки мінімізують суму квадратів нев'язок оцінювання. Нев'язками оцінювання, які також називають остатками, являються різниці

$$\varepsilon_i = y_i - \hat{y}_i,$$

а вектором нев'язок – вектор

$$\varepsilon = y - \hat{y}.$$

Остатки визиваються двома причинами – відмінністю вектора оцінок  $\hat{\beta}$  від вектора дійсних параметрів  $\beta$  і наявністю випадкових збурень  $\varepsilon$ :

$$\varepsilon = F(\beta - \hat{\beta}) + \varepsilon. \quad (2.31)$$

Сума квадратів остатків

$$Q_0 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \equiv \sum_{i=1}^n \varepsilon_i^2 \equiv (y - \hat{y})^T (y - \hat{y}) \equiv \varepsilon^T \varepsilon \quad (2.32)$$

виступає в МНК як критерій якості оцінок. Оцінкою методу найменших квадратів (МНК- оцінкою) називають вектор  $\hat{\beta}$ , що мінімізує функціонал (2.32).

Для знаходження оцінки перепишемо (2.32) у вигляді

$$Q_0 = Q(\hat{\beta}) = (y - F\hat{\beta})^T (y - F\hat{\beta}),$$

продиференціюємо його по  $\hat{\beta}$  і прирівнюємо нулю. В результаті отримуємо рівняння для обчислення оцінки:

$$\partial Q_0 / \partial \hat{\beta} = -2F^T y + 2F^T F\hat{\beta} = 0.$$

Його очевидним чином можна перетворити до вигляду

$$F^T F\hat{\beta} = F^T y. \quad (2.33)$$

Ця система лінійних алгебраїчних рівнянь (СЛАУ) являється основною для отримання МНК-оцінок параметрів регресії і називається системою нормальних рівнянь.

Так як матриця  $F$  має повний ранг (положення 7), то матриця  $(F^T F)$  не вироджена і, отже, має обернену матрицю  $(F^T F)^{-1}$ . Перемноживши праву і ліву частини системи (1.33) на  $(F^T F)^{-1}$ , отримуємо явні вирази для МНК-оцінки вектора параметрів;

$$\hat{\beta} = (F^T F)^{-1} F^T y, \quad (2.34)$$

звідки виходить вираз для оцінки вектора регресії:

$$\hat{\eta} = F\hat{\beta} = F(F^T F)^{-1} F^T y. \quad (2.35)$$

Матриця  $G = F^T F$  називається інформаційною, а її обернена  $S = (F^T F)^{-1}$  матрицею похибок.

Розглянемо основні властивості МНК-оцінок параметрів і вектора регресії. Почнемо із властивостей, які не залежать від виду розподілу збурень.

### 1. МНК-оцінки відносяться до класу лінійних.

Дійсно, оцінки  $\hat{\beta}$  і  $\hat{\eta}$  виходять із вектора спостережень  $y$  за допомогою лінійного перетворення - множення на матриці  $A = (F^T F)^{-1} F^T$  і  $B = FA$  відповідно:

$$\hat{\beta} = Ay; \quad \eta = By. \quad (2.36)$$

### 2. МНК-оцінки являються незміщеними:

$$M\{\hat{\beta}\} = M\{(F^T F)^{-1} F^T y\} = \beta; \quad (2.37)$$

$$M\{\hat{\eta}\} = M\{F\hat{\beta}\} = \eta. \quad (2.38)$$

### 3. Кореляційні матриці оцінок визначаються слідуючими виразами:

$$\begin{aligned} D\{\hat{\beta}\} &= M\{(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T\} = \\ &= M\{[(F^T F)^{-1} F^T (y - \eta)][(F^T F)^{-1} F^T (y - \eta)]^T\} = \\ &= M\{(F^T F)^{-1} F^T (y - \eta)(y - \eta)^T F[(F^T F)^{-1}]^T\} = \\ &= (F^T F)^{-1} F^T M\{(y - \eta)(y - \eta)^T\} F (F^T F)^{-1} = \\ &= (F^T F)^{-1} F^T \sigma_\varepsilon^2 E F (F^T F)^{-1} = \sigma_\varepsilon^2 (F^T F)^{-1} = \sigma_\varepsilon^2 S; \end{aligned} \quad (2.39)$$

$$\begin{aligned}
D\{\hat{\eta}\} &= M\{(\hat{\eta} - \eta)(\hat{\eta} - \eta)^T\} = M\{F(\hat{\beta} - \beta)(\hat{\beta} - \beta - \\
&- \beta)^T F^T\} = FM\{(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T\}F^T = FD\{\hat{\beta}\}F^T = \\
&= \sigma_\varepsilon^2 F(F^T F)^{-1} F^T = \sigma_\varepsilon^2 F S F^T. \quad (2.40)
\end{aligned}$$

При виводі враховано, що  $[(F^T F)^{-1}]^T = (F^T F)^{-1}$  (зважаючи на симетричність матриці),  
 $M\{(y - \eta)(y - \eta)^T\} = \sigma_\varepsilon^2 E$  (див. положення 4).

Елементами матриць  $D\{\hat{\beta}\}$  і  $D\{\hat{\eta}\}$  є дисперсії  $\sigma^2\{\hat{\beta}_i\}$  і  $\sigma^2\{\hat{\eta}_i\}$  і взаємно кореляційні моменти між оцінками параметрів  $(R\{\hat{\beta}_i, \hat{\beta}_j\})$  або оцінками відліків функції регресії  $(R\{\hat{\eta}_i, \hat{\eta}_j\})$ .

З виразів (2.39) і (2.40) виходить, що оцінки параметрів регресії і відліків функції регресії будуть некорельованими лише у тому випадку, якщо стовпчики матриці  $\mathbf{F}$  попарно ортогональні і матриця  $\mathbf{G}$  є діагональною.

Показано, що МНК-оцінки є ефективними в класі незміщених лінійних оцінок (теорема Гауса-Маркова) [3, 15, 27, 67]. Якщо збурення (відхилення) мають розподіл Гауса (положення 5), то МНК-оцінки є ефективними і в класі всіх незміщених оцінок (лінійних і нелінійних) і співпадають з оцінками методу максимальної правдоподібності.

Ефективність МНК-оцінок означає, що

$$\begin{aligned}
D\{\hat{\beta}\} &\leq D\{\tilde{\beta}\}; \quad \sigma^2\{\hat{\beta}_i\} \leq \sigma^2\{\tilde{\beta}_i\} \quad (i = \overline{1, m}); \\
tr\{D\{\hat{\beta}\}\} &\leq tr\{D\{\tilde{\beta}\}\}; \quad \det D\{\hat{\beta}\} \leq \det D\{\tilde{\beta}\}. \quad (2.41)
\end{aligned}$$

Аналогічно для оцінок регресії

$$\begin{aligned}
D\{\hat{\eta}\} &\leq D\{\tilde{\eta}\}; \quad \sigma^2\{\hat{\eta}_i\} \leq \sigma^2\{\tilde{\eta}_i\} \quad (i = \overline{1, n}); \\
tr\{D\{\hat{\eta}\}\} &\leq tr\{D\{\tilde{\eta}\}\}; \quad \det D\{\hat{\eta}\} \leq \det D\{\tilde{\eta}\}. \quad (2.42)
\end{aligned}$$

При цьому  $\tilde{\beta}, \tilde{\beta}_i, \tilde{\eta}$  і  $\tilde{\eta}_i$  - довільні оцінки  $\beta, \beta_i, \eta$  і  $\eta_i$  вказаних класів.

Відмітимо, що величини  $\det D\{\hat{\beta}\}$  і  $\det D\{\hat{\eta}\}$  називаються узагальненими дисперсіями оцінок векторів параметрів і регресії.

#### 4. МНК-оцінки спроможні (обґрунтовані):

$$\lim_{n \rightarrow \infty} P\{|\hat{\beta}_i - \beta_i| < \gamma\} = 1 \quad (i = \overline{1, m}),$$

де  $\gamma$  – довільне позитивне число.

**5. МНК-оцінки параметрів регресії дають можливість отримати незміщені ефективні і спроможні (обґрунтовані) оцінки вектора збурень (відхилень), дисперсії і кореляційної матриці збурень.**

Ними є вектор остатків

$$\hat{\varepsilon} = e \equiv (y - F\hat{\beta}), \quad (2.43)$$

величина

$$\hat{\sigma}_\varepsilon^2 \equiv s^2 = e^T e / (n - m) \quad (2.44)$$

і матриця

$$\hat{D}\{\varepsilon\} = s^2 S = (y - F\hat{\beta})^T (y - F\hat{\beta}) (F^T F)^{-1} / (n - m). \quad (2.45)$$

Тепер розглянемо властивості, зв'язані з припущенням про гаусів розподіл збурень.

**6. При виконанні положення 5 вектори оцінок  $\hat{\beta}$  і  $\hat{\eta}$  мають багатовимірні гаусові розподіли:**

$$\hat{\beta} \sim N(\beta; \sigma_\varepsilon^2 (F^T F)^{-1}); \hat{\eta} \sim N(\eta, \sigma_\varepsilon^2 F (F^T F)^{-1} F^T). \quad (2.46)$$

### 7. Випадкова величина

$$g = Q_0 / \sigma_\varepsilon^2, \quad (2.47)$$

де  $Q_0 = e^T e$  - так звана остаточна сума квадратів, має центральний  $\chi^2$  - розподіл з  $\nu = n - m$  степенями свободи.

### 2.3. Рекурентний алгоритм методу найменших квадратів

Традиційний алгоритм МНК-оцінок параметрів регресії по виборці даних повного об'єму (2.34) не завжди може бути використаний на практиці. До його недоліків відносяться велика необхідна ємкість пам'яті для збереження вектора  $y$  і матриці регресорів  $F$ , необхідність заново рішати задачу оцінювання при надходженні нових даних, неможливість отримання текучих оцінок в інтересах оперативного управління об'єктом в ході експерименту, необхідність у великому числі обчислювальних операцій при високій розмірності вектора параметрів обернення інформаційної матриці.

Розглянемо вільний від цих недоліків рекурентний алгоритм послідовного оцінювання вектора параметрів по мірі надходження нових вимірів.

Нехай спочатку проведено  $i$  вимірів. Це дає можливість сформуувати вектор спостережень і матрицю регресорів, що відповідають цьому виміру:

$$y^{(i)} = [y_1, \dots, y_i]^T; \quad F_i = \begin{bmatrix} f_1^T \\ \dots \\ f_i^T \end{bmatrix}. \quad (2.49)$$

Сформуємо оцінку вектора параметрів регресії, відповідну  $i$  вимірюванням:

$$\hat{\beta}^{(i)} = (F_i^T F_i)^{-1} F_i^T y^{(i)},$$

де  $S_i = (F_i^T F_i)^{-1}$  - матриця похибок на  $i$ - му кроці.

Після проведення  $(i+1)$ - го виміру можна сформуувати вектор результуючих змінних  $y^{(i+1)}$ , матрицю регресорів  $F_{i+1}$  і отримати оцінку

$$\hat{\beta}^{(i+1)} = (F_{i+1}^T F_{i+1})^{-1} F_{i+1}^T y^{(i+1)} \equiv S_{i+1} F_{i+1}^T y^{(i+1)}. \quad (2.50)$$

Цілком очевидно, що

$$y^{(i+1)} = \begin{bmatrix} y^{(i)} \\ y_{i+1} \end{bmatrix}; \quad F_{i+1} = \begin{bmatrix} F_i \\ f_{i+1}^T \end{bmatrix}; \quad F_{i+1}^T = [F_i^T \ f_{i+1}]. \quad (2.51)$$

Ці співвідношення дають можливість переписати вираз для оцінки (2.50) слідуочим чином:

$$\hat{\beta}^{(i+1)} = (F_i^T F_i + f_{i+1} f_{i+1}^T)^{-1} (F_i^T y^{(i)} + f_{i+1} y_{i+1}). \quad (2.52)$$

Тепер скористаємося слідкуючою формулою (часто називають лемою про обернення матриці)

$$(A + \gamma aa^T)^{-1} = A^{-1} - \frac{A^{-1}aa^T A^{-1}}{a^T A^{-1}a + \gamma^{-1}}. \quad (2.53)$$

В її справедливості можна впевнитись за допомогою безпосередньої перевірки виконання рівності

$$(A + \gamma aa^T) \left( A^{-1} - \frac{A^{-1}aa^T A^{-1}}{a^T A^{-1}a + \gamma^{-1}} \right) = E.$$

З врахуванням виразу (2.53) можна отримати рекурентний вираз для матриці похибок і оцінок параметрів:

$$S_{i+1} = S_i - \frac{S_i f_{i+1} f_{i+1}^T S_i}{f_{i+1}^T S_i f_{i+1} + 1}, \quad (2.54)$$

$$\begin{aligned} \hat{\beta}^{(i+1)} &= [S_i F_i^T y^{(i)} + S_i f_{i+1} y_{i+1} - S_i f_{i+1} f_{i+1}^T S_i * \\ &* (F_i^T y^{(i)} + f_{i+1} y_{i+1})] (f_{i+1}^T S_i f_{i+1} + 1)^{-1} = \hat{\beta}^{(i)} + \\ &c_{i+1} [y_{i+1} - f_{i+1}^T \hat{\beta}^{(i)}], \end{aligned} \quad (2.55)$$

де

$$c_{i+1} = S_i f_{i+1} / (f_{i+1}^T S_i f_{i+1} + 1) \quad - \text{вектор корекції.}$$

Із виразу (2.55) виходить, що оцінка вектора параметрів на  $(i+1)$ -му кроці представляє собою суму оцінки вектора на попередньому,  $i$ -му, кроці і поправки, яка, в свою чергу, формується із нев'язки (різниці між  $(i+1)$ -м виміром  $y_{i+1}$  і його прогнозом  $f_{i+1}^T \hat{\beta}^{(i)}$  по оцінці  $\hat{\beta}^{(i)}$  за допомогою вектора корекції.

Для функціонування алгоритму необхідно використати початкові умови для вектора оцінок  $\hat{\beta}^{(0)}$  і матриці похибок

$S_0$ . При цьому можливі різні варіанти.

По-перше, як початкові умови можна використати матрицю  $S_j$  і вектор  $\hat{\beta}^{(j)}$  ( $j \geq m$ ), сформовані за допомогою не рекурентного алгоритму. В цьому випадку рекурентний алгоритм починає працювати з  $(j+1)$ -го кроку. Недоліками такого вибору початкових умов є необхідність виконання (однократного) обернення інформаційної матриці і вимоги запам'ятовування вектора даних і матриці регресорів на початковому етапі. Перевагою є той факт, що на будь-якому кроці рішення регресійної задачі рекурентна оцінка співпадає з традиційною оцінкою (2.34) і володіє всіма розглянутими властивостями.

По-друге, для того щоб рекурентний алгоритм міг функціонувати, починаючи з першого кроку, можна задати довільні значення початкових умов. Звичайно вважають  $\hat{\beta}^{(0)} = 0$ ;  $S_0 = \alpha E$ , де  $\alpha$  - достатньо мале число ( $\alpha = 10^{-6} \dots 10^{-4}$ ). Як правило, при зростанні і рекурентні значення  $\hat{\beta}^{(i+1)}$  і  $S_{i+1}$  сходяться до значень, відповідним традиційному МНК.

Відмітимо, що рекурентний алгоритм дозволяє в ході експерименту рішення задачі оптимального останову вимірів при досягненні заданої точності оцінок. Наприклад, якщо задані рівні дисперсій оцінок параметрів  $D\{\hat{\beta}_i\}_{\text{тр}}$ , то обчислювальний процес закінчується на  $j$ -му кроці, якщо  $\sigma_\varepsilon^2 S_{ij} \leq D \hat{\beta}_{i \text{ тр}}$  ( $i = \overline{1, m}$ ), де  $S_{ij}$  -  $j$ -й діагональний елемент матриці  $S_j$  на  $j$ -му кроці.

## 2.4. Статистичний аналіз якості регресійної моделі

Отримані МНК-оцінки коефіцієнтів регресії (2.34) забезпечують високу якість отриманої моделі лише при умові, що її структура  $F\beta$  відповідає структурі дійсної залежності  $\eta_0 = F_0\beta_0$  між математичним сподіванням відгуку і факторами. Але на практиці, як правило, відсутня апріорна інформація про структуру дійсної моделі і дослідник вимушений по черзі розглядати різні види регресійних моделей і зупинитись на тій, яка узгоджується з експериментальними даними. Таку модель називають адекватною [3, 15]. Вона повинна задовольняти умові

$$M\{y\} = F\beta. \quad (2.56)$$

Проблеми раціонального вибору структури регресійної моделі розглядаються у розділі 5.

Адекватна модель не обов'язково повинна співпадати з дійсною. Більше того, адекватна модель не єдина – за допомогою довільного неособливого лінійного перетворення від моделі  $F\beta$  можна перейти до другої адекватної моделі  $F^*\beta^*$ :  $F^* = FR$ ,  $\beta^* = R^{-1}\beta$ ,  $R$  – невідроджена матриця. Однак загальним для всіх адекватних моделей є те, що для кожної із них існує неособливе лінійне перетворення  $R$ ,  $F$ , яке приводить її до дійсної моделі:

$$R_F : F^* R_F = F_0; \quad R_F^{-1} \beta^* = \beta_0. \quad (2.57)$$

Похибки у виборі структури моделі часто проявляються у тому, що, по-перше, оцінювана модель містить більше параметрів, ніж дійсна (так званий перебір параметрів), по-друге, модель, що перевіряється, вміщує менше параметрів,

ніж дійсна (недобір параметрів). Розглянемо наслідки цих похибок.

Почнемо з недобору параметрів. Нехай прийнята функція регресії має вигляд

$$\eta = F\beta, \quad (2.58)$$

в той же час як дійсною функцією регресії є функція

$$\eta_0 = F_0\beta_0. \quad (2.59)$$

Будемо вважати, що вектор дійсних параметрів  $\beta_0$  включає в себе вектор оцінюваних параметрів:

$$\beta_0 = [\beta^T \beta_1^T]^T, \quad (2.60)$$

де  $\beta_1$  – вектор параметрів, які не входять у функцію регресії (2.58).

В даному випадку дійсну матрицю регресорів можна представити у вигляді:

$$F_0 = [FF_1], \quad (2.61)$$

де  $F_1$  – матриця, що відображає частку в результуючому значенні  $y$  членів, відповідаючи вектору  $\beta_1$ .

Із виразів (2.60), (2.61) виходить, що вектор результуючих змінних

$$y = F\beta + F_1\beta_1 + \varepsilon.$$



Оцінка вектора параметрів вибраної моделі (2.58) має вигляд  $\hat{\beta} = (F^T F)^{-1} F^T y$ . Фактично це значить, що використовується оцінка вектора  $\hat{\beta}_0 = [\beta^T 0^T]^T$ .

Розглянемо математичне сподівання оцінки  $\beta$ :

$$M\{\hat{\beta}\} = F^T F)^{-1} F^T M\{y\} = (F^T F)^{-1} F^T (F\beta + F_1\beta_1) = \beta + (F^T F)^{-1} F^T F_1\beta_1 \equiv \beta + \Delta\beta(F, F_1, \beta_1). \quad (2.62)$$

Таким чином, видно, що при недоборі параметрів регресії отримана оцінка  $\hat{\beta}$  в загальному випадку є зміщеною, тобто, навіть компоненти вектора параметрів, які оцінюються, визначаються із систематичною похибкою  $\Delta\beta$ . Відповідно зміщеними виявляються і оцінки вектора регресії  $\hat{\eta}$  і дисперсії збурень  $S^2$ . Зміщення буде відсутнє лише в тому випадку, коли стовбці матриці  $F_0$  ортогональні: тоді матриці  $F$  і  $F_1$  будуть ортогональні ( $F^T F_1 = 0$ ) і  $\Delta\beta = 0$ .

Можна показати [27], що така оцінка не являється спроможною.

Тепер розглянемо перебір параметрів. Нехай  $\eta_0 = F_0\beta_0$  - дійсна функція регресії, а передбачувана модель має вигляд

$$\eta = F\beta,$$

де

$$F = [F_0 F_1]; \quad (2.63)$$

$$\beta = [\beta_0^T \beta_1^T]^T. \quad (2.64)$$

При цьому  $\beta_1$  - вектор «лишніх» параметрів.

В цих умовах оцінка вектора параметрів  $\hat{\beta}$  має вигляд

$$\hat{\beta} = (F^T F)^{-1} F^T y.$$

Найдемо математичне сподівання цієї оцінки з врахуванням істинної функції регресії:

$$M\{\hat{\beta}\} = \left[ \begin{array}{c} F_0^T \\ F_1^T \end{array} \right] [F_0 F_1]^{-1} \left[ \begin{array}{c} F_0^T \\ F_1^T \end{array} \right] F_0 \beta_0 = \left[ \begin{array}{cc} F_0^T F_0 & F_0^T F_1 \\ F_1^T F_0 & F_1^T F_1 \end{array} \right]^{-1} \left[ \begin{array}{c} F_0^T F_0 \beta_0 \\ F_1^T F_0 \beta_0 \end{array} \right] = \left[ \begin{array}{c} \beta_0 \\ 0 \end{array} \right]. \quad (2.65)$$

Із цього виразу видно, що істинні параметри оцінюються незміщено, а математичне сподівання оцінок «лишніх» (надлишкових) параметрів рівне нулю.

Звідси виходить, що оцінки вектора регресії і дисперсії, отримані в умовах перебору параметрів, також являються незміщеними. Доказано, що МНК-оцінки в умовах перебору являються спроможними (обґрунтованими). В той же час точність оцінок при переборі параметрів втрачається [15, 27].

$$D\{\hat{\beta}\} = D\{\hat{\beta}_0\} + \sigma_e^2 (F^T F)^{-1} F_0^T F_1 F_1^T F_0 (F_0^T F_0)^{-1}. \quad (2.66)$$

В даному випадку  $D\{\hat{\beta}_0\}$  - кореляційна матриця МНК-оцінки  $\hat{\beta}_0 = (F_0^T F_0)^{-1} F_0^T y$ , отриманої при умові співпадання істинної і вибраної функції регресії. Друга складова в (2.66) є невід'ємною визначеною матрицею, звідси

$$D\{\hat{\beta}\} \geq D\{\hat{\beta}_0\}. \quad (2.67)$$

Знак рівності в (2.67) забезпечується при умові ортогональної матриці регресорів.

Таким чином, відмітимо наступне.. Недобір параметрів є більш серйозним недоліком регресійної моделі, тому що при цьому оцінки параметрів, регресії і дисперсії збурень стають зміщеними і не являються спроможними. При надмірному ускладненні моделі (переборі параметрів) знижується ефективність оцінювання.

При перевірці адекватності моделі традиційно послідовно досліджуються два аспекти проблеми:

відповідність вибраного класу функцій регресії істинної функції регресії (ця під задача часто також називається перевіркою адекватності);

гіпотези про значимість коефіцієнтів регресії.

Для перевірки першої гіпотези використовуються методи дисперсійного аналізу [1, 3, 15] Розглянемо три можливих варіанти його реалізації.

Нульова гіпотеза  $H_0$  полягає у виконанні рівності

$$H\beta = M\{y\}.$$

Альтернативна гіпотеза  $H_1 : F\beta \neq M\{y\}$ . Гіпотеза  $H_0$

перевіряється при заданому рівні значимості  $\alpha$ .

Обчислюється оцінка дисперсії збурень  $s^2$ :

$$s^2 = Q_{0/(n-m)}, \quad (2.68)$$

де  $Q_0 = (y - F\hat{\beta})^T (y - F\hat{\beta}) = y^T y - \hat{\beta}^T F^T y$  - залишкова сума квадратів, яка характеризує розкид експериментальних даних відносно оціненої функції регресії.

В першому варіанті перевірки гіпотези застосовується апіорне (точне) значення дисперсії  $\sigma_\varepsilon^2$ . Статистика

$$g = s^2 / \sigma_\varepsilon^2 \quad (2.69)$$

при справедливості нульової гіпотези має  $\chi^2$  - розподіл з  $\nu = n - m$  степенями свободи. Тому гіпотеза  $H_0$  приймається, якщо

$$\chi_{\nu, \alpha/2}^2 \leq g \leq \chi_{\nu, 1-\alpha/2}^2, \quad (2.70)$$

В протилежному випадку вона відхиляється.

В другому варіанті перевірки гіпотези вважається, що на основі  $N$  додаткових експериментів  $\{y'_1, \dots, y'_N\}$  при фіксованих значеннях факторів визначена незалежна

оцінка дисперсії  $s_1^2 = [\sum_{i=1}^N (y'_i - (\sum_{j=1}^N y'_j) / N)] / (N - 1)$ . При

справедливості нульової гіпотези статистика

$$g = s^2 / s_1^2 \quad (2.71)$$

Має F-розподіл з  $\nu_1 = n - m$  і  $\nu_2 = N - 1$  степенями свободи. Тому  $H_0$  приймається, якщо

$$g \leq F_{\nu_1, \nu_2, 1-\alpha/2}, \quad (2.72)$$

в протилежному випадку вона відхиляється. Відмітимо, що при цьому приймається  $s^2 \geq s_1^2$ . В протилежному випадку відношення змінюється на протилежне, відповідно змінюється квантиль F-розподілу.

В третьому варіанті перевірки гіпотези обчислюється дисперсія випадкової величини  $Y$  за результатами основного регресійного експерименту:

$$\hat{\sigma}_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \hat{Y})^2 = \frac{1}{n-1} y^T y - \hat{y}^T \hat{y}, \quad (2.73)$$

де  $\hat{y} = [\hat{Y} \dots \hat{Y}]^T$ ;  $\bar{Y} = \left( \sum_{i=1}^n y_i \right) / n$ .

При справедливості нульової гіпотези відношення

$$g = s^2 / \hat{\sigma}_Y^2 \quad (2.74)$$

має F-розподіл з  $\nu_1 = n - m$  і  $\nu_2 = n - 1$  степенями свободи. Тому вона приймається, якщо

$$g \leq F_{\nu_1, \nu_2, 1-\alpha/2}. \quad (2.75)$$

При цьому знову приймається, що

Перевірка гіпотези про значимість параметрів регресії (гіпотези про рівність нулю відповідного параметра) дає можливість прибрати з моделі «лишні» параметри, що, як відмічалось, приводить до підвищення точності оцінок решти. Дані методи дозволяють перевірити, чи не відрізняються отримані оцінки коефіцієнтів від нуля тільки із-за випадкових збурень. Перевірка ґрунтується на тому, що, як відмічалось, оцінки  $\hat{\beta}_i$  мають гаусів розподіл з математичним сподіванням  $\beta_i$  і дисперсією  $D\{\beta_i\} = \sigma_\varepsilon^2 s_{ii}$ , де  $s_{ij}$  – елементи матриці похибок  $S$ . Оцінкою дисперсії  $D\{\beta_i\}$  являється величина  $s^2 s_{ii}$ .

Нульова гіпотеза  $H_0 : \beta_i = 0$ . Альтернативна гіпотеза  $H_1 : \beta_i \neq 0$ . Перевірка виконується при заданому рівні вірогідності помилкового признання значимим коефіцієнта, в дійсності рівного нулю. Застосовується статистика

$$g = \hat{\beta}_i / \sqrt{s^2 s_{ii}}, \quad (2.76)$$

яка при справедливості гіпотези  $H_0$  має t-розподіл Стьюдента з  $\nu = n - m$  степенями свободи.

Гіпотеза  $H_0$  приймається, якщо

$$|g| \leq t_{\nu, 1-\alpha/2}, \quad (2.77)$$

в протилежному випадку коефіцієнт  $\beta_i$  вважається значимим.

Незначимі коефіцієнти доцільно відкинути і заново рішення регресійну задачу. Але після цього необхідно знову перевірити гіпотезу про адекватність нової моделі. Якщо вона не підтвердила адекватності, то необхідно повернутись до попередньої моделі.

Поряд із розглянутими методами дисперсійного аналізу регресійної моделі для перевірки її адекватності широко використовуються і інші методи: критерій Бокса і Веца, перевірка значущості множинного коефіцієнта множинної кореляції, критерій Вальда-Вольфовиця, рангові критерії, графічні дослідження залишків [1, 3, 15, 74].

Для практичного використання, отриманих по експериментальним даним регресійних моделей, велике значення мають показники точності і надійності оцінок параметрів, функції регресії і дисперсії збурень.

Відому інформацію про точність оцінок несуть визначені раніше кореляційні матриці похибок оцінювання (2.39) і (2.40). Але більш зручним є використання довірчих інтервалів, які включають в себе істинні значення досліджуваних величин із заданою довірчою вірогідністю  $p$ . Довірча оцінка ґрунтується на тих же припущеннях, що і перевірка гіпотез про значення параметрів. Тому без додаткового обговорення приведемо вирази для індивідуальних довірчих інтервалів для параметрів  $\beta_i (i = \overline{1, m})$ , відліків функції регресії  $\eta_i$ , вектора регресії  $\eta$ , і дисперсії збурень [1, 3, 15]:

а) дисперсія збурень заздалегідь відома:

$$\hat{\beta}_i - t_{n-m, (1-p)/2} \sqrt{\sigma_\varepsilon^2 s_{ii}} \leq \beta_i \leq \hat{\beta}_i + t_{n-m, p/2} \sqrt{\sigma_\varepsilon^2 s_{ii}};$$

$$\hat{\eta}_i - t_{n-m, (1-p)/2} \sqrt{\sigma_\varepsilon^2 f_i^T (F^T F)^{-1} f_i} \leq \eta_i \leq \hat{\eta}_i +$$

$$+ t_{n-1; p/2} \sqrt{\sigma_\varepsilon^2 f_i^T (F^T F)^{-1} f_i}; \quad (2.78)$$

$$\hat{\eta} - t_{n-m, (1-p)/2} \left[ \begin{array}{c} \sqrt{\sigma_\varepsilon^2 f_1^T (F^T F)^{-1} f_1} \\ \dots\dots\dots \\ \sigma_\varepsilon^2 f_n^T (F^T F)^{-1} f_n \end{array} \right] \leq \eta \leq \hat{\eta} +$$

$$t_{n-m, p/2} \left[ \begin{array}{c} \sqrt{\sigma_\varepsilon^2 f_1^T (F^T F)^{-1} f_1} \\ \dots\dots\dots \\ \sigma_\varepsilon^2 f_n^T (F^T F)^{-1} f_n \end{array} \right];$$

б) дисперсія збурень оцінюється в ході регресійного експерименту. В даному випадку можуть бути використаними вирази, аналогічні (2.78), в яких дисперсія замінюється її оцінкою (2.44), при цьому число степенів свободи t-розподілу зменшується на одиницю.

Розглянуті індивідуальні довірчі інтервали використовуються, як правило тоді, коли матриця  $G = (F^T F)$  близька до діагональної. В других випадках через корельованість оцінок параметрів індивідуальні довірчі інтервали малоінформативні. Тоді застосовується довірча область в m-мірному просторі, яка із заданою вірогідністю p включає в себе вектор  $\beta$ .

Довірча область задається нерівністю [13, 15,71].

$$((\hat{\beta} - \beta)^T F^T F (\hat{\beta} - \beta) \leq ms^2 F_{m, n-m, p}, \quad (2.79)$$

де  $F_{v_1, v_2, p}$  – квантиль p-го порядку F-розподілу Снедекора-Фішера з  $v_1, v_2$  степенями свободи. При цьому вважається, що оцінка дисперсії  $s^2$  отримана із (2.44). Якщо  $s^2$  отрима-

на за додатковими даними, то застосовується квантиль розподілу з m, n-1 степенями свободи.

### Розділ 3. Обчислювальні алгоритми методу найменших квадратів

#### 3.1. Загальна характеристика чисельних методів

Вище було показано, що класичні оцінки параметрів лінійної регресії визначаються рішенням нормальної системи рівнянь (2.33)

$$F^T F \hat{\beta} = F^T y.$$

Яке символічно можна представити у вигляді (2.34)

$$\hat{\beta} = (F^T F)^{-1} F^T y.$$

В тих випадках, коли необхідно тільки отримати оцінки параметрів без їх подальшого статистичного аналізу точності і надійності, тобто тоді, коли немає явної необхідності в обчисленні матриці похибок  $S = (F^T F)^{-1}$ , оцінки визначаються безпосереднім рішенням нормальної системи рівнянь. Це приводить до суттєвої економії числа необхідних обчислювальних операцій і, отже, підвищує оперативність рішення задачі. Якщо крім отримання оцінок необхідно провести їх статистичний аналіз, то без обчислення матриці S не вдається, і тоді МНК-оцінки можуть бути отримані з використанням виразу (2.34). Але і в цьому випадку матриця  $G = (F^T F)$  вкрай рідко обертається безпосередньо по відомим формулам Крамера (тільки при  $m \leq 2 \dots 3$ ). При достатньо високій розмірності

вектора оцінюваних параметрів матриця  $S$  визначається шляхом  $m$ -кратного рішення системи рівнянь:

$$F^T F x_i = e_i \quad (i = \overline{1, m}), \quad (3.80)$$

де  $e_i = [e_{i1}, \dots, e_{im}]^T$ , а  $e_j = 1$  при  $j = i$  і  $e_j = 0$  при  $j \neq i$ .

Після знаходження набору рішень  $\hat{x}_i (i = \overline{1, m})$  обернена матриця отримується їх об'єднанням:

$$S \equiv [s_1, \dots, s_m] = [\hat{x}_1, \dots, \hat{x}_m], \quad (3.81)$$

тобто рішення  $\hat{x}_i$  є стовпчиками  $s_i$  матриці  $S$ .

Таким чином, основу класичного регресійного аналізу складають алгоритми рішення систем лінійних алгебраїчних рівнянь (СЛАУ) виду

$$Ax = b, \quad (3.82)$$

в яких матриця системи  $A$ , вектор невідомих  $x$  і вектор вільних членів  $b$  мають свій конкретний сенс, впливаючий із вирішуваної задачі. Характерним для системи (3.82) є її сумісність і визначеність, симетричність матриці

$A(A = F^T F)$  і її невиродженість (див. розд. 2). Розглянемо основні методи чисельного рішення подібних систем.

При розробці і аналізі різних методів чисельного рішення СЛАУ важливе місце займає врахування виникаючих похибок рішення, що визиваються похибками заокруглення при виконанні основних арифметичних операцій в ЕОМ, неточністю вихідної інформації про систему, і зв'язані з цим проблеми економії машинної пам'яті і зниження необхідного числа операцій [7].

Вибір методу рішення, а також його конкретна реалізація у вигляді обчислювальної програми самим суттєвим чином залежить від типу матриці системи і відповідного способу її зберігання в запам'ятовуючому

пристрої ЕОМ. При цьому можна виділити два основних типи матриць: матриці загального виду і розріджені матриці.

Матриці загального виду містять елементи, більша частина яких відмінна від нуля, тому для їх зберігання необхідно  $m^2$  чарунків пам'яті. Виключення складають так звані обчислювальні матриці, елементи яких можуть бути достатньо просто обчислені по деяким вихідним даним, займаю чим в пам'яті порівняно мало місця. В цьому випадку замість того, щоб запам'ятовувати всі елементи матриці, їх вигідніше кожний раз заново обчислювати. У зв'язку з різноманітністю і специфікою обчислюваних матриць, свої особливості мають і відповідні методи вирішення СЛАУ. Тому основані на них програми розробляються для конкретних задач і не входять в програмне забезпечення універсальних ЕОМ.

Значна, а іноді і більша частина елементів розріджених матриць рівна нулю. Можна виділити декілька різновидностей таких матриць. Одну групу складають розріджені матриці, нульові елементи яких розташовані без будь-якої системи. Тому необхідно запам'ятати не тільки значення ненульових елементів, але і їх «адрес» - номери строчки і стовпця. Однак, особливий інтерес представляють розріджені матриці, що відносяться до другої групи. Це розріджені матриці спеціальної структури - діагональні, трикутні, майже трикутні (матриці Гессенберга), стрічкові матриці, зокрема трьох діагональні, і деякі інші. До цієї ж групи можна віднести і симетричні матриці, які формально не відносяться до розріджених, але потребують для запам'ятовування не  $m^2$  (як в загальному випадку), а  $m(m+1)/2$  чарунків пам'яті.

Для основних видів матриць спеціальної структури розроблені особливі модифікації більшості алгоритмів, спочатку створених для рішення СЛАУ з матрицями

загального виду , а також оригінальні спеціальні алгоритми, які дають можливість ефективно використовувати матриці. Так, існують алгоритми і програми для рішення СЛАУ з симетричними, стрічковими, трьохдіагональними, майже трикутними і другими спеціальними матрицями. Ці алгоритми працюють більш ефективно і дають можливість рішати задачі великої розмірності , ніж при представленні матриць систем в загальній формі.

Вибір методу рішення СЛАУ суттєвим чином зв'язаний зі способом зберігання матриць, який, в свою чергу, визначається видом матриці. Основними являються три способи: загальний, симетричний і діагональний.

Існуючі методи рішення СЛАУ можна розбити на два класи: прямі і ітераційні.

Метод рішення відносять до класу прямих (точних, кінцевих) якщо в припущенні відсутності похибок заокруглення він дає точне рішення задачі після кінцевого числа операцій.

Ітераційні методи принципово дають можливість отримувати не рішення СЛАУ, а тільки послідовність, що до нього сходиться. Деякий, достатньо близький до межі член цієї послідовності приймається за наближене рішення. Інакше кажучи, ітераційні методи мають деяку теоретично обумовлену похибку рішення (методичну похибку).

Прямі методи частіше використовуються у випадках, коли матриця СЛАУ являється матрицею загального виду або спеціальною матрицею, наприклад трикутною або симетричною. Для великих розріджених матриць , при зберіганні яких використовуються їх особливі властивості, частіше застосовують ітераційні методи. У зв'язку зі згаданою особливістю структури інформаційної матриці в подальшому будуть розглянуті прямі методи.

### 3.2. Вплив помилок округлення і похибок вихідних даних

Серед головних компонентів похибок рішення СЛАУ, виділяються похибки, викликані округленням чисел в ЕОМ, і похибки, породжені неточністю вихідної інформації , заданої коефіцієнтами системи і її правою частиною.

Заокруглення чисел, неминує при їх представленні з використанням любого кінцевого числа розрядів , приводить до того, що наближені арифметичні операції додавання (віднімання) і множення (ділення) в ЕОМ мають інші властивості, ніж точні: вони не асоціативні , не дистрибутивні, в машинній арифметиці існують дільники нуля (добуток ненульових співмножників може виявитися рівним нулю) і т.і.

При виконанні великого числа арифметичних операцій, обумовлюючих результат, похибки округлення накопичуються. У зв'язку з цим при інших рівних умовах , віддають перевагу алгоритмам , які мають менше число операцій. А саме тому, наприклад, правило Крамера або еквівалентний йому метод безпосереднього обернення матриці системи. (3.82), потребує виконання приблизно  $n!n$  операцій множення , не витримують конкуренції з прямими методами, для яких необхідне число множень має порядок  $n^3$  (див.нижче).

Безпосередній кількісний аналіз поширення похибок округлень і перетворення їх в похибку кінцевого результату (прямий аналіз похибок округлення) представляє собою складну задачу. Тому на практиці часто використовують так званий обернений аналіз похибок округлень [7]. При його застосуванні вважають, що отримане наближене рішення СЛАУ являється точним рішенням деякої другої системи із зміненою матрицею.

Замість оцінки степеня відмінності точного і наближеного рішень визначається степінь відмінності коефіцієнтів вихідної і зміненої системи. На практиці елементи матриці системи (див. розд.1) часто отримують в ході фізичних вимірів, принципово виконуваних з деякими похибками. Тому, якщо здається, що відмінність елементів матриць вихідної і зміненої систем не перевищує похибок вимірів, то можна вважати, що обчислення проводяться з достатньо високою точністю і похибки округлення суттєво не впливають на результат. В протилежному випадку необхідно змінити алгоритм рішення з метою зменшення числа необхідних обчислювальних операцій і (або) збільшити число розрядів в представленні чисел, наприклад, шляхом переходу до чисел подвійної довжини в ЕОМ (представлення чисел з подвійною точністю).

При рішенні практичних задач коефіцієнти і вільні члени СЛАУ, як правило, відомі не точно, а з деякими похибками. Крім того, при введенні цих даних в ЕОМ часто проводиться їх додаткове спотворення в результаті округлень. У зв'язку з цим важливо вміти оцінювати отриману похибку рішення, розуміти причини її виникнення, а при неприйнятному її рівні вибирати такі методи, які забезпечували б мінімальну чутливість результату до похибок вихідних даних, змінювати схему експериментів (що породжують матрицю  $A$  і вектор  $b$ )

і підвищити точність вимірів.

Розглянемо випадок, коли з похибкою відома права частина векторно-матричного рівняння  $Ax = b$ : замість вектора  $b$  в СЛАУ використовується вектор  $b_\varepsilon = b + \varepsilon$ , де  $\varepsilon$  - вектор похибок, який, природно, не може бути відокремлений від вектора  $b$ . В цих умовах рівняння (3.82) перепишемо у вигляді

$$Ax \equiv b_\varepsilon \quad \text{або} \quad Ax = b + \varepsilon.$$

Лінійність рівняння і адитивність похибки правої частини дозволяє представити його рішення  $\hat{x}_\varepsilon = A^{-1}b_\varepsilon$ . У вигляді суми рішення рівняння (3.82) з точно відомою правою частиною  $\hat{x}_T$  («точного рішення») і похибки рішення  $b_\varepsilon \hat{x}$ , викликані збуренням  $\varepsilon$ :

$$\hat{x}_\varepsilon = \hat{x}_T + \delta_\varepsilon \hat{x}.$$

Приймаємо, що число рівнянь дорівнює числу невідомих і СЛАУ визначена, тобто  $A$  представляє собою не вироджену квадратну матрицю. Нагадаємо, що запис рішення системи рівнянь у вигляді  $\hat{x} = A^{-1}b$  має чисто символічний сенс і не обов'язково означає, що при його отриманні обчислюється обернена матриця і множиться на вектор правих частин.

Очевидно, що похибка рішення і похибка правої частини зв'язані рівнянням  $A\delta_\varepsilon \hat{x} = \varepsilon$ , звідки  $\delta_\varepsilon \hat{x} = A^{-1}\varepsilon$ .

Так як  $\|\delta_\varepsilon \hat{x}\| \leq \|A^{-1}\| \cdot \|\varepsilon\|$ ;  $\|b\| \leq \|A\| \cdot \|\hat{x}_T\|$  і  $\|\hat{x}_T\| \geq \frac{\|b\|}{\|A\|}$ , то

відносна похибка рішення  $\gamma_{x\varepsilon}$ , обумовлена похибкою  $\varepsilon$  заданою правою частиною СЛАУ, задовольняє нерівності

$$\gamma_{x\varepsilon} := \frac{\|\delta_\varepsilon \hat{x}\|}{\|\hat{x}_T\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|\varepsilon\|}{\|b\|} = \text{cond}A \gamma_b. \quad (3.83)$$

Із нього виходить, що відносна похибка рішення може в  $\text{cond}A$  раз перевершувати відносну похибку правої частини  $\gamma_b := \|\varepsilon\|/\|b\|$ . Виразом  $\text{cond}$  позначають число обумовленості матриці. Підчеркнемо, що цей результат ніяк не зв'язаний з методом рішення СЛАУ і справедливий при використанні будь-якої векторно-матричної норми. Видно, який великий вплив на точність рішення виявляє число обумовленості матриці системи.

Тепер оцінимо похибку рішення, обумовлену збуренням матриці системи. В цьому випадку замість матриці  $A$  в рівнянні (3.82) застосовується матриця  $A_\delta = A + \delta A$ , де  $\delta A$  – матриця збурення. По тим же причинам, що і раніше рішення рівняння  $A_\delta x = b$ , що має вигляд  $\hat{x}_{\delta A} = A_\delta^{-1} b$ , можна представити як суму «точного рішення» і похибки, визваної збуренням матриці  $\hat{x}_{\delta A} = \hat{x}_T + \delta_A \hat{x}$ .

Це рішення задовольняє умову  
 $(A + \delta A)(\hat{x}_T + \delta_A \hat{x}) = b$ .

Будемо вважати матрицю  $A_\delta$ . Із цього рівняння можна визначити похибки рішення [7]:

$$\begin{aligned} \hat{x}_T + \delta_A \hat{x} &= (A + \delta A)^{-1} b \Rightarrow \delta_A \hat{x} = (A + \delta A)^{-1} b - \hat{x}_T = \\ &= (A + \delta A)^{-1} \delta A \hat{x}_T. \end{aligned}$$

Звідси виходить співвідношення для відносної похибки рішення  $\gamma_{xA}$ , обумовлене похибкою задавання матриці системи рівнянь:

$$\begin{aligned} \gamma_{xA} &:= \frac{\|\delta_A \hat{x}\|}{\|\hat{x}_T\|} \leq \|A + \delta A\| \cdot \|A + \delta A\|^{-1} \frac{\|\delta A\|}{\|A + \delta A\|} \approx \\ &\approx \text{cond}A_{\delta A}. \end{aligned} \quad (3.84)$$

Величина  $\|\delta A\| / \|A + \delta A\| \approx \|\delta A\| / \|A\| := \gamma_A$  по суті представляє собою відносну похибку задавання матриці системи. Як видно із рівняння (3.84), і в цьому випадку число обумовленості матриці системи виступає в якості своєрідного «коефіцієнта підсилення» відносної похибки задавання вихідних даних.

В загальному випадку, при одночасному збуренні матриці системи і її правої частини рівняння (3.82) приймає вигляд

$$(A + \delta A)x = (b + \varepsilon) \quad \text{або} \quad A_\delta x = b_\varepsilon.$$

Можна показати, що відносна похибка  $\gamma_{xA\varepsilon}$  рішення  $\hat{x}_{\delta A\varepsilon}$  цього рівняння містить три доданків

$$\gamma_{xA\varepsilon} := \frac{\|\hat{x}_{\delta A\varepsilon} - \hat{x}_T\|}{\|\hat{x}_T\|} = \gamma_{xA} + \gamma_{x\varepsilon} + \gamma_{xA}\gamma_{x\varepsilon},$$

де  $\gamma_{xA}$  і  $\gamma_{x\varepsilon}$  – відносні похибки, введені раніше. Так як, як правило  $\gamma_{xA}\gamma_{x\varepsilon} < \gamma_{xA}$  і  $\gamma_{xA}\gamma_{x\varepsilon} < \gamma_{x\varepsilon}$ , а  $\text{cond}A \approx \text{cond}A_\delta$ , то цей вираз можна переписати:

$$\gamma_{xA\varepsilon} \leq \text{cond}A_\delta [\gamma_A + \gamma_b]. \quad (3.85)$$

Таким чином, повна відносна похибка традиційних методів СЛАУ визначається числом обумовленості матриці системи і відносними похибками вихідних даних. В багатьох практичних регресійних задачах числа обумовленості можуть досягати дуже великих значень (порядка  $10^2$ - $10^4$ ) [1, 5, 15, 21, 44], а точність вихідних даних обмежена – відносні похибки мають порядок  $10^{-5}$  –  $10^{-1}$ . Тому, як витікає із отриманих виразів, традиційні методи часто не дозволяють отримати задовільні по точності рішення. Для рішення погано обумовлених СЛАУ розроблені спеціальні методи, які використовують ідеї регуляризації, вони будуть розглянуті в розділі 5.

### 3.3. Прямі методи рішення лінійних рівнянь

Для пояснення суті різних обчислювальних методів в сучасній лінійній алгебрі широко використовується ідея представлення матриць СЛАУ у вигляді добутку або суми других матриць спеціального виду. Такі представлення називають *розкладами* або *факторизаціями*. [7].



Найбільш використовувані методи рішення СЛАУ можна розділити на дві групи: методи, що використовують розклади на *трикутні множники*, і методи зв'язані з використанням *ортогональних матриць*. До першої групи відносяться *метод Гауса і його модифікації*, *метод квадратного корня і схема Халецького*, до другої-*метод QR-розкладу*, *метод сингулярних розкладів* і деякі інші.

Мета **факторизації** при рішенні СЛАУ полягає в тому, щоб перейти від системи (3.82)

$$Ax = b$$

до еквівалентної системи

$$Ux = b' \quad (3.86)$$

або

$$Dx = b'', \quad (3.87)$$

де  $U$  і  $D$  – трикутна і діагональна матриці.

Розглянемо, яким чином може бути отримано рішення цих еквівалентних систем рівнянь.

Запишемо структуру СЛАУ (3.86) у розгорнутому вигляді. Для визначеності будемо вважати матрицю  $U$  *верхньою трикутною*:

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1m} \\ & u_{22} & u_{23} & \dots & u_{2m} \\ & & u_{33} & \dots & u_{3m} \\ & & & \dots & \\ 0 & & & & u_{mm} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_m \end{bmatrix} = \begin{bmatrix} b'_1 \\ b'_2 \\ b'_3 \\ \dots \\ b'_m \end{bmatrix}. \quad (3.88)$$

Із виразів (3.88) видно, що останнє рівняння  $u_{mm}x_m = b'_m$  утримує єдиний невідомий аргумент  $x_m$  і може бути безпосередньо вирішене відносно нього:

$$\hat{x}_m = b'_m / u_{mm}.$$

В передостаннє рівняння входять два невідомих вже-  $x_{m-1}$  і  $x_m$ , але останнє вже визначено, тому цього рівняння достатньо, щоб визначити

$$\hat{x}_{m-1} = \frac{b'_{m-1} - u_{(m-1)m}\hat{x}_m}{u_{(m-1)(m-1)}}.$$

Міркуючи по аналогії, на кожній ітерації будемо визначати із чергового рівняння значення чергового невідомого:

$$\hat{x}_i = \left( b'_i - \sum_{j=i+1}^m u_{ij}\hat{x}_j \right) / u_{ii} \quad (i = m-1, 1). \quad (3.89)$$

Ця процедура називається підстановкою. Очевидно, що для її реалізації необхідно виконати  $N \sim m^2/2$  операцій множення.

Аналогічно рішається СЛАУ з нижньою трикутною матрицею. Відповідна процедура називається прямою підстановкою.

І ще простіше рішити систему рівнянь (3.87), яка в розгорнутому вигляді виглядає наступним чином:

$$\begin{bmatrix} d_{11} & 0 & 0 & \dots & 0 \\ 0 & d_{22} & 0 & \dots & 0 \\ 0 & 0 & d_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & d_{mm} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_m \end{bmatrix} = \begin{bmatrix} b''_1 \\ b''_2 \\ b''_3 \\ \dots \\ b''_m \end{bmatrix}. \quad (3.90)$$

В даному випадку СЛАУ розпадається на сукупність із  $m$  незалежних рівнянь з одним невідомим кожне, які визначаються так:

$$x_i = b_i / d_{ii} \quad (i = \overline{1, m}). \quad (3.91)$$

Для рішення системи необхідно виконати  $m$  операцій множення (ділення).

### 3.4. Методи LU – розкладання. Метод Гауса

При використанні методу LU – розкладання Матриця  $\mathbf{A}$  представляється у вигляді добутку нижньої трикутної матриці  $\mathbf{L}$  і верхньої трикутної матриці  $\mathbf{U}$ :  $\mathbf{A}=\mathbf{LU}$  (позначення матриць введені по початковим буквам слів *lower i upper*) / Алгоритми відрізняються способом визначення матриць  $\mathbf{L}$  і  $\mathbf{U}$  по матриці  $\mathbf{A}$ .

Серед методів LU- розкладання найбільш відомим є метод Гауса. Його ідея полягає в тому, що послідовно за  $m-1$  кроків обчислень визначається матриця  $\mathbf{C}=\mathbf{L}^{-1}$ , яка забезпечує приведення вихідної системи до еквівалентної системи (3.86) з верхньою трикутною матрицею:

$$\mathbf{CAx} = \mathbf{Cb} \Leftrightarrow \mathbf{L}^{-1}\mathbf{LUx} = \mathbf{Cb} \Leftrightarrow \mathbf{Ux} = \mathbf{b}'. \quad (3.92)$$

Етап переходу від вихідної системи до системи (3.86) називається прямим ходом (прямим виключенням), а розглянутий раніше етап отриманого рівняння – оберненим ходом (оберненою підстановкою). Запишемо розгорнуту форму матричного представлення СЛАУ  $\mathbf{Ax}=\mathbf{b}$ :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}.$$

Процедура прямого ходу, яка включає в себе як операції по формуванню матриці  $\mathbf{C}$ , так і множення на неї вихідної матриці системи  $\mathbf{A}$ , виконується за  $m-1$  кроків. Розглянемо один із можливих варіантів його реалізації (схему єдиного ділення) [6]. Для спрощення будемо вважати, що  $a_{ij} \neq 0 \quad (i = \overline{1, m})$ .

На першому кроці обчислимо  $m-1$  множників:

$$M_i^{(1)} = a_{i1} / a_{11} \quad (i = \overline{2, m}),$$

і віднімемо із кожного  $j$ -го рівняння перше, помножене на  $M_j$ . Розглядаючи нові коефіцієнти при невідомих і вільні члени перетворених рівнянь

$$a_{ij}^{(1)} = a_{ij} - M_i^{(1)} a_{1j};$$

$$b_i^{(1)} = b_i - M_i^{(1)} b_1 \quad (i = \overline{2, m}; j = \overline{1, m}),$$

можна переконатись, що для всіх рівнянь, починаючи з другого, коефіцієнти при першому невідомому

$$a_{i1}^{(1)} = 0 \quad (i = \overline{2, m}).$$

Перетворена система рівнянь запишеться в наступному вигляді:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ 0 & a_{22}^{(1)} & \dots & a_{2m}^{(1)} \\ \dots & \dots & \dots & \dots \\ 0 & a_{m2}^{(1)} & \dots & a_{mm}^{(1)} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \dots \\ b_m^{(1)} \end{bmatrix}.$$

Продовжуючи таким же чином на другому і послідовних кроках, можна виключити  $x_2$  із останніх  $m-2$  рівнянь [обчислюються множники

$M_i^{(2)} = a_{i2}^{(1)} / a_{22}^{(1)} \quad i = \overline{3, m}$ ], після чого із останніх  $m-3$  рівнянь [обчислюються множники  $M_i^{(3)} = a_{i3}^{(2)} / a_{33}^{(2)} \quad (i = 4, m)$ ] і т.д.

На  $k$ -му кроці виключається  $x_k$  за допомогою множників

$M_i^k = a_{ik}^{(k-1)} / a_{kk}^{(k-1)}$  ( $i = \overline{k+1, m}$ ). При цьому коефіцієнти і праві частини залишених  $m - k$  рівнянь визначаються наступними очевидними виразами:

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - M_i^{(k)} a_{kj}^{(k-1)}; b_i^{(k)} = b_i^{(k-1)} - M_i^{(k)} b_k^{(k-1)}$$

( $i = \overline{k+1, n}; j = \overline{k, n}$ ).

При цьому система рівнянь приймає вигляд

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1(k-1)} & a_{1k} & a_{1(k+1)} & \dots & a_{1m} \\ 0 & a_{22}^{(1)} & \dots & a_{2(k-1)}^{(1)} & a_{2k}^{(1)} & a_{2(k+1)}^{(1)} & \dots & a_{2m}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{kk}^{(k-1)} & a_{k(k+1)}^{(k-1)} & \dots & a_{km}^{(k-1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{mk}^{(k-1)} & a_{m(k+1)}^{(k-1)} & \dots & a_{mm}^{(k-1)} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_k \\ \dots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \dots \\ b_k^{(k-1)} \\ \dots \\ b_m^{(k-1)} \end{bmatrix}.$$

Після виконання (m-1)-го кроку проходить виключення  $x_{(m-1)}$  із останнього рівняння і остаточно СЛАУ записується у формі (3.86) наступним чином:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1(m-1)} & a_{1m} \\ 0 & a_{22}^{(1)} & \dots & a_{2(m-1)}^{(1)} & a_{2m}^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{(m-1)(m-1)}^{(m-2)} & a_{(m-1)m}^{(m-2)} \\ 0 & 0 & \dots & 0 & a_{mm}^{(m-1)} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_{m-1} \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \dots \\ b_{m-1}^{(m-2)} \\ b_m^{(m-1)} \end{bmatrix}.$$

Нагадаємо, що після завершення прямого ходу виконується обернене ний (зворотній) хід методу Гауса, розглянутий вище. Елементи  $a_{ij}^{(i-1)}$ , які стоять вздовж головної діагоналі матриці системи, називаються ведучими елементами.

Дамо матричну інтерпретацію процесу виключення, відповідну обчислювальній схемі (3.82). Кожний  $k$ -й крок ( $k = \overline{1, m-1}$ ) прямого ходу еквівалентний множенню системи рівнянь зліва на матрицю

$$C_k = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & M_{k+1}^{(k)} & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 & M_{k+2}^{(k)} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & M_m^{(k)} & 0 & 1 & \end{bmatrix}.$$

При використанні таких матриць процедура прямого ходу записується наступним чином:

$$CAx = Cb,$$

де  $C = C_{m-1} C_{m-2} \dots C_1$ , що співпадає з виразом (3.92).

### Модифікація методу Гауса

Насправді, в описаній формі метод Гауса майже ніколи не використовується. Це зв'язано з тим, що окремі елементи  $a_{ii}^{i-1}$  (центральні або ведучі елементи), на які виконується ділення рівнянь, можуть бути рівні нулю.

Для подолання цієї проблеми достатньо поступити наступним чином: при  $a_{ii}^{i-1}$  можна поміняти місцями  $i$ -е рівняння з одним із подальших. При не виродженій матриці  $A$  знайдеться хоча б одна така перестановка. При цьому отримана система являється еквівалентною вихідній.

Правила, які визначають порядок перестановок, називають стратегіями перестановок.

Подальша модифікація направлена на те, щоб за допомогою перестановок уникнути ділення не тільки на нульові, але і на малі ведучі елементи. Це зв'язано з тим, що похибка результату ділення двох чисел обернено пропорційна абсолютній величині дільника.

Серед різних стратегій вибору ведучого елемента часто використовують так зване часткове упорядкування. Його суть полягає в тому, що при виборі першого і любого чергового, і-го ведучого рівняння перевага віддається тому, що не побували ведучими рівняннями, в якому модуль елемента із і-го стовпчика максимальний. Воно стає ведучим, помінявшись місцями з рівнянням, яке займало його місце раніше. Такі перестановки можливі на кожному кроці. Це *стратегія вибору ведучого елемента по стовпчику*. Відомі також алгоритми, які використовують схему вибору *ведучого елемента по строчці*, також виконують перестановку рівнянь.

Перестановка рівнянь може бути описана шляхом множення вихідної системи на деяку перестановочну матрицю **P**, елементи якої можуть бути тільки нулями, або одиницями, причому в кожній строчці і в кожному стовпці є рівно одна одиниця. Тому, модифікації методу Гауса, в яких використовується впорядкування рівнянь, представляють собою алгоритми, які використовують **LU**-розклади матриці **PA**:

$$PA = LU.$$

Подальшим вдосконаленням методу виключення являється метод Гауса з вибором головного елемента (метод головних елементів), в якому ведучий елемент визначається по всій матриці системи. Цей метод також називають методом Гауса з повним впорядкуванням.

Пояснимо його суть з використанням *розширеної матриці системи* (3.82) [6, 7, 82].

$$A = \begin{bmatrix} a_{11} & \dots & a_{1q} & \dots & a_{1m} & b_1 \\ a_{p1} & \dots & \{a_{pq}\} & \dots & a_{pm} & b_p \\ a_{m1} & \dots & a_{mq} & \dots & a_{mm} & b_m \end{bmatrix}.$$

Серед елементів матриці  $a_{ij}$  виберемо найбільший по модулю, що називається головним елементом. Нехай, їм буде, наприклад, елемент  $a_{pq}$ . Строчка з номером «р», що містить головний елемент, називається головною строчкою, а відповідне рівняння системи - головним рівнянням.

Дальше обчислюємо множники  $M_i^{(1)} = a_{ij} / a_{pq}$  для всіх  $i \neq p$ . Після перетворюємо матрицю наступним чином: із кожної і-ї строчки віднімаємо головну строчку, помножену на  $M_i^{(1)}$ . В результаті отримаємо матрицю, в якій всі елементи q-того стовпчика, крім  $a_{pq}$ , дорівнюють нулю. Відкидаючи цей стовпець і головну строчку, отримаємо нову матрицю  $\bar{A}_1$  з меншим на одиницю числом строчок і стовпців.

Над матрицею  $\bar{A}_1$  повторюються ті ж операції, після чого отримуємо матрицю  $\bar{A}_2$  і т.д. Так продовжується до тих пір, поки не отримується матриця строчка, яка має коефіцієнт перед останнім невідомим і вільний член рівняння. В подальшому об'єднуються всі головні строчки, починаючи з останньої. Якщо тепер змінити порядок проходження невідомих у векторі **x**, то матриця системи стає трикутною.

Відмітимо, що класичний метод Гауса являється окремим випадком методу головних елементів, якщо в

якості головного елемента завжди вибрати лівий верхній елемент матриці системи.

В розглянутих модифікаціях методу Гауса **LU**-розкладання проводилось в неявній формі [див. вираз (3.92)]. В той же час існують реалізації алгоритму, в яких **LU**-розкладання проводилось в неявній формі [див. вираз (3.92)]. В той же час існують реалізації алгоритму, в яких **LU**-розкладання використовується безпосередньо [6, 82].

Значення елементів трикутних матриць **L** і **U** можуть бути отримані із рівності **A=LU**. Система рекурентних співвідношень для визначення елементів  $l_{ij}$  і  $u_{ij}$  виглядає наступним чином [6, 7, 82]:

$$\begin{aligned} l_{ij} &= 0 \quad \text{при} \quad i < j; \\ u_{ij} &= 0 \quad \text{при} \quad i > j \quad (i, j = \overline{1, m}); \\ u_{11} &= a_{11}, \quad u_{1j} = a_{1j}, \quad l_{j1} = a_{j1} / u_{11} \quad (j = \overline{2, m}); \\ u_{ii} &= a_{ii} - \sum_{p=1}^{i-1} l_{ip} u_{pi} \quad (i = \overline{2, m}); \\ u_{ij} &= a_{ij} - \sum_{p=1}^{i-1} l_{ip} u_{pj} \cdot l_{ji} = \left( a_{ji} - \sum_{p=1}^{j-1} l_{jp} u_{pi} \right) / u_{ii} \\ &\quad (i = \overline{2, m}; \quad j = \overline{i+1, m}). \end{aligned} \quad (3.93)$$

Ці формули часто називають компактною схемою методу Гауса для трикутного розкладання матриці системи або схемою Халецького.

Після визначення **LU**-розкладання на основі використання послідовності еквівалентних перетворень

$$\left. \begin{aligned} Ax = b &\Leftrightarrow L U x = b \Leftrightarrow L^{-1} L U x = L^{-1} b \equiv b' \Leftrightarrow \\ \left\{ \begin{aligned} U x &= b'; \\ L b' &= b \end{aligned} \right. \end{aligned} \right\} \quad (3.94)$$

Рішення вихідної СЛАУ з матрицею **A** загального виду можна звести до послідовного рішення двох систем з трикутними матрицями:  $Lb' = b$ ;  $Ux = b'$ . Рішення СЛАУ з верхньою трикутною матрицею отримується також просто, як і для системи з нижньою трикутною матрицею (див. вище).

Другі варіанти методу **LU**-розкладання, зокрема відомий алгоритм Краута-Дуліттла, описані в роботах [6, 79, 82, 85].

### 3.5. Метод квадратних коренів

Цей метод (називають іноді, як і схему розглянуту вище, методом Халецького або методом Банахевича) використовується при рішенні систем рівнянь із симетричною позитивно визначеною матрицею.

Метод оснований на факті існування трикутного розкладання симетричної позитивно визначеної матриці **A** слідуючого виду;

$$\mathbf{A} = \mathbf{U}^T \mathbf{U}, \quad (3.95)$$

де **U** – верхня трикутна матриця.

Видно, що це розкладання представляє собою окремий випадок **LU**-розкладання, використовуючи чого вказану специфіку матриці **A**.

Елементи матриці **U** отримуються шляхом рішення системи рівнянь

$$\sum_{k=1}^i u_{ki} u_{kj} = a_{ij} \quad (i = \overline{1, m}; \quad j = \overline{i, m}).$$

Ці рівняння безпосередньо походять із виразу (3.95).

Рішення системи рівнянь приводить до слідуючих рекурентних формул, добре пристосованих до реалізації на ЕОМ [6, 7, 82]:

$$u_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} u_{ki}^2};$$

$$u_{ij} = \left( a_{ij} - \sum_{k=1}^{i-1} u_{ki} u_{kj} \right) / u_{ii} \quad (i = \overline{1, m}; j = \overline{i+1, m}); \quad (3.96)$$

$$u_{ij} = 0 \quad \text{при} \quad i > j.$$

Після отримання розкладання (3.95) рішення системи зводиться по аналогії з (3.94) до рішення двох систем рівнянь з трикутними матрицями:

$$\begin{cases} U^T b' = b; \\ Ux = b'. \end{cases}$$

Іноді при рішенні методом квадратних коренів використовується розкладання матриці системи у вигляді  $A = L^T L$ , (3.97) де  $L$ - нижня трикутна матриця. Елементи матриці  $L$  визначаються за формулами, які нескладно отримати по аналогії з (3.96):

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=i+1}^m l_{ki}^2}; \quad l_{ij} = \left( a_{ij} - \sum_{k=i+1}^m l_{ki} l_{kj} \right) / l_{ii}$$

$$(j = \overline{1, i-1}; i = \overline{m, 1}); \quad l_{ij} = 0 \quad \text{при} \quad i < j.$$

Для отримання рішення вихідної системи послідовно рішення дві системи з трикутними матрицями:

$$\begin{cases} L^T b' = b; \\ Lx = b'. \end{cases}$$

### 3.6. Методи QR-розкладення

Як і при рішенні других задач лінійної алгебри, при визначенні коренів СЛАУ велику роль грає представлення матриці системи  $A$  у вигляді добутку ортогональної матриці  $Q$  і верхньої трикутної матриці  $R$ :

$$A = QR.$$

Це розкладення часто називають  $QR$  – розкладенням матриці  $A$ , а основані на його використанні методи рішення алгебраїчних задач - методами  $QR$  – розкладення.  $QR$  – розкладення існує для будь-якої квадратної матриці  $A$ , причому, якщо матриця  $A$  не вироджена, то її  $QR$  – розкладення, в якому діагональні елементи матриці додатні, єдине [6, 7, 53, 82].

Якщо  $QR$  – розкладення матриці СЛАУ знайдено, то система (3.82) зводиться до еквівалентної системи

$$Rx = Q^T b \equiv b', \quad (3.98)$$

яка може бути знайдена методом зворотної підстановки. У виразах (3.98) враховано, що для ортогональної матриці  $Q$  виконується рівність  $Q^{-1} = Q^T$ .

Для отримання  $QR$  – розкладення матриці системи використовуються методи, основними із яких являються: *метод обертання* (Гівенса) і *метод відображень* (Хаусхолдера). Так як обидва ці методи детально розглядаються в підручниках полінійній алгебри, то ми дамо лише їх коротке пояснення [6, 7, 82].

При використанні метода Гівенса визначається послідовність матриць спеціального виду – матриць обертання –  $P_0^{(1)}, \dots, P_0^{(N)}$ , така, що матриця  $R = P_0^{(N)} \dots P_0^{(1)} A$  буде верхньою трикутною з додатними елементами на головній діагоналі.

Матрицею обертання називається матриця



трьохвимірному просторі розглянути дзеркальну площину, для якої вектор  $w$  являється нормальним, а площина проходить через початок координат, то перетворення дзеркального відображення від такої площини задається розгляданою матрицею відображення.

Нехай задані будь-які ненульові вектори  $a$  і  $s$ . Можна показати, що існує таке відображення (матриця  $P_B$  і визначаючий її вектор  $w$ ), яке переводить вектор  $a$  у вектор  $\alpha s$ , де  $\alpha$  - деяке число:  $P_B a = \alpha s$ . Вектор  $w$  при цьому задається співвідношеннями:

$$w = \frac{1}{\rho}(a - \alpha s); \quad \rho^2 = 2[a(a - \alpha s)^T]; \quad \alpha \pm \|a\|/\|s\|.$$

Звідси, зокрема, витікає, що завжди існує матриця відображення, яка дає можливість перетворити довільний вектор (наприклад, вектор-стовпець матриці системи) у вектор заданого виду.

Нехай  $A$  – задана матриця СЛАУ. Помножимо її зліва на послідовність матриць відображення  $P_0^{(1)}, \dots, P_0^{(m-1)}$ .

Кожну із цих матриць  $P_B^{(i)}$  ( $i = \overline{1, m-1}$ ) будемо вибирати із умови перетворення в нуль під діагональних елементів  $i$ -го стовпчика отриманого раніше добутку  $P_B^{(i-1)}, \dots, P_B^{(1)} A$ . В результаті цих операцій буде отримана необхідна верхня трикутна матриця  $R$ .

### 3.7. Метод сингулярного розкладення

Він оснований на такій факторизації матриці СЛАУ, яка дозволяє перейти до еквівалентної системи з діагональною матрицею, тобто до системи із  $m$  незалежних рівнянь з одним невідомим в кожному (див. вище).

Введемо основні поняття і визначення, які будуть використані також в шостому розділі [6, 7, 74, 79, 82].

Нехай  $A$  - довільна матриця розміром  $n \times m$ . Розглянемо квадратні матриці  $A^T A$  і  $AA^T$  порядку  $m$  і  $n$  відповідно. Можна показати, що ненульові власні значення цих матриць  $\lambda_i$  співпадають ( $\lambda_i > 0; i = \overline{1, r}$ , де  $r$  – ранг матриць  $A, A^T A, AA^T$ ). Арифметичні значення квадратних коренів із загальних власних значень матриць  $A^T A$  і  $AA^T$   $\mu_i = \sqrt{\lambda_i}$  називаються *сингулярними* (головними) числами матриці  $A$ . Будемо вважати, що перші  $r$  сингулярних чисел пронумеровані і розташовані в порядку убутання:  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > 0$ . Решта сингулярних чисел приймається рівною нулю:  $\mu_{r+1} = \dots = 0$ .

Сингулярні числа являються важливими характеристиками матриці  $A$  і мають наступні властивості:

1. Сингулярні числа матриці не змінюються при множенні її справа або зліва на ортогональну матрицю.
2. Квадратна матриця являється не виродженою тоді і тільки тоді, коли всі її сингулярні числа відмінні від нуля.
3. Модуль визначника матриці  $A$  дорівнює добутку всіх сингулярних чисел.

Справедлива наступна теорема, яка являється основою для розглядаємого методу рішення СЛАУ [7, 74, 79].

Для будь-якої матриці  $A$  розміру  $n \times m$  існує розкладення

$$A = U D V^T, \tag{3.99}$$

де  $U, V$  – ортогональні матриці порядку  $n$  і  $m$  відповідно;  $D$  – прямокутна діагональна матриця розміром  $n \times m$  із незростаючими елементами.

Розкладення (3.99) називається сингулярним розкладенням матриці  $A$ . Для матеріальної матриці  $A$  сингулярне розкладення являється матеріальним, тобто матриці  $U, D, V^T$  мають матеріальні елементи.

Матриці, які входять в сингулярне розкладення (3.99), мають наступний вигляд:



$$D = \text{diag}(\mu_1, \dots, \mu_r, 0, \dots, 0);$$

$$U = [u_i] \quad (i = \overline{1, n}); \quad V = [v_i] \quad (i = \overline{1, m}).$$

При цьому  $u_i$ ,  $v_i$  – власні вектори матриць  $AA^T$  і  $A^T A$  відповідно.

Сингулярне розкладення (3.99) матриці системи (3.82), якщо матриця  $A$  квадратна і не вироджена, дозволяє перейти до еквівалентної системи:

$$DV^T x = U^T b \equiv b'.$$

Звідси витікає, що рішення вихідної СЛАУ можна звести до послідовного рішення двох систем:

$$\begin{cases} Dz = b'; \\ V^T x = z; \end{cases} \quad (3.100)$$

перша із цих систем, що має діагональну матрицю обчислюється за формулами (3.93). При рішенні другої системи слід врахувати, що  $(V^T)^{-1} = V$ . В результаті рішення системи (3.100) можна представити у вигляді

$$\hat{x} = VD^{-1}U^T b = \sum_{i=1}^m \mu_i^{-1} v_i u_i^T b$$

або в рекурентній формі:

$$\hat{x}^{(i)} = \hat{x}^{(i-1)} + \mu_i v_i u_i^T b \quad (i = \overline{1, m}); \quad \hat{x}^{(0)} = 0.$$

Для виконання сингулярного розкладення матриці СЛАУ використовуються відомі методи визначення власних значень і власних векторів матриць  $A^T A$  і  $AA^T$  [6, 7, 79, 82].

### 3. 8. Порівняння методів

Розглянуті методи рішення СЛАУ [6, 7, 15, 53, 63, 79, 82], природно, мають різну стійкість до похибок округлення, потребують різної ємкості пам'яті і часу розрахунку.

Час рішення СЛАУ визначається числом арифметичних операцій (в першу чергу, операцій множення і ділення), необхідних для виконання факторизації матриці системи. Тому переваги мають методи  $LU$  – розкладення (табл.3.1): методи Гауса, компактна схема і, особливо, метод квадратного кореня. Додаткової оперативної пам'яті ці методи не потребують. По числу операцій і по ємкості необхідної пам'яті метод Хаусхолдера дещо кращий метода Гівенса і по цим характеристикам порівняно небагато відстає від метода Гауса. Найбільші обчислювальні труднощі виникають при реалізації метода сингулярного розкладення.

Таблиця 3.1. Порівняльна таблиця

Спосіб розкладення	Число операцій
Метод Гауса	$m^3/3$
Компактна схема методу Гауса	$m^3/3$
Метод Халецького (квадратного кореня)	$m^3/6$
Метод Гівенса	$m^3$
Метод Хаусхолдера	$2/3m^3$
Метод сингулярного розкладення	$3m^3$

Загальним недоліком методів  $LU$  – розкладення являється їх низька обчислювальна стійкість до похибок вихідних даних при поганій обумовленості матриці СЛАУ: в цих умовах класичний метод Гауса без вибору ведучого елемента практично непрацездатний, другі модифікації цього методу більш стійкі, але поступаються методам

других груп. Найбільш стійким із методів трикутного розкладу являється метод квадратного кореня. Причина цього недоліку полягає в тому, що при LU – розкладенні число обумовленості матриці системи росте, що приводить до росту похибок рішення.

Методи Гівенса і Хаусхолдера близькі між собою по рівню похибки округлення і по цьому показнику лише небагато поступаються модифікаціям методу Гауса. Разом з тим вони вільні від відміченого вище основного недоліку методів LU – розкладення, так як норми матриць **A** і **R** співпадають і перехід до еквівалентної системи з трикутною матрицею не приводить до росту числа обумовленості.

Метод сингулярних розкладень в класичному варіанті не забезпечує переваг по рівню обчислювальної стійкості в порівнянні з методами **QR** – розкладення.

Але, як буде показано в розділі 5, він являється доброю базою для створення високоефективних алгоритмів, призначених спеціально для роботи з погано обумовленими і виродженими матрицями.

#### Розділ 4. Регресійний аналіз найпростіших поліноміальних моделей

Розглянемо процедури оцінки параметрів і статистичного аналізу регресійних поліноміальних моделей нульового, першого і другого порядку.

##### 4.1. Поліноміальна модель нульового порядку

В цьому випадку регресія не містить дійсних факторів. Єдиним фактором являється єдина змінна  $x_i'$ , тотожно рівна одиниці [див. вираз [1.7]]:

$$Y = \beta_1 x_i' + \varepsilon = \beta_1 + \varepsilon. \quad (4.101)$$

Так як  $M\{\varepsilon\} = 0$ , то  $\beta_1$  представляє собою математичне сподівання випадкової величини  $Y : M\{Y\} = \beta_1$ . Реалізації результуючої змінної

$$Y = \beta_1 + \varepsilon_i \quad (i = \overline{1, n}).$$

Звідси нескладно отримати тривіальний вираз для векторів і матриць, які входять в загальну форму регресійної моделі [1.11]:

$$F = [1 \dots 1]^T; \quad F^T = [1 \dots 1]; \quad y = [y_1, \dots, y_n]^T; \quad \beta = \beta_1.$$

Так як  $(F^T F) = 1/n$ , то

$$\hat{\beta}_1 = \hat{M}\{Y\} = \frac{1}{n} [1 \dots 1] y = \frac{1}{n} \sum_{i=1}^n y_i. \quad (4.102)$$

Цей вираз, природно, співпадає з використаною раніше формулою для оцінки математичного сподівання випадкової (гаусів розподіл) величини

$$\hat{\mu}_x \equiv m_x = \frac{1}{n} \sum_{i=1}^n x_i.$$

Дисперсія оцінки має вигляд

$$\sigma^2\{\hat{\beta}_1\} = \sigma_\varepsilon^2 / n, \quad (4.103)$$

А оцінка дисперсії збурень

$$s^2 \equiv \hat{\sigma}_\varepsilon^2 = \frac{1}{(n-1)} (y - F\beta_1)^T (y - F\beta_1) = \frac{1}{n-1} * \sum_{i=1}^n (y_i - \hat{\beta}_1)^2 = \frac{1}{n-1} \sum_{i=1}^n \left( y_i - \frac{1}{n} \sum_{i=1}^n y_i \right)^2. \quad (4.104)$$

Отримані вирази є незміщеними і асимптотично ефектив-

ними.

Аналогічно може бути отриманий рекурентний алгоритм оцінювання:

$$\hat{\beta}_1^{(i+1)} \equiv \hat{M}\{Y\}_{i+1} = \hat{\beta}_1^{(i)} + c_{i+1}[y_{i+1} - \hat{\beta}_1^{(i)}]; \quad (4.105)$$

$$c_{i+1} = s_i / (s_i + 1); \quad s_{i+1} = s_i - s_i^2 / (s_i + 1) = 1 / (1 + 1/s_i).$$

Так як для першого відліку дисперсія оцінки  $\hat{\beta}_1^{(1)} = \sigma_\varepsilon^2 s_1 = \sigma_\varepsilon^2$ , то початкова умова  $s_1=1$ . В даному випадку  $s_{i+1} = 1/(1+i)$ ;  $c_{i+1} = 1/(i+1)$  і отримаємо вираз

$$\hat{\beta}_1^{(i+1)} = \hat{\beta}_1^{(i)} + \frac{1}{1+i}[y_{i+1} - \hat{\beta}_1^{(i)}]. \quad (4.106)$$

Аналогічно отримуємо рекурентну оцінку дисперсії збурень  $D \equiv \sigma_\varepsilon^2$ :

$$\hat{D}^{(i+1)} = \frac{i}{i+1} \hat{D}^{(i)} + \frac{1}{i+1}[y_{i+1} - \hat{\beta}_1^{(i+1)}]^2. \quad (4.107)$$

Вирази (4.106) і (4.107) входять в автоматизовану систему опрацювання експериментальних даних для отримання текучих оцінок числових характеристик з широко використовуваними рекурентними процедурами. Вони основані на використанні при формуванні текучої оцінки оцінок параметрів, отриманих на попередніх кроках, і їх корекції в результаті використання знову поступаючих експериментальних даних. Рекурентні алгоритми потребують суттєво меншої ємкості пам'яті ЕОМ, ніж алгоритми інших реалізацій.

## 4.2. Поліноміальна модель першого порядку

Для визначеності розглянемо лінійну по факторам модель наступного виду:

$$Y = \beta_2 x + \beta_1 + \varepsilon, \quad (4.108)$$

де  $x \equiv x_2'$  - єдина реальна змінна [див.(1.7)].

В даному випадку матриці, що входять в рівняння регресії, мають вигляд:

$$F = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}; \quad F^T F = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix};$$

$$(F^T F)^{-1} = \frac{1}{\Delta} \begin{bmatrix} \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i & \\ -\sum_{i=1}^n x_i & n \end{bmatrix},$$

де

$$\Delta = n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 - \text{визначник матриці похибок}; \quad x_i = x_{i2}'.$$

Таким чином, оцінки параметрів будуть:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2}; \quad (4.109)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}. \quad (4.110)$$

Останню оцінку зручніше знаходити із рівняння для середніх значень пояснюючої і результуючої змінної:

$$\hat{M}\{Y\} = \beta_2 \bar{x} + \beta_1. \quad (4.111)$$

Звідси

$$\hat{\beta}_1 = \frac{1}{n} \left( \sum_{i=1}^n y_i - \hat{\beta}_2 \sum_{i=1}^n x_i \right). \quad (4.112)$$

Після визначення оцінок параметрів може бути найдена оцінка функції регресії

$$\hat{\eta}(x) = \hat{\beta}_2 x + \hat{\beta}_1. \quad (4.113)$$

Рекурентні оцінки параметрів

$$\begin{aligned} \hat{\beta}_2^{(i+1)} &= \hat{\beta}_2^{(i)} + c_{2i} (y_{i+1} - \hat{\beta}_2^{(i)} x_{i+1} - \hat{\beta}_1^{(i)}); \\ \hat{\beta}_1^{(i+1)} &= \hat{\beta}_1^{(i)} + c_{1i} (y_{i+1} - \hat{\beta}_2^{(i)} x_{i+1} - \hat{\beta}_1^{(i)}), \end{aligned} \quad (4.114)$$

Отримують із загальних співвідношень (2.55), якщо прийняти  $f_i^T = [1 \quad x_i]$ .

Для перевірки адекватності лінійної моделі обчислюється дисперсійне співвідношення

$$F = \max(D, D_0) / \min(D, D_0).$$

При цьому

$$D_0 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{\beta}_2 x_i - \hat{\beta}_1) \quad (4.115)$$

- остаточна дисперсія;

$$D = \frac{1}{n-1} \sum_{i=1}^n (y_i - \hat{y})^2 = \frac{1}{n-1} \sum_{i=1}^n \left( y_i - \frac{1}{n} \sum_{i=1}^n y_i \right)^2 \quad (4.116)$$

- загальна дисперсія результуючої змінної.

Лінійна модель (4.108) вважається адекватною, якщо

$$F \leq F_{n-2, n-1, 1-\alpha}. \quad (4.117)$$

При перевірці адекватності моделі доцільно також оцінити значимість параметрів. Для цього вичислюються і порівнюються з квантилями t-розподілу Стьюдента статистики

$$g_i \equiv \left| \hat{\beta}_i \right| / \sqrt{\hat{\sigma}^2 \{ \hat{\beta}_i \}} \geq t_{n-2, \alpha} \quad (i = 1, 2). \quad (4.118)$$

При виконанні цієї нерівності значення відповідних параметрів вважаються що не суперечать експериментальним даним. В протилежному випадку рекомендується змінити модель регресії і знову рішення задачі оцінки її (тепер вже єдиного) параметра. Якщо незначимим виявився коефіцієнт  $\beta_2$ , то необхідно перейти до регресійної задачі розглянутої раніше. При встановленні не значимим коефіцієнта  $\beta_1$  рівняння регресії першого порядку буде вміщати тільки коефіцієнт  $\beta_2$ :

$$Y = \beta_2 x + \varepsilon. \quad (4.119)$$

Якщо параметри регресії виявилися значимими або якщо було прийняте рішення не змінювати модель і не відкидати незначимі коефіцієнти (вище відмічалось, що краще залишити не значимі коефіцієнти, ніж помилково –

вірогідність цієї похибки рівна  $\alpha$  – відкинути значимі), то в подальшому доцільно провести інтервальне оцінювання параметрів і значень функції регресії [див. вирази 2.78].

### 4.3. Поліноміальна модель другого порядку

На закінчення розглянемо процедуру двохфакторного регресійного аналізу на прикладі рівняння

$$Y = \beta_1 + \beta_2 x + \beta_3 x^2 + \varepsilon. \quad (4.120)$$

При цьому роль першого реального фактора  $x_2'$  грає величина  $x$ , а другого  $x_3'$  - її квадрат. В цьому випадку

$$\beta = [\beta_1 \quad \beta_2 \quad \beta_3]^T; \quad (4.121)$$

$$F = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \dots & \dots & \dots \\ 1 & x_n & x_n^2 \end{bmatrix}.$$

При обчисленні оцінок параметрів застосувати обернену матрицю  $(F^T F)^{-1}$  часто недоцільно і необхідно перейти до рішення системи нормальних рівнянь  $F^T F \beta = F^T y$ , яка, як можна показати, в скалярному вигляді буде

$$n\beta_1 + \left(\sum_{i=1}^n x_i\right)\beta_2 + \left(\sum_{i=1}^n x_i^2\right)\beta_3 = \sum_{i=1}^n y_i;$$

$$\left(\sum_{i=1}^n x_i\right)\beta_1 + \left(\sum_{i=1}^n x_i^2\right)\beta_2 + \left(\sum_{i=1}^n x_i^3\right)\beta_3 = \sum_{i=1}^n x_i y_i; \quad (4.122)$$

$$\left(\sum_{i=1}^n x_i^4\right)\beta_1 + \left(\sum_{i=1}^n x_i^3\right)\beta_2 + \left(\sum_{i=1}^n x_i^2\right)\beta_3 = \sum_{i=1}^n x_i^2 y_i.$$

Для її рішення можуть використовуватись методи, розглянуті в третьому розділі.

## Розділ 5. Особливості регресійного аналізу при порушенні базових положень

### 5.1. Регресійний аналіз при неоднорідних і корельованих збуреннях

Серед базових передумов класичного регресійного аналізу (див. розділ 2) особливо часто порушується умова про однорідність і некорельованість збурень. Покажемо, яким чином повинна бути модифікована процедура регресійного аналізу і до яких наслідків приведе використання класичного алгоритму в неадекватних умовах [1, 3, 15, 27, 67, 78].

Нехай, вектор збурень  $\varepsilon$  що входить в рівняння регресії  $Y = F\beta + \varepsilon$  характеризується кореляційною матрицею

$$V \equiv D\{\varepsilon\} = \sigma_\varepsilon^2 \Omega, \quad (5.123)$$

де  $\Omega$  – деяка позитивно визначена матриця.

Основна ідея модифікації полягає в тому, щоб знайти таке перетворення рівняння регресії з корельованими чи неоднорідними збуреннями, які дозволили б отримати класичну регресію, розглянуту раніше:

$$Z = Q\beta + \xi; \quad D\{\xi\} = \sigma_\varepsilon^2 E, \quad (5.124)$$

де  $E$  - одинична матриця.

Таке перетворення можна виконати, якщо використати представлення кореляційної матриці  $\Omega$  у вигляді

$$\Omega = TT^T. \quad (5.125)$$

Відмітимо, що представлення (5.125) не являється єдиним.

Помножимо праву і ліву частини вихідного рівняння регресії на  $T^{-1}$  і отримаємо

$$T^{-1}Y = T^{-1}F\beta + T^{-1}\varepsilon. \quad (5.126)$$

Приймаючи позначення

$$T^{-1}Y=Z; \quad T^{-1}F=Q; \quad T^{-1}\varepsilon=\xi, \quad (5.127)$$

прийдемо до моделі (4.124). Покажемо, що перетворені збурення  $\xi$  являються некорельованими і однорідними. Для цього визначимо кореляційну матрицю вектора  $\xi$ :

$$\begin{aligned} D\{\xi\} &:= M\{\xi\xi^T\} = M\{(T^{-1}\varepsilon)(T^{-1}\varepsilon)^T\} = \\ &= T^{-1}M\{\varepsilon\varepsilon^T\}(T^T)^{-1} = T^{-1}V(T^T)^{-1} = \sigma_\varepsilon^2 E, \end{aligned} \quad (5.128)$$

що і потрібно було показати. При цьому враховано, що  $M\{\xi\}=\mathbf{0}$ , тому що  $M\{\varepsilon\}=\mathbf{0}$ .

Перетворена модель (5.124) відповідає всім умовам класичного регресійного аналізу, тому для отримання оцінок параметрів може застосовуватись МНК:

$$\hat{\beta}_\Omega = (Q^T Q)^{-1} Q^T z.$$

Цю оцінку можна перетворити з врахуванням (5.127):

$$\begin{aligned} \hat{\beta}_\Omega &= [(T^{-1}F)^T (T^{-1}F)]^{-1} (T^{-1}F)^T (T^{-1}y) = \\ &= (F^T (T^{-1})^T T^{-1}F)^{-1} F^T (T^{-1})^T T^{-1}y = \\ &= (F^T \Omega^{-1} F)^{-1} F^T \Omega^{-1} y. \end{aligned} \quad (5.129)$$

Відповідним чином можна модифікувати і решту основних співвідношень регресійного аналізу:

$$\hat{\eta} = F(F^T \Omega^{-1} F)^{-1} F^T \Omega^{-1} y; \quad (5.130)$$

$$\hat{\varepsilon} = y - F(F^T \Omega^{-1} F)^{-1} F^T \Omega^{-1} y; \quad (5.131)$$

$$D\{\hat{\beta}_\Omega\} = \sigma_\varepsilon^2 (F^T \Omega^{-1} F)^{-1}; \quad (5.132)$$

$$D\{\hat{\eta}\} = \sigma_\varepsilon^2 F(F^T \Omega^{-1} F)^{-1} F^T; \quad (5.133)$$

$$s^2 = \frac{1}{n-m} (y - F\hat{\beta})^T \Omega^{-1} (y - F\hat{\beta}). \quad (5.134)$$

Отримані оцінки при недиагональній матриці  $V$  називають оцінками **узагальненого методу найменших квадратів (УМНК)**, або **(оцінками Ейткена)**, а при діагональній матриці – **оцінками зваженого методу найменших квадратів (ЗМНК)**. У відповідності з теоремою Гауса-Маркова вони являються **незміщеними і ефективними** в класі незміщених лінійних оцінок, а при розподілі Гауса збурень співпадають з максимально правдоподібними оцінками і являються ефективними в класі всіх незміщених оцінок.

Для аналізу адекватності моделі, точності і надійності оцінок (5.130)-(5.134) можуть застосовуватись вирази, розглянуті у другому розділі для звичайної моделі, в якій необхідно лише відповідним чином змінити звичайну суму

квадратів на узагальнену. Наприклад, довірча область для всіх параметрів задається нерівністю [пор.з виразом (2.79)]

$$(\hat{\beta}_{\Omega} - \beta)^T Q^T Q (\hat{\beta}_{\Omega} - \beta) \leq \frac{m}{n-m} (z^T z - \hat{\beta}_{\Omega} Q^T z) F_{m, n-m, p} \quad (5.135)$$

а індивідуальні довірчі інтервали для параметрів функції регресії (ФР) мають вигляд [пор. з виразом (2.78) ]

$$\begin{aligned} (\hat{\beta}_{\Omega i} - t_{n-m, (1-p)/2} \sqrt{s^2 v_{ii}} \leq \beta_i \leq \hat{\beta}_{\Omega i} + \\ + t_{n-m, (1-p)/2} \sqrt{s^2 v_{ii}}, \end{aligned} \quad (5.136)$$

де  $s^2$  визначається із (5.134) ;  $v_{ii}$  –  $i$ -й діагональний елемент модифікованої матриці похибок (5.132).

Застосування звичайного, класичного МНК для аналізу моделі з кореляційною матрицею збурень (5.123) приводить до того [15], що отримані оцінки параметрів і функції регресії (ФР) хоч і остаються *спроможними (обґрунтованими) і незміщеними* але перестають бути *ефективними*. При цьому оцінка дисперсії становиться зміщеною , а із-за того, що класична оцінка кореляційної матриці  $\mathbf{D}\{\hat{\beta}\}$  відрізняється від істинної кореляційної матриці похибок оцінювання і його елементи часто занижені значення , статистичний аналіз якості дає, як правило , хибні , «оптимістичні» результати.

Основна трудність застосування розглянутих оцінок закладається в тому, що матриця  $\mathbf{\Omega}$  наперед, як правило, невідома. Тому, оцінки отримують в два етапи [15]: спочатку отримують оцінку кореляційної матриці  $\mathbf{\Omega}$ , а після – оцінки у відповідності з отриманими раніше виразами.

## 5.2.Регресійний аналіз в умовах мультиколінеарності

При рішенні регресійних задач часто стикаються з наявністю сильного лінійного зв'язку між всіма або деякими пояснюючим змінними. Це явище називається мультиколінеарністю [1, 15, 27, 67, 78].

Мультиколінеарність може проявлятися у функціональній (явній) і вірогіднісній (скритій) формі. Функціональна форма виникає, коли по крайній мірі один із факторів зв'язаний з другим лінійними функціональними співвідношеннями. Вірогіднісна форма мультиколінеарності має місце, коли по крайній мірі між двома випадковими пояснюючими змінними існує порівняно сильна кореляція. У всіх випадках, коли виявляється мультиколінеарність необхідний ретельний предметно-змістовний аналіз причин його виникнення.

З математичної точки зору мультиколінеарність проявляється в тому, що інформаційна матриця  $\mathbf{G}=\mathbf{F}^T \mathbf{F}$  стає погано обумовленою , а при строгій мультиколінеарності – виродженою ( $\det \mathbf{F}^T \mathbf{F}=\mathbf{0}$ ), тобто матриця регресорів  $\mathbf{F}$  має неповний ранг ( $r<m$ ). Це приводить, як витікає із результатів (розділи 2 і 3), до різкого зниження точності оцінок параметрів і функцій регресії , які часто перестають задовольняти пред'явленим вимогам. Оцінки стають вкрай чутливими до вибіркового спостережень. Застосування критеріїв значущості становиться ненадійним.

Для встановлення наявності і степеня мультиколінеарності застосовуються різні способи і показники: коефіцієнт множинної детермінації, елементи кореляційної матриці стандартизованих регресорів і т.і. [3, 15, 78]. В лінійній алгебрі для дослідження мультиколінеарності широко застосовуються розглянуті в третьому розділі числа обумовленості матриці регресорів або інформаційної

матриці, які достатньо просто можуть бути розраховані із застосуванням мікро-ЕОМ.

Для боротьби з мультиколінеарністю або з метою пониження степеня її небажаного впливу на результати аналізу в даний час застосовується ряд методів [1, 13, 27, 28, 67, 78]:

- 1) метод центрування факторів [15];
- 2) метод виключення корельованих незалежних змінних; задача відсіву факторів рішається в ході пошуку найкращої структури моделі (див. нижче);
- 3) метод лінійного перетворення змінних. Метод оснований на такому лінійному перетворенні моделі регресії, при якому інформаційна матриця стає ортогональною. Подальшим його розвитком являється метод головних компонент, який фактично також реалізує ідею виключення змінних, але не у вихідному, а в ортогональному базисі;
- 4) алгебраїчні методи, основані на застосуванні для отримання оцінок параметрів регресії у вихідному базисі чисельно стійких методів рішення вихідної системи

$$F\beta = y \Leftrightarrow Ax = b. \quad (5.137)$$

Відмова від використання нормальної системи рівнянь (2.33) зв'язана з тим, що число обумовленості матриці  $F^T F$  суттєво більше числа обумовленості матриці  $F$  вихідної системи ( $\text{cond } F^T F = \text{cond}^2 F$ ). В даному випадку застосовується визначення числа обумовленості матриці як відношення *максимального сингулярного числа* до *мінімального*, яке дозволяє оцінювати степінь виродженості не тільки квадратних, але і прямокутних матриць. Звідси витікає, що перехід від вихідної системи до нормальної погіршує обумовленість задачі і приводить до пониження точності рішення. Крім того, в умовах строгої мультиколінеарності нормальна система не являється визначеною, що не дозволяє застосувати для її рішення стандартні методи.

Розглянемо два алгебраїчні методи, які широко застосовуються на практиці.

### 5.2.1. Метод псевдо-обертання

Попередньо дамо декілька визначень. Нормальним псевдо рішенням СЛАУ (5.137) називається вектор  $\hat{x}_H$ , який забезпечує мінімум норми нев'язки  $\|b - A\hat{x}\|$ , що має мінімальну норму  $\hat{x}$  [7, 73]. Очевидно, що якщо матриця  $A^T A$  не вироджена, то нормальне псевдо рішення співпадає з узагальненим рішенням СЛАУ. При виродженій матриці  $A^T A$  існує нескінченне число узагальнених рішень і нормальне псевдо рішення співпадає з тим із них, яке має найменшу норму.

Нормальне псевдо-рішення формально можна отримати за допомогою псевдо оберненої матриці  $A^+$ :

$$\hat{x}_H = A^+ b. \quad (5.138)$$

Матриця  $A^+$  розміру  $m \times n$  називається псевдо оберненою (*матрицею Мура-Пенроуза, узагальненою оберненою матрицею*) для матриці  $A$  розміру  $n \times m$  якщо вона задовольняє наступним умовам:

$$\begin{aligned} AA^+A &= A; & A^+AA^+ &= A^+; \\ AA^+ &= (AA^+)^T; & A^+A &= (A^+A)^T. \end{aligned} \quad (5.139)$$

Можна показати, що така матриця завжди існує і єдина [7]. Псевдо-обернена матриця має наступні екстремальні властивості [7]: для будь-якої матриці  $X$  розміру  $m \times n$  виконується співвідношення  $\|AA^+ - E\| \leq \|Ax - E\|$ , причому, якщо для будь-якої матриці



$X$ , відмінної від  $A^+$ , має місце рівність, то  $\|A^+\| < \|x\|$ . Якщо  $A$  квадратна не вироджена матриця, то  $A^+ = A^{-1}$ , яка, очевидно, задовольняє умовам (5.139). Якщо матриця  $A$  – прямокутна і має повний ранг ( $r=m$ ), то

$$A^+ = (A^+ A)^{-1} A^T \quad (5.140)$$

і нормальне псевдо-рішення співпадає з рішенням по методу найменших квадратів (2.34). Якщо матриця  $A$  має неповний ранг, то такого простого представлення псевдо-оберненої матриці немає.

В даний час для псевдо-обертання матриць систем, особливо при їх поганій обумовленості, як правило, застосовується метод *сингулярного розкладення* [74, 79]. При застосуванні цього методу, по-перше, не робиться перехід до нормальної системи рівнянь, який, як відмічалось, погіршує обумовленість системи, і, по-друге, отримуються алгоритми, які можуть бути узагальнені і на випадок виродженої СЛАУ.

Спочатку розглянемо псевдо-обернену матрицю для прямокутної матриці розміру  $n \times m$

$$D = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & d_m \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \quad (5.141)$$

Якщо матриця має ранг  $r (r \leq m)$ , то  $r$  її діагональних елементів відмінні від нуля. Як можна перевірити безпосередньою підстановкою у вираз (5.139), псевдо-оберненою для неї буде матриця  $D^+$  розмірів  $m \times n$ :

$$D^+ = \begin{bmatrix} d_1^+ & 0 & \dots & 0 & \dots & 0 \\ 0 & d_2^+ & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_m^+ & \dots & 0 \end{bmatrix}, \quad (5.142)$$

де

$$d_i^+ = \begin{cases} 1/d_i & \text{при } d_i \neq 0; \\ 0 & \text{при } d_i = 0. \end{cases}$$

Відмітимо, що матриця  $DD^+$  представляє собою матрицю порядку  $n$  у якій на головній діагоналі розташовано  $r$  одиниць, а решта елементів являються нулями.

Тепер розглянемо псевдо-обернену матрицю для довільної матриці  $A$ . Так як для матриці  $A$  завжди, як і для виродженої матриці, може бути отримано сингулярне розкладення (див.розділ 3)

$$A = UDV^T,$$

то з врахуванням (5.141), (5.142) псевдо-оберненою для  $A$  являється матриця

$$A^+ = VD^+U^T, \quad (5.143)$$

Яка визначена як для не вироджених, так і для вироджених СЛАУ.

Нормальне псевдо-рішення може бути отримано як в традиційній формі

$$\hat{x} = VD^+U^T b = \sum_{i=1}^r \mu_i^{-1} v_i u_i^T b, \quad (5.144)$$

так і у вигляді рекурентного співвідношення

$$\hat{x}^{(i)} = \hat{x}^{(i-1)} + \mu_i^{-1} v_i u_i^T b \quad (i = \overline{1, r}); \quad (5.145)$$

$$\hat{x}^{(0)} = 0,$$

де  $r$ -ранг матриці  $\mathbf{A}$ ;  $\mu_i$  – її сингулярні числа.

Безпосереднє використання псевдо-обертання дає задовільні результати лише при малому рівні похибок вихідних даних:  $\gamma_A, \gamma_b \rightarrow 0$  (див.розділ 3). В протилежному випадку метод псевдо-обертання неефективний – норма вектора похибки рішення  $\|\hat{x}_T - A^+ b\|$ , як правило, виявляється неприпустимо великою. Детально це питання розглядається в роботах [5, 7, 73].

Для збуреної СЛАУ вводиться ефективна псевдо-обернена матриця  $A_\tau^+$ , яка має структуру псевдо-оберненої матриці (5.143), але для якої елементи матриці (5.142) визначаються по-другому:

$$d_{\tau i}^+ = \begin{cases} 1/d_i & \text{при } d_i > \tau; \\ 0 & \text{при } d_i \leq \tau. \end{cases} \quad (5.146)$$

При цьому враховано, що

$$d_i \geq 0, \text{ тому що } d_i = \mu_i \quad (i = \overline{1, m}),$$

де  $\mu_i$  - сингулярні числа матриці  $\mathbf{A}$ .

Матриця  $A_\tau^+$  задовольняє лише трьом останнім умовам Мура-Пенроуза (5.139), а перше замінюється нерівністю

$$\|AA_\tau^+ A - A\| \leq \tau.$$

Величина  $\tau$  повинна бути узгоджена з рівнем похибок заданих вихідних даних: при зростанні збурень  $\tau$  необхідно збільшувати, при підвищенні точності даних  $\tau \rightarrow 0$ .

Ефективне псевдо-рішення

$$\hat{x}_\tau = A_\tau^+ b = \sum_{i=1}^{r'} \mu_i^{-1} v_i u_i^T b; \quad (5.147)$$

$$\hat{x}_\tau^{(i)} = \hat{x}_\tau^{(i-1)} + \mu_i^{-1} v_i u_i^T b \quad (i = \overline{1, r'}), \quad (5.148)$$

$$\hat{x}_\tau^{(0)} = 0,$$

де  $r'$ - число, яке не перевищує рангу матриці  $r' \leq r$ , і так же, як і  $\tau$ , узгоджене з рівнем похибки вихідних даних.

Очевидно, що при малому рівні збурень

( $\tau \rightarrow 0, r' \rightarrow r$ )  $\hat{x}(\tau) \approx \hat{x}_H \approx \hat{x}_T$ . При зростанні похибок заданих вихідних даних  $\hat{x}(\tau) \neq \hat{x}_H \neq \hat{x}_T$ . Але при

правильному завданні параметра  $\tau$  точність ефективного псевдо-рішення може суттєво перевищити точність нормального псевдо-рішення:  $\|\hat{x} - A_\tau^+ b\| \ll \|\hat{x}_T - A_\tau^+ b\|$ .

Недоліком цього методу являється відсутність конкретних рекомендацій по вибору параметрів  $\tau$  або  $r'$ .

## 5.2.2 Метод регуляризації

Від цього недоліку вільний другий метод, оснований на ідеї регуляризації. В даний час метод регуляризації *А.Н.Тихонова* являється основним при рішенні широкого класу некоректних задач<sup>1</sup>, до числа яких відноситься і задача рішення збурених СЛАУ [5, 7, 15, 44, 45, 73, 74].

У відповідності з методом регуляризації в якості наближеного рішення СЛАУ (5.137) в умовах збурення матриці системи і (або) вектора правих частин слід брати

<sup>1</sup>Задача називається коректною (по АДАМАРУ), якщо, по-перше, рішення існує, по-друге, рішення єдине, по-третє, рішення стійке до похибок вихідних даних [73].

вектор  $\hat{x}(\alpha)$ , який забезпечує мінімум функціонала

$$\Phi_\alpha(\hat{x}, b) = Q\{A\hat{x}, b\} + \alpha F(\hat{x}). \quad (5.149)$$

Функціонал (5.149) називається згладжуючим. Складова  $Q\{A\hat{x}, b\} = \|b - A\hat{x}\|$ , яка входить у (5.149), представляє собою квадрат норми нев'язки рішення  $\hat{x}$ , а  $F(\hat{x})$  звичайно називають *стабілізуючим функціоналом* або стабілізатором і вибирають виходячи із фізичного сенсу задачі. Параметр  $\alpha$  ( $\alpha > 0$ ), який називається параметром регуляризації, повинен бути визначеним чином узгоджений з похибкою завдання вихідних даних.

Відмітимо, що згладжуючий функціонал являється узагальненням функціоналів, які використовувалися раніше. Так, при самостійному використанні функціоналу  $Q\{A\hat{x}, b\}$  пошук його мінімуму приводить до традиційних нерегуляризованих рішень за методом найменших квадратів (2.34) або (5.144). Для виродженої матриці системи раніше було отримано нормальне псевдо-рішення, яке відповідає одночасній мінімізації функціоналу  $Q\{A\hat{x}, b\}$  і стабілізатора окремого виду  $F(\hat{x}) = \|\hat{x}\|^2$ .

Отримане в результаті мінімізації згладжуючого функціоналу (5.149) наближене рішення СЛАУ (5.137)  $\hat{x}(\alpha)$  володіє тією чудовою властивістю, що при узгодженні параметра регуляризації з похибками вихідних даних для нього справедливі дві умови [73]:

- 1)  $\hat{x}_\alpha \rightarrow \hat{x}_T$  при  $\gamma_A, \gamma_b \rightarrow 0$ ;
- 2)  $\|\hat{x}_T - \hat{x}_\alpha\| \leq \|\hat{x}_T - \hat{x}\|$  при  $\forall \hat{x}$ .

Наближене рішення  $\hat{x}_\alpha$  називається регуляризованим рішенням СЛАУ. Воно не являється єдиним в тому сенсі,

що різним стабілізаторам  $F(\hat{x})$  відповідають різні, взагалі говорячи, рішення  $\hat{x}(\alpha)$ . Вибір функціоналу  $F(\hat{x})$  звичайно підказується характером вирішуваної задачі, і, природно, оснований на використанні задалегідь заданої інформації про рішення.

Зокрема, якщо відомий вектор  $\hat{x}_0$ , близький по апіорним даним до точного рішення  $\hat{x}_T$ , то стабілізатор доцільно задати у вигляді

$$F(\hat{x}) = \|\hat{x}_0 - \hat{x}\|^2 = (\hat{x}_0 - \hat{x})^T (\hat{x}_0 - \hat{x})$$

і шукати регуляризоване рішення шляхом мінімізації по  $\alpha$  функціонала

$$\Phi_\alpha(\hat{x}, b) = (b - A\hat{x})^T (b - A\hat{x}) + \alpha(\hat{x}_0 - \hat{x})^T (\hat{x}_0 - \hat{x}). \quad (5.150)$$

Можна показати [5], що в цьому випадку

$$\hat{x}_\alpha = (A^T A + \alpha E)^{-1} (A^T b + \alpha \hat{x}_0). \quad (5.151)$$

Вектор (5.151) називається регуляризованим рішенням СЛАУ, нормальним відносно вектора  $\hat{x}_0$ . Якщо вектор  $\hat{x}_0$  задати не вдається, то можна прийняти його рівним нулю. В результаті отримується рішення

$$\hat{x}_\alpha = (A^T A + \alpha E)^{-1} A^T b. \quad (5.152)$$

Воно мінімізує квадрат норми нев'язки рішення  $\|b - A\hat{x}\|$  при фіксованому квадраті норми рішення  $\|\hat{x}\|^2$ . Вірно і навпаки – рішення (5.152) має мінімальний квадрат норми в класі рішень з заданим значенням норми нев'язки.

Через відмічену раніше обчислювальну неефективність застосування обернених матриць на практиці регуляризоване рішення отримують шляхом рішення СЛАУ

$$(A^T A + \alpha E)x = A^T b. \quad (5.153)$$

Так як матриця  $(A^T A + \alpha E)$  симетрична, для рішення системи (5.153) доцільно застосувати метод квадратного кореня (див.розділ3).

Центральним при використанні методу регуляризації являється питання про оптимальне (узгоджене) значення параметрів регуляризації  $\alpha$ . Для визначення цього значення використовується ряд методів, зокрема, метод нев'язки [73, 74]. Показано, що параметр регуляризації, задовольняючий умову  $\|b - A\hat{x}_\alpha\| \approx \delta_\varepsilon$ , забезпечує отримання рішень, за порядком точності близьким до оптимального. При цьому  $\delta_\varepsilon$  - оцінка рівня похибки вихідних даних:  $\|\varepsilon\| \leq \delta_\varepsilon$ .

Розглянемо більш детально особливості застосування регуляризованих оцінок виду (5.152) для рішення регресійної задачі, яка являється по своїй суті статистичною. Відмітимо, що в регресійному аналізі оцінка (5.152) часто називається *гребневою або рідж-оцінкою*. Можна показати, що гребнева оцінка вектора параметрів лінійної регресії  $\hat{\beta}_\alpha = (F^T F + \alpha E)^{-1} F^T y$  має наступні властивості [5, 15, 27, 28, 67]:

1. Гребнева оцінка являється зміщеною, тому, що при  $\alpha \neq 0$

$$M\{\hat{\beta}_\alpha\} = (E - \alpha(F^T F)^{-1})^{-1} \beta \neq \beta. \quad (5.154)$$

2. Гребнева оцінка- результат лінійного перетворення МНК-оцінки. Дійсно, так як  $F^T y = F^T F\beta$ , то

$$\hat{\beta}_\alpha = (F^T F + \alpha E)^{-1} F^T F\beta = R_\alpha \hat{\beta}, \quad (5.155)$$

де  $R_\alpha = [E + \alpha(F^T F)^{-1}]^{-1}$ , при цьому  $R_\alpha \neq E$  при  $\alpha > 0$ .

Звідси, зокрема, витікає, що при не виродженій матриці

$$F^T F \lim_{\alpha \rightarrow 0} \hat{\beta}_\alpha = \hat{\beta}.$$

3. Для будь-якої матриці регресорів  $F$  існує таке значення параметра регуляризації  $\alpha = \alpha_0 > 0$ , при якому середня сума квадратів похибки оцінки (середній квадрат норми похибки) буде меншим повної дисперсії МНК-оцінки:

$$L\{\hat{\beta}_{\alpha_0}\} \leq L\{\hat{\beta}\}, \quad (5.156)$$

де

$$L\{\hat{\beta}\} \equiv tr\{D\{\hat{\beta}\}\} = \sigma_\varepsilon^2 tr\{(F^T F)^{-1}\};$$

$$L\{\hat{\beta}_{\alpha_0}\} = tr\{D\{\hat{\beta}_{\alpha_0}\}\} + (R_{\alpha_0} x_T - x_T).$$

Дійсно, при поганій обумовленості матриці  $(F^T F)$  її мінімальне власне число прямує до нуля, тому величина

$$L\{\hat{\beta}\} \equiv \sigma_\varepsilon^2 tr\{(F^T F)^{-1}\} = \sigma_\varepsilon^2 \sum_{i=1}^m 1/\lambda_i$$

приймає вкрай великі значення. В той же час середній квадрат норми похибки гребневої оцінки на відмінність від незміщеної МНК-оцінки вміщає дві складові: повну дисперсію гребневої оцінки, яка монотонно убуває з ростом  $\alpha$ , і повний квадрат зміщення, який при збільшенні параметра регуляризації  $\alpha$  монотонно росте. Звідси витікає, що при деякому значенні  $\alpha = \alpha_0$  середня сума квадратів  $L\{\hat{\beta}_\alpha\}$  буде мати мінімум. Другими словами, із виразу (5.156) витікає, що допускаючи невелике зміщення оцінки, можна зменшити дисперсію регуляризованої оцінки настільки, що при цьому зменшиться і повний середній квадрат похибки.

При використанні гребневих оцінок для вибору оптимального значення  $\alpha_0$  також широко застосовується метод нев'язки. У найпростішому випадку  $\alpha$  підбирають за умови

$$\sigma_{\varepsilon H}^2 \leq \|y - F\hat{\beta}_\alpha\|^2 \leq \sigma_{\varepsilon B}^2, \quad (5.157)$$

де  $\sigma_{\varepsilon H}^2$  і  $\sigma_{\varepsilon B}^2$  – оцінки нижньої і верхньої межі дисперсії збурень. Більш коректним являється статистичний підхід, коли поступово зменшують, починаючи з  $\alpha \gg 0$ , перевіряючи кожний раз гіпотезу про рівність векторів  $y$  і  $F\hat{\beta}_\alpha$ . Якщо гіпотеза при будь-якому  $\alpha = \hat{\alpha}_0$  приймається, то отримана оцінка близька до оптимальної. При корельованих збуреннях ( $D\{\varepsilon\} = \sigma_\varepsilon^2 \Omega$ , де  $\Omega$  – невід’ємно визначена матриця з одиничною головною діагоналлю) гребенева оцінка прийме вигляд:

$$\hat{\beta}_\alpha = (F^T \Omega^{-1} F + \alpha E)^{-1} F^T \Omega^{-1} y. \quad (5.158)$$

Гребневі оцінки більш загального вигляду, а також їх зв’язок з другими відомими методами рішення регресійних задач розглянуті в роботах [5, 44, 45].

### 5. 3. Вибір найкращої структури регресійної моделі

Однією із основних задач регресійного аналізу являється обґрунтоване рішення про те, які фактори або регресори слід включити в модель. Вище відмічалось, що небажаним виявляється як недобір, так і перебір факторів. Обчислювальні проблеми, які обмірковувалися вище, викликані явищем мультиколінеарності, багато в чому зв’язані з наявністю в моделі корельованих або лінійно залежних змінних. Обґрунтоване спрощення моделі дає можливість понизити обчислювальні затрати, підвищити оперативність опрацювання матеріалів, підвищити прогностичні можливості моделі. Зменшення числа регресорів збільшує число степенів свободи  $\nu = n - m$

остаточної дисперсії і робить зв’язані з нею перевірки гіпотез більш надійними. Розглянемо деякі найбільш поширені методи пошуку найкращих регресійних моделей. (Більш детально ці питання обмірковуються в роботах [1, 3, 10, 15, 27, 28, 32, 67].)

#### 5.3.1. Побудова і перевірка всіх можливих регресій

Очевидний підхід до побудови оптимальної моделі полягає в побудові і статистичній перевірці всіх можливих рівнянь регресії, які можна отримати, вибираючи по  $i$  ( $i = \overline{0, m}$ ) регресорів. Так як для кожного регресора є дві можливості – або він включається в модель, або не включається, - то всього маємо  $2^m$  можливих рівнянь регресії. Кожна модель перевіряється на адекватність, наприклад, з використанням остаточної дисперсії (2.76) коефіцієнта множинної детермінації, статистики Меллоуза або інших. Звичайно цей метод використовується при  $m \leq 10 \dots 15$ . Для скорочення числа перевірок поступають наступним чином.

Всі моделі розбиваються на  $m$  підмножин, які мають 1, 2, ...,  $m$  регресорів. Всередині кожної підмножини моделі впорядковуються, наприклад, по зростанню коефіцієнта множинної детермінації (квадрат коефіцієнта кореляції змінних  $Y$  і  $X$ ):

$$R^2 = 1 - [(y - F\hat{\beta})^T (y - F\hat{\beta})] / [(y - \hat{y})^T (y - \hat{y})] \equiv 1 - Q_0 / Q,$$

де  $Q_0$ - остаточної сума квадратів;  $Q$ -повна сума квадратів відхилень від середнього. Далі розглядаються «лідери» різних підмножин і робиться спроба виявити закономірності в появі регресорів. Розгляд починається з моделей, що мають

мале число регресорів. Якщо при ускладненні моделі немає помітного приросту коефіцієнта детермінації, то введення нового регресора мало доцільно. Для прийняття рішення про обґрунтованість моделі з  $p$  регресорами ( $p < m$ ) обчислюються і порівнюються коефіцієнти детермінації для даної і для самої складної моделі з  $m$  регресорами:

$R_p^2$  і  $R_m^2$ . Відмінності між моделями з  $p$  і  $m$  регресорами вважаються не значимими, якщо виконується нерівність

$$\frac{(n-m)(R_m^2 - R_p^2)}{1 - R_m^2} \leq (m-1)F_{(m-1), (n-1), \alpha}, \quad (5.159)$$

де  $F_{v_1/v_2q}$  – квантиль порядку  $q$  F-розподілу з  $v_1, v_2$  степенями свободи.

Аналогічна перевірка може бути виконана за допомогою остаточної дисперсії  $s^2$ , остаточної суми квадратів  $Q_0$  статистики Меллоуза або окремого коефіцієнта кореляції. Пошук найкращої (адекватної) моделі закінчується, як тільки відмінності між деякою спрощеною і повною моделями будуть не значимими.

### 5.3.2. Метод виключення

При великому числі факторів і регресорів аналіз всіх можливих регресій стає утрудненим. Тому необхідний перехід до більш економічних процедур пошуку структури адекватної регресійної моделі. На практиці використовується декілька стратегій пошуку: пошук вздовж перспективних гілок, метод включення і др. [15, 32, 67]. Але найбільш широко застосовується метод виключення і розглянутий далі кроковий метод. Розрахунки методом виключення починають із звернення до повної регресійної моделі. Після цього обчислюють

частинні F-критерії для оцінки значимості кожного із отриманих коефіцієнтів регресії. Частинний критерій для  $i$ -го регресора визначається в даному випадку наступним чином:

$$F_j = \frac{(Q_{0i} - Q_0)(n - m')}{Q_0}, \quad (5.160)$$

де  $Q_0$  – остаточної сума квадратів для редукованої моделі;  $Q_{0i}$  остаточної сума квадратів для редукованої моделі без  $i$ -го регресора;  $m'$  – число регресорів в редукованій моделі (на першому кроці виключення  $m' = m$ ).

Після отримання значень критерію для всіх регресорів визначається «кандидат» на виключення – регресор з найменшим значенням F-критерію. Це значення  $F_{\min}$  порівнюється з критичним значенням F-розподілу з  $1, n - m'$  степенями свободи при заданому рівні значимості  $\alpha$ . Якщо  $F_{\min} < F_{1, n - m', 1 - \alpha}$ , то даний регресор виключається із моделі, у протилежному випадку процедура виключення закінчується. Після виключення регресора коефіцієнти для редукованої регресійної моделі розраховуються заново, виконується перевірка гіпотези про адекватність нової моделі і при позитивному рішенні процедура виключення повторюється.

### 5.3.3 Метод крокової регресії

В розрахунках цим методом використовується як процедура виключення, так і включення регресорів. Процедура починається з розгляду самої простої регресійної моделі, зокрема що містить один регресор. В подальшому розглядаються всі регресори, які ще не включені в модель, і для них розраховуються частинні F-

критерії. Вибирається регресор з самим великим значенням критерію  $F_{\max}$ , яке порівнюється з критичним значенням  $F_E$ . Для включення в модель. Якщо  $F_{\max} > F_E$ , то відповідний регресор включається в модель. В протилежному випадку процедура завершається. Після включення нового регресора коефіцієнти моделі оцінюються заново і перевіряється гіпотеза про адекватність моделі. В подальшому виконується процедура виключення регресора, який стає незначимим після включення в модель нового регресора: визначається регресор з мінімальним значенням частинного критерію  $F_{\min}$ , який порівнюється з критичним значенням  $F_0$  для виключення із моделі (див. вище). Якщо  $F_{\min} < F_0$ , то даний регресор виключається, в протилежному випадку модель залишається без змін і переходять до пошуку «кандидата» на нове включення. Кроковий алгоритм завершається, якщо не вдається виконати ні нового виключення, ні нового включення регресорів.

На закінчення відмітимо, що розглянуті в даному пункті методи фактично являються засобом для формування хорошої обумовленої матриці рівняння регресії і доповнюють розглянуті раніше алгебраїчні методи рішення регресійної задачі при заданій матриці  $F$ . В той же час очевидні і визначені аналогії між деякими методами рішення погано обумовлених СЛАУ, наприклад, методом псевдо-обертання або методом головних компонент і методом виключення регресорів.

#### 5.4. Регресійний аналіз в умовах похибок в регресорах

В деяких практичних випадках фактори або регресори задаються (вимірюються) з похибками (див. розд. 1). В цих умовах класичні МНК-оцінки параметрів і функції регресії

втрачають свої оптимальні властивості, що необхідно враховувати при їх застосуванні або при пошуку альтернативних алгоритмів оцінювання. Обмежимося коротким розглядом особливостей регресійного аналізу в умовах активного експерименту. Більш повно це питання досліджується в роботах [1, 15, 27, 32, 42, 67], де також обмірковуються особливості опрацювання результатів пасивного експерименту і розглядаються спеціальні алгоритми регресійного аналізу.

В умовах активного експерименту, характерного для наукових досліджень в техніці, значення факторів ціле направлено задаються, але через похибки управління фактично на досліджуваній об'єкт взаємодіють відмінні від них значення.

$$x_{\delta i j} = x_{ij} + \delta_{ij}, \quad (5.161)$$

де  $\delta_{ij}$  - похибка заданого  $j$ -фактору в  $j$ -експерименті.

В цих умовах рівняння регресії можна записати в наступному вигляді:

$$y = F_{\delta} \beta + v, \quad (5.162)$$

де  $y$  – вектор значень результуючої змінної, що спостерігаються;  $\beta$  – вектор оцінюваних параметрів;  $F_{\delta}$  – фактична матриця регресорів;  $v$  – вектор збурень.

Матриця регресорів  $F_{\delta}$  представляється наступним чином

$$F_{\delta} = \begin{bmatrix} f_1(x_{\delta 1}) \dots f_m(x_{\delta 1}) \\ \dots \dots \dots \\ f_1(x_{\delta n}) \dots f_m(x_{\delta n}) \end{bmatrix}, \quad (5.163)$$

де  $f_i(x)$  – відомі функції;  $x_{\delta i} = x_i + \delta_i$ .

Фактична матриця невідома, тому при оцінюванні параметрів застосовується «запланована» матриця регресорів  $F$ , яка визначається заданими значеннями

факторів. Рівняння регресії, застосоване при формуванні оцінки має вигляд

$$y = F\beta + \varepsilon, \quad (5.164)$$

де  $\varepsilon$ -вектор відхилень, визваних як збуренням моделі, (5.162), так і відмінністю матриць  $F_\delta$  і  $F$ :

$$\varepsilon = v - \Delta F\beta. \quad (5.165)$$

При цьому  $\Delta F = F_\delta - F$  – матриця, що характеризує відхилення фактичної і розрахункової матриці регресорів. В загальному випадку  $\Delta F$  залежить не тільки від вектора збурень факторів  $\delta = [\delta_1, \dots, \delta_k]^T$ , але і від значень вектора факторів  $x$ . Тільки для лінійних функцій  $f_i(x)$  матриця  $\Delta F$  буде функцією вектора  $\delta$ . Наприклад, якщо застосовується модель виду (1.6), то

$$F = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}; \quad F_\delta = \begin{bmatrix} x_1 + \delta_1 \\ \dots \\ x_n + \delta_n \end{bmatrix}; \quad \Delta F = \begin{bmatrix} \delta_1 \\ \dots \\ \delta_n \end{bmatrix}. \quad (5.166)$$

Модель (5.164) представляє собою рівняння регресії з відомою матрицею плану, але реальні властивості вектора відхилень  $\varepsilon$  не будуть відповідати базовим передумовам, описаним у розділі 2, навіть в тому випадку, якщо фактичний вектор відхилень  $v$  буде цим передумовам задовольняти ( $M\{v\} = 0$ ;  $D\{v\} = \sigma_v^2 E$ ).

Дійсно, можна показати [1, 15, 67], що математичне сподівання вектора  $\varepsilon$

$$M\{\varepsilon\} = -M\{\Delta F\}\beta, \quad (5.167)$$

А його кореляційна матриця

$$D\{\varepsilon\} = \sigma_v^2 E + M\{[\Delta F - M\{\Delta F\}]\beta\beta^T[\Delta F - M\{\Delta F\}]\}. \quad (5.168)$$

Звідси витікає, що оцінки параметрів регресії фактичної моделі (5.162), отримані звичайним методом найменших квадратів для моделі (5.164) з застосуванням виразу (2.34),

являються зміщеними (через зміщення векторів  $\varepsilon$ ) і неефективними (через корельованість елементів цього вектора). Незміщеними, однак, і також неефективними класичні оцінки будуть тільки при лінійних функціях  $f_i(x)$ , наприклад для моделі (1.6) при умові, що  $M\{\delta_i\} = 0$ . Безпосереднє використання для оцінки параметрів регресії узагальненого МНК, пристосованого для роботи в умовах корельованих відхилень, неможливо, тому що кореляційна матриця  $D\{\varepsilon\}$  залежить від оцінюваних параметрів [див.(5.168)] і в зв'язку з цим невідома. Можливий шлях рішення цієї проблеми – використання ітераційних процедур, на кожному кроці яких оцінюється вектор параметрів  $\beta$  і кореляційна матриця збурень. Відмітимо, що в тих випадках, коли похибки задавання факторів невеликі, пониження ефективності класичних МНК-оцінок на практиці можна нехтувати або використовувати для отримання оцінок розглянуті вище методи рішення рівняння регресії, стійкі до похибок вихідних даних

## Розділ 6. Результати експериментальних досліджень

### 6.1. Побудова математичної моделі класичним МНК

Побудуємо математичну модель залежності купівельної спроможності від середньої заробітної плати, корупції і інфляції за даними, приведеними в табл. 6.1 поліномом другого порядку виду  $y = ax^2 + bx + c$  [52].

При знаходженні емпіричної формули, як правило, вибирають лише частину вітки параболи, яка кращим чином підходить до даних експерименту.

**Таблиця 6.1. Купівельна спроможність У, заробітна плата Х**



№п/п	Роки	Куп.спром. КСП(У)	Сер.з.плата СЗП(Х)
1	1998	17,68866	152,83
2	1999	20,11089	177,39
3	2000	21,60697	231,04
4	2001	37,17733	311,62
5	2002	49,76956	375,98
6	2003	44,73678	372,72
7	2004	59,83743	524,14
8	2005	90,11909	735,57
9	2006	115,5926	928,81
10	2007	140,3828	1194,91
11	2008	171,5988	1573,99
12	2009	188,4182	1650,43
13	2010	239,1131	<b>1982,63</b>

Таблиця 6.2. Розрахункова таблиця в MS EXCEL

Номер строчки excel →	N...СЗП= Х	О КСП1=У	Р	ХiУi	Q	Х^2
3	152,83	17,68865741	2703,357512	23357,0089		
4	177,39	20,11089319	3567,555139	31468,69037		
5	231,04	21,60697029	4992,14644	53381,02188		
6	311,62	37,17733052	11585,07581	97104,94694		
7	375,98	49,76956034	18712,31782	141360,3338		
8	372,72	44,73677893	16674,29224	138920,1984		
9	524,14	59,83743179	31363,1915	274722,7396		
10	735,57	90,11908554	66288,89575	541063,2249		
11	928,81	115,5926424	107363,6022	862688,0161		
12	1194,91	140,3827628	167744,767	1427809,908		
13	1573,99	171,5988008	270094,7964	2477444,52		
14	1650,43	188,4181565	310970,978	2723919,185		
15	<b>1982,63</b>	239,1130783	474072,7624	3930821,717		

Суми 16 | 10212,06 | 1196,152149 | 1486133,738 | 12724061,51

Продовження таблиці 6.2.

→	R	Y^2	S	X^3	T	X^4	U	YX^2
3	312,888600	9	3569651,67	545549864,8	413154,1285			
4	404,448025	466,861165	5582362,104	990278473,4	632863,4709			
5	466,861165	2	12333329,23	2849533497	1153402,154			
6	1382,15390	4	30259519,88	9429370721	3610102,707			
7	2477,00913	7	53148540,49	19982743963	7035441,662			
8	2001,37938	9	51778336,35	19298821523	6214842,205			
9	3580,51824	3	143993176,7	75472583653	16438703,19			
10	8121,44957	9	397989876,3	2,92749E+11	48760123,05			
11	13361,6589	7	801273256,2	7,44231E+11	99720387,33			
12	19707,3200	8	1706104337	2,03864E+12	200439899,6			
13	29446,1484	2	3899472900	6,13773E+12	425126508,6			
14	35501,4017	4495637940	7,41974E+12	513235831,3				
15	57175,0642	1	7793365061	1,54514E+13	939910881			
16	173938,301	4	19394508287	3,2213E+13	2262692140			

Робочі формули (3.12), (3.13) [52]

$$\begin{aligned}
A &= n[x^2] - [x][x] = 61126562,117, \\
\text{при цьому комп'ютерна формула} &= M15 * Q16 - N16^2, \\
B &= [x][x^2] - n[x^3] = -1,2219E + 11, \\
\text{при цьому комп'ютерна формула} &= N16 * Q16 - M15 * S16, \\
C &= [x][x^3] - [x^2][x^2] = 3,61562E + 13, \\
\text{при цьому комп'ютерна формула} &= N16 * S16 - Q16 * Q16, \\
D &= [x^2][x^4] - [x^3][x^3] = 3,37335E + 19, \\
\text{при цьому комп'ютерна формула} &= Q16 * T16 - S16 * S16, \\
E &= [x^2][x^3] - [x][x^4] = -8,21844E + 16, \\
\text{комп'ютерна формула} &= Q16 * S16 - N16 * T16, \\
F &= n[x^4] - [x^2][x^2] = 2,56867E + 14, \\
\text{комп'ютерна формула} &= M15 * T16 - Q16 * Q16.
\end{aligned}
\tag{6.169}$$

$$\begin{aligned}
a &= ([x^2y]A + [xy]B + [y]C) / (nD + [x]E + [x^2]C) = -5,27679E-07 \\
b &= ([x^2y]B + [xy]F + [y]E) / (nD + [x]E + [x^2]C) = 0,117281813 \\
c &= ([x^2y]C + [xy]E + [y]D) / (nD + [x]E + [x^2]C) = 0,117281813
\end{aligned}$$

При цьому комп'ютерні формули будуть

$$a = (U16 * O19 + P16 * P20 + O16 * P21) / (M15 * P22 + N16 * P23 + Q16 * P21),$$

$$b = (U16 * P20 + P16 * P24 + O16 * P23) / (M15 * P22 + N16 * P23 + Q16 * P21),$$

$$c = (U16 * P21 + P16 * P23 + O16 * P22) / (M15 * P22 + N16 * P23 + Q16 * P21).$$

Таким чином на основі проведених розрахунків отримана формула

$$y = -5,27679E - 07x^2 + 0,117281813x + 0,398235929. \tag{6.170}$$

Виконаємо контроль зрівноваження за формулою

$$[y^2] - a[yx^2] - b[yx] - c[y] = [\varepsilon\varepsilon]. \tag{6.171}$$

В нашому випадку

Контроль зрівноваження [y^2]-a[yx^2]-b[yx]- c[y]=[VV]	359,46663
---	-----------

В табл. 6.3 [εε] = 359,4666293

**Таблиця 6.3. Результати зрівноваження**

→	O КСП1=У	Z Y"зрівн	AA ε=V=Y"-YV^2=εε	AB
3	17,6886574	1 18,31009	0,621432971	0,386178937
4	20,1108931	9 21,18674	1,07584682	1,15744638
5	21,6069702	9 27,46725	5,860278567	34,34286488
6	37,1773305	2 36,89396	-0,283367254	0,080297001
7	49,7695603	4 44,41916	-5,350399038	28,62676987
8	44,7367789	44,03821	-0,698570993	0,488001433

	3			
9	59,8374317	9	61,72536	1,887928113
10	90,1190855	4	86,38171	-3,737374209
11	115,592642	4	108,8755	-6,717108147
12	140,382762	8	139,786	-0,596741026
13	171,598800	8	183,6913	12,09254054
14	188,418156	5	192,5263	4,108147109
15	239,113078	3	230,8505	-8,262613455
16	1196,15214	9	Σ	2,23821E-13
				<b>359,4666293</b>

Другим контролем процедури зрівноваження був розрахунок за формулою

$$= \text{ЛИНЕЙН}(O3:O15; X3:Y15; 1; 1), \quad (6.172)$$

При цьому виділяється необхідний масив і натиском клавіш F2, Ctrl+Shift+Enter отримується результат

a	b	c	F(0,05;2;10)=	4,102821015
-5,27679E-07	0,117282	0,398235929	a,b,c	Коефіц.
6,08638E-06	0,012477	4,521413692	стандарт S	ai=S√dii
0,994372634	5,995554	#Н/Д	R^2	μ
883,5150814	10	#Н/Д	Fкритерій	n-m-1
63518,83765	359,4666	#Н/Д	[(Y'-Ycp)^2]	[VV]
0,086698357	9,400095	0,088077747	t(0,05;10)=	2,228138842

Першим масивом набирається вектор Y, який знаходиться в діапазоні O(3:O15)

17,68865741
20,11089319
21,60697029
37,17733052
49,76956034
44,73677893
59,83743179
90,11908554
115,5926424
140,3827628
171,5988008
188,4181565
239,1130783

другим масивом набирається матриця коефіцієнтів початкових рівнянь X(X3:Y15)

СЗП=X	X^2
152,83	23357,01
177,39	31468,69
231,04	53381,02
311,62	97104,95
375,98	141360,3
372,72	138920,2
524,14	274722,7
735,57	541063,2
928,81	862688
1194,91	1427810
1573,99	2477445
1650,43	2723919
<b>1982,63</b>	<b>3930822</b>

Порівняння результатів опрацювання експериментальних даних двома різними методами встановлює їх повну автентичність.

Встановлений коефіцієнт детермінації  $R^2 = 0.994$  говорить про високу адекватність побудованої математичної моделі експериментальним даним.

### Аналіз моделі за критерієм Фішера-Снедекора

$F_{критерій} = 883,515,$

$F(0.05; 2; 10) = 4.102,$

$F > F(0.05; 2; 10).$

*Модель адекватна експериментальним даним.*

### Аналіз моделі за критерієм Стьюдента

0,086698357	9,400095	0,088077747	$t(0,05; 10) =$	2,228138842
t(a)	t(b)	t(c)		
Статистично незначимі коефіцієнти "a" і "c"				

Статистично значимим є коефіцієнт b

$t(b) = 9,4$  при  $t(0.05; 10) = 2.28$ , тобто  $t(b) > t(0.05; 10).$

### 6.2. Оцінка точності зрівноваженої моделі

Необхідно виконати оцінку точності визначених коефіцієнтів a, b, c в побудованій нами математичній моделі в розділі 3.

Знайдемо середню квадратичну похибку одиниці ваги

$$\mu = \sqrt{\frac{[\varepsilon\varepsilon]}{n-3}} = \sqrt{\frac{359,4666293}{10}} = 5,996..$$

Тобто, на основі проведених нами теоретичних і експериментальних досліджень, середня квадратична похибка одиниці ваги складає шість не корумпованих гривень.

Розрахуємо ваги коефіцієнтів

$$S = n([x^2][x^4] - [x^3][x^3]) + [x]([x^2][x^3] - [x][x^4]) + [x^2]([x][x^3] - [x^2][x^2]) = 5,93159E+19.$$

При цьому комп'ютерна формула розрахунку має вигляд

$$= M15 * P22 + N16 * P23 + Q16 * P21.$$

Тобто

$$S = nD + [x]E + [x^2]c = \Delta$$

i є визначник системи нормальних рівнянь.

Вага (ступінь довір'я до функції) встановленого нами коефіцієнта a при  $x^2$  буде

$$P_a = \frac{S}{n[x^2] - [x][x]} = \frac{5,93159E+19}{61126562,117} = 9,70378E+11.$$

Комп'ютерна формула розрахунку ваги коефіцієнта «a» має вигляд

=T19/O19.

Вага коефіцієнта «в» розраховується за формулою

$$P_b = \frac{S}{n[x^4] - [x^2][x^2]} = \frac{S}{F}.$$

І в нашому випадку

$$P_b = S/F = 230920,1098.$$

Вага коефіцієнта «с»

$$P_c = \frac{S}{n[x^4][x^2] - [x^3][x^3]} = \frac{S}{D}.$$

І в нашому випадку

$$P_c = S/D = 1,758369286.$$

Середні квадратичні похибки зрівноважених коефіцієнтів будуть

$$m_a = \mu \sqrt{\frac{1}{P_a}} = 5,995553597 \sqrt{\frac{1}{9,70378E+11}} = 6,08638E-06.$$

$$m_b = \mu \sqrt{\frac{1}{P_b}} = \frac{5,995553597}{\sqrt{230920,1098}} = 0,012476662.,$$

$$m_c = \mu \sqrt{\frac{1}{P_c}} = \frac{5,995553597}{\sqrt{1,758369286}} = 4,521413692..$$

Таким чином, для поліному

$$y = ax^2 + bx + c, \text{ тобто}$$

$$y = -5,27679E-07x^2 + 0,117281813x + 0,398235929$$

середні квадратичні похибки отриманих коефіцієнтів будуть  $m_a = 6,086E-06$ ,  $m_b = 0.0125$ ,  $m_c = 4.521$ .

Контрольні визначення функцією ЛИНЕЙН дають повністю автентичні результати

-5,27679E-07	0,117282	0,398235929
6,08638E-06	0,012477	4,521413692
0,994372634	5,995554	#Н/Д
883,5150814	10	#Н/Д
63518,83765	359,4666	#Н/Д

У програмованому мікрокалькуляторі CITIZEN SRP-350 приведена програма [5x<sup>2</sup>], яка повністю реалізує дану проблему

Підрахувати середню квадратичну похибку зрівноваженої функції

$$y = ax^2 + bx + c, \text{ або}$$

$$y = -5,27679E-07x^2 + 0,117281813x + 0,398235929$$

для всіх випадків рівняння апроксимуючої функції.

### Реалізація

Середню квадратичну похибку функції будемо розраховувати за формулою

$$m_{\varphi} = \sqrt{m_a^2 x^4 + m_b^2 x^2 + m_c^2 + 2 \frac{[\varepsilon\varepsilon] E}{n-3 S}}$$

Де

$$E = ([x][x^2] - n[x^3])x^3 + ([x][x^3] - [x^2][x^2])x^2 + ([x^2][x^3] - [x][x^4])x^3$$

$$S = n([x^2][x^4] - [x^3][x^3]) + [x]([x^2][x^3] - [x][x^4]) + [x^2]([x][x^3] - [x^2][x^2])$$

Раніше, в [52,п.4.2] нами були встановлені середні квадратичні похибки зрівноважених коефіцієнтів

$$m_a = 6,086E - 0.6, \quad m_b = 0.0125, \quad m_c = 4.521.$$

Там же був визначений параметр  $S = 5,93159E+19$ , який є не чим іншим як визначником системи нормальних рівнянь.

**Таблиця 6.4. Розрахунок середньої квадратичної похибки зрівноваженої функції**

Роки	Середня заробітн. плата Х(грн.)	Купівельна спроможність У(грн.) з врах. корупції і інфляції	Похибка моделі $m_{\varphi}$ (грн.)
1998	152,83	17,68866	3,06115
1999	177,39	20,11089	2,87409
2000	231,04	21,60697	2,52125
2001	311,62	37,17733	2,15455
2002	375,98	49,76956	2,01506
2003	372,72	44,73678	2,01889
2004	524,14	59,83743	2,12322
2005	735,57	90,11909	2,64795
2006	928,81	115,5926	2,98706
2007	1194,91	140,3828	3,01623
2008	1573,99	171,5988	2,78200
2009	1650,43	188,4182	2,91437
2010	<b>1982,63</b>	239,1131	5,02999
n=13;Σ			

Комп'ютерна формула для розрахунку зрівноваженої функції купівельної спроможності громадян України буде

$$Y = (\$U\$23 * T3 + \$U\$24 * Q3 + \$U\$25 + 2 * (\$AB\$21^2)) * (\$P\$20 * S3 + \$P\$21 * Q3 + \$P\$23 * N3) / (\$T\$19)^{0,5}. \quad (6.173)$$

Як видно із табл. 6.1, при середній заробітній платі у 1998 році в 152,83 гривні, реальна купівельна спроможність громадян України складала всього 17,68 гривні в місяць.

При цьому середня квадратична похибка побудованої математичної моделі становила 3,06 гривні.

У 2007 році при середній заробітній платі в 1194,91 гривні, реальна купівельна спроможність громадян України складала всього 140,38 гривні в місяць.

При цьому середня квадратична похибка побудованої математичної моделі становила 3,02 гривні.

А в 2010 році при середній заробітній платі в 1982,63 гривні, реальна купівельна спроможність громадян України складала всього 239,11 гривні в місяць.

При цьому середня квадратична похибка побудованої математичної моделі становила 5,03 гривні.

В даному випадку, не корумповані 239,11 гривні 2010 року дуже близькі до корумпованих 231,04 гривні 2000 року.

І, як в кожному серйозному академічному дослідженні, нам залишається лише привести контроль розрахунку середньої квадратичної похибки зрівноваженої функції.

При цьому матриця коефіцієнтів початкових рівнянь  $X$  має вигляд

1	152,83	23357,0089
1	177,39	31468,6904
1	231,04	53381,0219
1	311,62	97104,9469
1	375,98	141360,334
1	372,72	138920,198
1	524,14	274722,74
1	735,57	541063,225
1	928,81	862688,016
1	1194,91	1427809,91
1	1573,99	2477444,52
1	1650,43	2723919,18

1	1982,63	3930821,72
---	---------	------------

Матриця коефіцієнтів нормальних рівнянь  $N=X^T X$

13	10212,06333	12724061,5
10212,0633	12724061,51	1,9395E+10
12724061,5	19394508287	3,2213E+13

Комп'ютерна формула встановлення матриці коефіцієнтів нормальних рівнянь

=МУМНОЖ((ТРАНСП(АF3:АН15));АF3:АН15) (6.174)

Обернена матриця  $Q=N^{-1}$

0,56870875	-0,00138554	6,0955E-07
-0,0013855	4,3305E-06	-2,06E-09
6,0955E-07	-2,06E-09	1,0305E-12

Комп'ютерна формула встановлення оберненої матриці

=МОБР(АF22:АН24) (6.175)

Допоміжна матриця  $Q'=XQ$

0,37119417	-0,00077182	3,188E-07
0,34210406	-0,00068216	2,7655E-07
0,28112779	-0,00049497	1,8862E-07
0,19614237	-0,00023612	6,7698E-08
0,13394164	-4,856E-05	-1,9282E-08
0,13696995	-5,7648E-05	-1,5082E-08
0,00995057	0,000318326	-1,8706E-07
-0,1206451	0,000685267	-3,4813E-07
-0,1923391	0,000859552	-4,1476E-07
-0,2165589	0,000847757	-3,8055E-07

-0,1019804	0,000327134	-7,9767E-08
-0,0576514	0,000150424	1,6766E-08
0,2177444	-0,00089718	5,7618E-07

0,236281789
0,703842181

Комп'ютерна формула розрахунку допоміжної матриці Q'

$$=МУМНОЖ(АF3:АН15;АF28:АН30). \quad (6.176)$$

Обернена вага зрівноваженої функції обчислюється як добуток двох векторів построчно

$$\frac{1}{Py'} = X_{1 \times m} Q'_{m \times 1} TP. \quad (6.177)$$

Комп'ютерна формула розрахунку обернених ваг зрівноважених функцій

$$=МУМНОЖ(АF3:АН3;(ТРАНСП(АF33:АН33))) \quad (6.178)$$

Вектор обернених ваг зрівноваженої функції

1/P(y)
0,260682516
0,229795961
0,176837063
0,129138161
0,112958291
0,113388293
0,125409204
0,195057454
0,248215589
0,253087977
0,21530552

1/P(y)^0,5
0,510571
0,47937
0,42052
0,359358
0,336093
0,336732
0,354132
0,441653
0,498212
0,503079
0,46401
0,486088
0,838953

Розраховуються середні квадратичні похибки зрівноваженої функції зрівноваженої функції

$$\dots \dots \dots m_{y'} = \mu_0 \sqrt{\frac{1}{P_{y'}}}. \quad (6.179)$$

Середня квадратична похибка (стандарт) одиниці ваги  $\mu$  (мю)

$$\mu = \sqrt{\frac{\sum_{i=1}^n V_{1 \times n} V_{n \times 1}}{n - m - 1}}, \quad (6.180)$$



де n- число років, у нас-13, m-ступінь поліному, у нас -2;  
тобто

$$n-m-1=10.$$

Вектор середніх квадратичних похибок зрівноваженої функції буде (рядом приведені дані попередніх розрахунків)

m(y)контр.	m(Y)
3,061154	3,06115
2,874091	2,87409
2,52125	2,52125
2,15455	2,15455
2,015062	2,01506
2,018893	2,01889
2,123215	2,12322
2,647955	2,64795
2,987059	2,98706
3,016234	3,01623
2,781998	2,78200
2,914368	2,91437
5,029988	5,02999

Представляє інтерес побудови математичної моделі купівельної спроможності українців поліномом третього степеня.

Дану проблему вдалось вирішити за допомогою функції ЛИНЕЙН. Комп'ютерна формула при цьому

$$=ЛИНЕЙН(O3:O15;A3:A15;1;1) \quad (6.181)$$

При цьому виділяється необхідний масив і натиском клавіш F2, Ctrl+Shift+Enter отримується результат

Апроксимація кубічним поліномом

Функція ЛИНЕЙН (по X;X^2;X^3)					
a	b	c	d	F(0,05;3;9)	3,86254
2,632E-08	-8,248E-05	0,186049	-12,68148	a,b,c,d	Коефіц.
9,026E-09	2,8472E-05	0,025391	5,637707	стандарт S	$\sigma_i = S \sqrt{d_{ii}}$
0,997107	4,53114	#Н/Д	#Н/Д	R^2	$\mu$
1034,089	9	#Н/Д	#Н/Д	Fкритерій	n-m-1
63693,52	184,7818	#Н/Д	#Н/Д	$[(Y'-Y_{cp})^2]$	[VV]
2,9168901	-2,897106	7,327291	-2,24940	t(0,05;9)=	2,262157

Першим масивом набирається вектор Y, який знаходиться в діапазоні O(3:O15)

17,68865741
20,11089319
21,60697029
37,17733052
49,76956034
44,73677893
59,83743179
90,11908554
115,5926424
140,3827628
171,5988008
188,4181565
239,1130783

другим масивом набирається матриця коефіцієнтів початкових рівнянь X(A3:AE15)

СЗП=X	X^2	X^3
152,83	23357,009	3569651,67
177,39	31468,690	5582362,104
231,04	53381,022	12333329,23
311,62	97104,947	30259519,88

375,98	141360,334	53148540,49
372,72	138920,198	51778336,35
524,14	274722,740	143993176,7
735,57	541063,225	397989876,3
928,81	862688,016	801273256,2
1194,91	1427809,908	1706104337
1573,99	2477444,520	3899472900
1650,43	2723919,185	4495637940
<b>1982,63</b>	<b>3930821,717</b>	<b>7793365061</b>

Таким чином, нами вперше отримана математична модель залежності купівельної спроможності жителів України від середньої заробітної плати з врахуванням корупції і інфляції у вигляді кубічного поліному

$$Y' = 2.632E - 0.8X^3 - 8.248E - 0.5X^2 + 0.186049X - 12.68148. \quad (6.182)$$

Встановлений коефіцієнт детермінації  $R^2 = 0.997$  (при апроксимації квадратичним поліномом  $R^2 = 0.994$ ), що говорить про високу адекватність побудованої математичної моделі експериментальним даним

За критерієм Стьюдента всі коефіцієнти можна вважати статистично значимими.

2,9168901	-2,897106	7,327291	-2,24940	$t(0,05;9)=$	2,262157
t(a)	t(b)	t(c)	t(d)		

Середні квадратичні похиби встановлених нами коефіцієнтів забезпечують високу точність самих коефіцієнтів

9,026E-09	2,8472E-05	0,025391	5,637707
m(a)	m(b)	m(c)	

m(d)

За критерієм Фішера-Снедекора F критерій дорівнює 1034,089 (при  $F=883.52$  апроксимованої моделі квадратичним поліномом).

Середня квадратична похибка одиниці ваги  $\mu=4,53$  при  $\mu=5,995$  в апроксимованій моделі квадратичним поліномом.

Тобто, апроксимація кубічним поліномом зменшує середню квадратичну похибку одиниці ваги на 1,5 гривни.

Накінець, нам залишається перевірити дію шостої теореми, яку ми вивели в книзі 1 по конструюванню математичних моделей.

Ми стверджуємо:

**Теорема 6. Якщо в емпіричні значення функції Y ввести абсолютні похибки зрівноваження, поділені на корінь квадратний із відповідної ваги функції, взяті із попередньо-го зрівноваження, то значно поліпшуються оцінки і характеристики в порівнянні з характеристиками попередньої моделі і отримана нова модель буде близькою до попередньої, що обумовлює адекватність її застосування.**

Будуємо математичну модель при  $Y=Y+V/P^{0.5}$

Вектор Y з введеними абсолютними похибками зрівноваження, поділеними на корінь квадратний із відповідної ваги функції, взятої із попереднього зрівноваження

$Y''=Y+V/P^{0,5}$
18,00594292

20,62662229
24,07133459
37,07550023
47,97133038
44,50154788
60,50600683
88,46846266
112,2460958
140,0825552
177,2098636
190,4150785
232,1811337

1650,43	2723919,18
<b>1982,63</b>	3930821,72

Комп'ютерна формула побудови математичної моделі квадратичним поліномом з врахуванням матеріалів попереднього зрівноваження (тобто представленого вище вектора  $Y''=Y+V/P^{0,5}$ .

$$= \text{ЛИНЕЙН}(AM3:AM15;AG3:AH15;1;1). \quad (6.183)$$

При цьому вектором  $Y''$  (AM3:AM15) представлені приведені вище значення. Діапазоном **X(AG3:AH15)** представлена матриця коефіцієнтів початкових рівнянь

152,83	23357,0089
177,39	31468,6904
231,04	53381,0219
311,62	97104,9469
375,98	141360,334
372,72	138920,198
524,14	274722,74
735,57	541063,225
928,81	862688,016
1194,91	1427809,91
1573,99	2477444,52

Контроль розрахунків функцією ЛИНЕЙН				
Функція ЛИНЕЙН (по X;X^2)				
a	b	c	F(0,05;2;10)=	4,102821
-2,24412E-06	0,119896	-0,19007	a,b,c	Коефіц.
2,93756E-06	0,006022	2,182239	стандарт S	$a_i=S\sqrt{d_{ii}}$
0,998664864	2,893726	#Н/Д	R^2	$\mu$
3739,936276	10	#Н/Д	Fкритерій	n-m-1
62633,81716	83,73648	#Н/Д	$[(Y'-Y_{cp})^2]$	[VV]
-0,763938967	19,91034	-0,0871	t(0,05;10)=	2,228139

Конструювання моделі

Таким чином, нами отримана зконструйована математична модель купівельної спроможності громадян України поліномом другого степеня за матеріалами повторного зрівноваження

$$Y'' = -2.24412E - 0.6X^2 + 0.119896X - 0.19007. \quad (6.184)$$

При цьому коефіцієнт детермінації  $R^2=0.99866$  при  $R^2=0.99437$  в попередній моделі. Середня квадратична похибка одиниці ваги склала  $\mu=2,89$  гривні при  $\mu=5,995$  в попередній моделі, тобто середня квадратична похибка одиниці ваги зменшилась на 3,1 гривні. Критерій Фішера  $F=3739,936276$ , при  $F=883.52$  в попередній моделі.

Середні квадратичні похибки зрівноважених коефіцієнтів

2,93756E-06	0,006022	2,182239
m(a)	m(b)	m(c)

при середніх квадратичних похибках у попередній моделі

6,08638E-06	0,012477	4,521413692
m(a)	m(b)	m(c)

Тобто, значно збільшилась точність знайдених коефіцієнтів математичної моделі. Звичайно, дещо змінилися і самі коефіцієнти моделі.

Статистична значимість конструйованої моделі по критерію Стьюдента

-0,763938967	19,91034	-0,0871	t(0,05;10)=	2,228139
t(a)	t(b)	t(c)		

при статистичній значимості попередньої моделі

0,086698357	9,400095	0,088077747	t(0,05;10)=	2,228138842
t(a)	t(b)	t(c)		

Статистична значимість коефіцієнта b зростає в 19,91/9,40=2,1 рази, коефіцієнти a і c, як і раніше, виявилися статистично не значимими.

### 6.3. Побудова математичної моделі узагальненим методом найменших квадратів (УМНК)

Для застосування узагальненого методу найменших квадратів (УМНК) необхідно знати ковариаційну матрицю вектора збурень  $\Omega$ , що зустрічається вкрай рідко на практиці. Якщо ж вважати всі  $n(n+1)/2$  елементів симетричної ковариаційної матриці  $\Omega$ , невідомими параметрами узагальненої моделі (в доповнення до  $(m+1)$  параметрам  $\beta_i$ ), то загальне число параметрів значно

перевищить число спостережень n, що зробить оцінку цих параметрів нерозрешимою задачею.

Із матричної алгебри відомо, що всяка не вироджена симетрична ( $n \times n$ ) матриця  $A$  допускає представлення у вигляді  $A=PP^T$ , де  $P$  – деяка не вироджена ( $n \times n$ ) матриця.

Тому, існує така не вироджена ( $n \times n$ ) матриця  $P$ , що

$$\Omega=PP^T, \quad (6.185)$$

(представлення матриці  $\Omega$  у вигляді (6.169) не єдине, але для нас це не має значення).

Приймаючи до уваги властивості обернених квадратних матриць, тобто

$$(AB)^{-1}=B^{-1}A^{-1} \quad \text{і} \quad (P^T)^{-1}=(P^{-1})^T,$$

А це означає, що

$$\Omega^{-1}=(P^{-1})^T P^{-1}. \quad (6.186)$$

Замітимо, що якщо обидві частини рівності (6.185) помножити зліва на матрицю  $P^{-1}$ , а справа на матрицю  $(P^T)^{-1}=(P^{-1})^T$ , то в добутку отримаємо одиничну матрицю  $E$ .

Дійсно,  $P^{-1} \Omega (P^T)^{-1}=P^{-1} (PP^T) (P^T)^{-1}=(P^{-1}P) P^T (P^T)^{-1}=E_n$ , тобто

$$P^{-1} \Omega (P^{-1})^T=E_n. \quad (6.187)$$

Отримавши значення обернених ваг зрівноваженої функції  $1/P_y$ , із результатів зрівноваження по звичайному методу найменших квадратів (МНК), знаходимо кореляційну матрицю  $\Omega_0^{-1}$  за формулою

$$\Omega_0^{-1} = \sqrt{1/P_y} * \sqrt{1/P_y}^T. \quad (6.188)$$

13x13    13x1    1x13

Допоміжна матриця  $\varepsilon^T \Omega^{-1}$

-0,9975	0,87931	0,67667	0,49415	0,43224	0,43388	0,47988	0,746387678	0,949797369	0,968441571	0,82387	0,9041327	-2,683253289
Матриця $\Omega^{-1} = P^{-1} \cdot (P^{-1})^T$												
0,067955	0,059904	0,046098	0,033664	0,029446	0,029558	0,032692	0,050848	0,064705	0,065976	0,056126	0,061595	0,183479
0,059904	0,052806	0,040636	0,029675	0,025957	0,026056	0,028819	0,044823	0,057039	0,058159	0,049476	0,054297	0,16174
0,046098	0,040636	0,031271	0,022836	0,019975	0,020051	0,022177	0,034493	0,043894	0,044755	0,038074	0,041783	0,124465
0,033664	0,029675	0,022836	0,016677	0,014587	0,014643	0,016195	0,025189	0,032054	0,032683	0,027804	0,030513	0,090893
0,029446	0,025957	0,019975	0,014587	0,01276	0,012808	0,014166	0,022033	0,028038	0,028588	0,024321	0,02669	0,079505
0,029558	0,026056	0,020051	0,014643	0,012808	0,012857	0,01422	0,022117	0,028145	0,028697	0,024413	0,026792	0,079807
0,032692	0,028819	0,022177	0,016195	0,014166	0,01422	0,015727	0,024462	0,031129	0,03174	0,027001	0,029632	0,088268
0,050848	0,044823	0,034493	0,025189	0,022033	0,022117	0,024462	0,038047	0,048416	0,049367	0,041997	0,046089	0,13729
0,064705	0,057039	0,043894	0,032054	0,028038	0,028145	0,031129	0,048416	0,061611	0,06282	0,053442	0,058649	0,174705
0,065976	0,058159	0,044755	0,032683	0,028588	0,028697	0,03174	0,049367	0,06282	0,064054	0,054491	0,0598	0,178134
0,056126	0,049476	0,038074	0,027804	0,024321	0,024413	0,027001	0,041997	0,053442	0,054491	0,046356	0,050873	0,151541
0,061595	0,054297	0,041783	0,030513	0,02669	0,026792	0,029632	0,046089	0,058649	0,0598	0,050873	0,055829	0,166305
0,183479	0,16174	0,124465	0,090893	0,079505	0,079807	0,088268	0,13729	0,174705	0,178134	0,151541	0,166305	0,495394

Застосовуючи *звичайний* метод найменших квадратів (МНК), невідомі параметри регресійної моделі знаходимо, мінімізуючи остаточно суму квадратів

$$S = \varepsilon^T \varepsilon = \sum_{i=1}^n (\hat{y}_i - y_i)^2. \quad (6.189)$$

Застосовуючи *узагальнений* метод найменших квадратів (УМНК), мінімізуючи остаточно суму квадратів  $S = \varepsilon^T \Omega^{-1} \varepsilon$  знаходимо спочатку кореляційну матрицю  $\Omega^{-1}$  за формулою  
 $= \text{МУМНОЖ}(\text{АJ3:AJ15}; \text{ТРАНСП}(\text{AJ3:AJ15})), \quad (6.190)$   
де в діапазоні **AJ3:AJ15** знаходяться обернені ваги зрівноваженої функції  $1/P_Y$ .

Допоміжна матриця  $\varepsilon^T \Omega^{-1}$  знаходиться за формулою

$$\text{МУМНОЖ}(\text{ТРАНСП}(\text{AA3:AA15}); \text{AZ3:BL15}). \quad (6.191)$$

Тоді, остаточно суму квадратів узагальненого методу  $S = e^T \Omega^{-1} e$

$$= \text{МУМНОЖ}(\text{BX18:CJ18}; \text{AA3:AA15}) \quad (6.192)$$

І в нашому випадку, отримали:  $S = 14,64212$ . При цьому середня квадратична похибка одиниці ваги буде  $\mu(\Omega) = 1,210046$ . Нагадаємо, що в звичайному МНК  $S[VV] = 359,4666$ , а середня квадратична похибка одиниці ваги була  $\mu[VV] = 5,995554$ .

Матриця коваріаційна регресійної функції (функції зрівноважених елементів) ЗМНК $S(u) = \mu \cdot \Omega^{-1} = \mu^* P^{-1*} (P^{-1})^{TM}$ (звичайний МНК)												
						Оцінки обґрунтовані, незміщені, неефективні						
9,370667	8,798036	7,717936	6,59541	6,168415	6,180145	6,499489	8,105799	9,143849	9,233158	8,516127	8,921331	15,39757
8,798036	8,260398	7,246302	6,192373	5,79147	5,802483	6,102313	7,610463	8,585079	8,668931	7,995716	8,376159	14,45664
7,717936	7,246302	6,356702	5,43216	5,080475	5,090135	5,353156	6,676156	7,531123	7,604681	7,014114	7,347851	12,66186
6,59541	6,192373	5,43216	4,642086	4,341551	4,349807	4,574573	5,705151	6,435768	6,498627	5,993955	6,279152	10,83736
6,168415	5,79147	5,080475	4,341551	4,060474	4,068195	4,278409	5,335792	6,019108	6,077898	5,605898	5,872632	10,13574
6,180145	5,802483	5,090135	4,349807	4,068195	4,075931	4,286545	5,345939	6,030554	6,089455	5,616558	5,883799	10,15501
6,499489	6,102313	5,353156	4,574573	4,278409	4,286545	4,508042	5,622178	6,342169	6,404114	5,906781	6,187831	10,67975
8,105799	7,610463	6,676156	5,705151	5,335792	5,345939	5,622178	7,011665	7,909597	7,986852	7,366606	7,717115	13,31918
9,143849	8,585079	7,531123	6,435768	6,019108	6,030554	6,342169	7,909597	8,922522	9,00967	8,309994	8,70539	15,02487
9,233158	8,668931	7,604681	6,498627	6,077898	6,089455	6,404114	7,986852	9,00967	9,097668	8,391158	8,790417	15,17162
8,516127	7,995716	7,014114	5,993955	5,605898	5,616558	5,906781	7,366606	8,309994	8,391158	7,739515	8,107768	13,99342
8,921331	8,376159	7,347851	6,279152	5,872632	5,883799	6,187831	7,717115	8,70539	8,790417	8,107768	8,493542	14,65924
15,39757	14,45664	12,66186	10,83736	10,13574	10,15501	10,67975	13,31918	15,02487	15,17162	13,99342	14,65924	25,30078

Матриця коваріаційна регресійної функції (функції зрівноважених елементів) УМНК. $KOB = C(u) = \mu \cdot \Omega^{-1} = \mu^* P^{-1*} (P^{-1})^{TM}$ (узагальнений МНК)												
						Оцінки обґрунтовані, незміщені, ефективні						
0,381694	0,358369	0,314374	0,26865	0,251257	0,251735	0,264743	0,330173	0,372455	0,376093	0,346886	0,363392	0,627188
0,358369	0,33647	0,295163	0,252233	0,235903	0,236352	0,248565	0,309996	0,349695	0,353111	0,325689	0,341185	0,588861
0,314374	0,295163	0,258927	0,221268	0,206942	0,207336	0,21805	0,271939	0,306764	0,309761	0,285705	0,299299	0,516569
0,26865	0,252233	0,221268	0,189086	0,176844	0,17718	0,186336	0,232387	0,262147	0,264708	0,244151	0,255768	0,441437
0,251257	0,235903	0,206942	0,176844	0,165395	0,165709	0,174272	0,217342	0,245176	0,24757	0,228344	0,239209	0,412858
0,251735	0,236352	0,207336	0,17718	0,165709	0,166024	0,174603	0,217756	0,245642	0,248041	0,228779	0,239664	0,413643
0,264743	0,248565	0,21805	0,186336	0,174272	0,174603	0,183626	0,229008	0,258335	0,260858	0,2406	0,252048	0,435017
0,330173	0,309996	0,271939	0,232387	0,217342	0,217756	0,229008	0,285605	0,322181	0,325328	0,300063	0,31434	0,542529
0,372455	0,349695	0,306764	0,262147	0,245176	0,245642	0,258335	0,322181	0,36344	0,36699	0,33849	0,354596	0,612006
0,376093	0,353111	0,309761	0,264708	0,24757	0,248041	0,260858	0,325328	0,36699	0,370574	0,341796	0,358059	0,617984
0,346886	0,325689	0,295705	0,244151	0,228344	0,228779	0,2406	0,300063	0,33849	0,341796	0,315253	0,330253	0,569992
0,363392	0,341185	0,299299	0,255768	0,239209	0,239664	0,252048	0,31434	0,354596	0,358059	0,330253	0,345967	0,597113
0,627188	0,588861	0,516569	0,441437	0,412858	0,413643	0,435017	0,542529	0,612006	0,617984	0,569992	0,597113	1,030574

"Зважучи" кожний залишок  $e(i) = y(i) - \hat{y}(i)$  за допомогою коефіцієнта  $1/\sigma(i)$  ми забезпечуємо рівномірний внеску залишків в загальну суму, що приводить в кінцевому рахунку до отримання найбільш ефективних оцінок параметрів моделі.

Кореляційна матриця $KOP = 1/P(Y^{*0,5} \text{ УМІНК МУМНОЖ}((1/P^{*0,5})^{*} \text{ ТРАНСП}(1/P^{*0,5})))$												
0,260693	0,244753	0,214705	0,183478	0,171599	0,171925	0,180809	0,225495	0,254373	0,256857	0,23691	0,248182	0,428345
0,244753	0,229796	0,201585	0,172266	0,161113	0,161419	0,16976	0,211715	0,238828	0,241161	0,222433	0,233016	0,402169
0,214705	0,201585	0,176837	0,151117	0,141334	0,141602	0,148919	0,185724	0,209508	0,211555	0,195126	0,20441	0,352797
0,183478	0,172266	0,151117	0,129138	0,120778	0,121007	0,12726	0,158712	0,179037	0,180785	0,166746	0,17468	0,301484
0,171599	0,161113	0,141334	0,120778	0,112958	0,113173	0,119021	0,148436	0,167446	0,169081	0,15595	0,163371	0,281966
0,171925	0,161419	0,141602	0,121007	0,113173	0,113388	0,119247	0,148719	0,167764	0,169403	0,156247	0,163681	0,282502
0,180809	0,16976	0,148919	0,12726	0,119021	0,119247	0,125409	0,156403	0,176433	0,178156	0,164321	0,172139	0,2971
0,225495	0,211715	0,185724	0,158712	0,148436	0,148719	0,156403	0,195057	0,220037	0,222186	0,204932	0,214682	0,370526
0,254373	0,238828	0,209508	0,179037	0,167446	0,167764	0,176433	0,220037	0,248216	0,25064	0,231176	0,242175	0,417977
0,256857	0,241161	0,211555	0,180785	0,169081	0,169403	0,178156	0,222186	0,25064	0,253088	0,233434	0,244541	0,422059
0,23691	0,222433	0,195126	0,166746	0,15595	0,156247	0,164321	0,204932	0,231176	0,233434	0,215306	0,22555	0,389283
0,248182	0,233016	0,20441	0,17468	0,163371	0,163681	0,172139	0,214682	0,242175	0,244541	0,22555	0,236282	0,407805
0,428345	0,402169	0,352797	0,301484	0,281966	0,282502	0,2971	0,370526	0,417977	0,422059	0,389283	0,407805	0,703842

Вихідні дані	Введення ваг у вихідні дані									
СЗП=X	КСП1=У	X0	X*1/P(y) <sup>0,5</sup>	(X*1/P(y) <sup>0,5</sup> ) <sup>2</sup>	Y*1/P(y) <sup>0,5</sup>	1/P(y) <sup>0,5</sup>				
152,83	17,68865741	1	78,03053155	6088,763853	9,031311522	0,510570775				
177,39	20,11089319	1	85,03750896	7231,37793	9,6405666497	0,479370379				
231,04	21,60697029	1	97,15834062	9439,743153	9,086162969	0,420519992				
311,62	37,17733052	1	111,9819374	12539,9543	13,35997057	0,359357985				
375,98	49,76956034	1	126,3638464	15967,82168	16,72718501	0,336092682				
372,72	44,73677893	1	125,5066697	15751,92414	15,06429528	0,336731782				
524,14	59,83743179	1	185,6145473	34452,76015	21,19032665	0,35413162				
735,57	90,11908554	1	324,8667656	105538,4154	39,80137286	0,441653093				
928,81	115,5926424	1	462,7446532	214132,6141	57,58968703	0,498212394				
1194,91	140,3827628	1	601,1335308	361361,5218	70,62354975	0,5030785				
1573,99	171,5988008	1	730,347507	533407,4809	79,62360392	0,464010259				
1650,43	188,4181565	1	802,2546335	643612,4969	91,5878523	0,486088252				
<b>1982,63</b>	<b>239,1130783</b>	<b>1</b>	<b>1663,333439</b>	<b>2766678,13</b>	<b>200,6046408</b>	<b>0,838953027</b>				

Розрахунок по даним з врахування ваг		
εε	ε'=Y'''-Y	Y'''
0,447077	0,668638073	9,69995
0,69521	0,83379225	10,47436
7,448463	2,72918718	11,81535
0,009566	0,097804878	13,45778
2,800218	-1,673385049	15,0538
0,01117	-0,105690073	14,95861
0,216357	0,465142207	21,65547
6,069059	-2,463545935	37,33783
20,19116	-4,493457072	53,09623
2,189101	-1,479560922	69,14399
22,21321	4,713089103	84,33669
1,666397	1,290889913	92,87874
0,339778	-0,582904551	200,0217
64,29676	-5,32907E-14	

2,535681	Контроль
μ	

Розрахунок по даним без врахування ваг		
εε	ε=Y''-Y	Y''
0,095320482	0,30874015	17,9974
0,39204246	0,62613294	20,73703
26,40795247	5,13886685	26,74584
1,800409855	-1,3417935	35,83554
43,78152678	-6,616761	43,1528
3,824808294	-1,9557117	42,78107
0,12231061	0,34972934	60,18716
26,66794633	-5,1641017	84,95498
56,67544323	-7,5283095	108,0643
0,058822053	0,24253258	140,6253
285,2561885	16,889529	188,4883
98,6427868	9,93190751	198,3501
8,497576888	2,91506036	242,0281
552,2231347	13,7958213	
7,431171743		
μ		

a	b	c	F(0,05;2;10)=	4,102821015
6,0408E-06	0,109534654	1,116121317	a,b,c	Коефіц.
3,084E-06	0,005107977	1,316376438	стандарт S	ai=S√dii
0,99818048	2,535680559	#Н/Д	R^2	μ
2742,98297	10	#Н/Д	Fкритерій	n-m-1
35272,983	64,29675896	#Н/Д	[(Y'-Ycp)^2]	[VV]
1,95874618	21,44384366	0,84787397	t(0,05;10)=	2,228138842



Висновки. 1. З введенням ваг регресійної моделі (зрівноважених функцій), визначених із попереднього зрівноваження в X і Y отримана математична модель:		
$Y'' =$	$0,109534654 X +$	$1,116121$
2. При розрахунку зрівноважених значень функції за даною формулою по значенням X і Y отримана середня квадратична похибка одиниці ваги $\mu =$		$7,431171743$
3. При цьому отримані значення зрівноваженої функції близькі до експериментальних даних, що обумовлює адекватність їх застосування.		
4. При розрахунку зрівноважених значень функції за даною формулою по значенням X(P) і Y(P) отримана середня квадратична похибка одиниці ваги $\mu =$		$2,535680559$
5. При цьому отримані значення зрівноваженої функції далекі до експериментальних даних, що не дає змоги їх застосування.		
6. При попередньому зрівноваженні $\mu =$	$5,9955536$	, тобто результати були кращими.

## Розглянемо метод максимальної правдоподібності

$\Omega^{-1}X$		
3,078114	2754,21	3795785
2,890014	2585,903	3563829
2,535219	2268,442	3126312
2,166487	1938,511	2671609
2,026226	1813,01	2498646
2,030079	1816,457	2503397
2,134978	1910,318	2632754
2,662625	2382,442	3283424
3,003608	2687,544	3703908
3,032945	2713,794	3740084
2,797411	2503,045	3449636
2,930515	2622,142	3613772
5,057855	4525,626	6237108

Матриця коефіцієнтів початкових рівнянь		
1	152,83	23357,0089
1	177,39	31468,69037
1	231,04	53381,02188
1	311,62	97104,94694
1	375,98	141360,3338
1	372,72	138920,1984
1	524,14	274722,7396
1	735,57	541063,2249
1	928,81	862688,0161
1	1194,91	1427809,908
1	1573,99	2477444,52
1	1650,43	2723919,185
1	<b>1982,63</b>	3930821,717
X0	X	X <sup>2</sup>

$X^T \Omega^{-1} X$		
36,34607661	32521,44359	44820262,77
32521,44359	29099269,89	40103906200
44820262,77	40103906200	5,52702E+13

Визначник= -1,24401E-09

Q=	$(X^T \Omega^{-1} X)^{-1}$	Оберн.матр.
-8,0442E+13	2,40188E+11	109047180,3
2,74878E+11	-268435456	28130,69019
-134217728	0	108,841125

$\Omega^{-1} Y$
323,6664
303,8875
266,5805
227,808
213,0594
213,4646
224,4948
279,9774
315,832
318,9168
294,1503
308,1462
531,8379

Метод максимальної правдоподібності

$X^T \Omega^{-1} Y$	$b' = (X^T \Omega^{-1} X)^{-1} * X^T \Omega^{-1} Y$	
3821,8218	-64	c
3419658,29	0,1875	b
4712889899	-6,10352E-05	a

$$Y'' = -6,10352E-05X^2 + 0,1875X - 64. \quad (6.193)$$

### Висновки

1. Розглянуті теоретичні основи побудови і дослідження математичної моделі за матеріалами експерименту.
2. Наведені приклади приведення задачі опрацювання експериментальних даних до задачі регресійного аналізу.
3. Встановлені основні положення регресійного аналізу.
4. Відображені обчислювальні алгоритми звичайного методу найменших квадратів (ЗМНК)
5. Приведений регресійний аналіз найпростіших поліноміальних моделей.
6. Позначені особливості регресійного аналізу при порушенні базових положень (узагальнений метод найменших квадратів УМНК).
7. Встановлено, що при застосуванні кореляційної матриці узагальненого методу найменших квадратів, середня квадратична похибка одиниці ваги зменшилась у шість раз.
8. Проведена оцінка параметрів регресійної моделі «зваженим» методом найменших квадратів, що приводить в кінцевому рахунку до ефективних оцінок параметрів моделі.
9. Для узагальненої регресійної моделі, на відміну від класичної, коефіцієнт детермінації не являється задовільною мірою якості моделі. В загальному випадку  $R^2$  може виходити навіть за межі інтервалу  $[0;1]$ , а додавання (видалення) пояснюючої змінної не обов'язково приводить до його збільшення (зменшення).

### Літературні джерела

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей. -М.: Финансы и статистика, 1985.
2. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных.-М.: Финансы и статистика, 1983.
3. Асатурян В.И. Теория планирования эксперимента. -М.: Радио и связь, 1983.
4. Астанин Л. Ю., Дорский Ю.Д., Костылев А.А. Применение программируемых калькуляторов для инженерных и научных расчётов. -Л.: Энергоатомиздат, 1986.
5. Астанин Л.Ю., Костылев А.А. Основы сверхширокополосных радиолокационных измерений.- Радио и связь, 1989
6. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные Методы.- Наука, 1987.
7. Беклемишев Д.В. Дополнительные главы линейной алгебры. -М.:Наука, 1983.
8. Болнокин В.Е., Чинаев П.И. Анализ и синтез систем автоматического управления на ЭВМ: Алгоритмы и программы. -М.: Радио и связь, 1986.
9. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1965.
10. Браверманн Э.М., Мучник И.В. Структурные методы обработки эмпирических данных.- М.: Наука, 1983.
11. Бухтияров А.М., Маликова Ю.П., Фролов Г.Д. Практикум по программированию на фортране (ОС ЕС ЭВМ).- 3-е изд.- М.: Наука, 1988.
12. Вайну Я.Я. Корреляция рядов динамики.- Статистика, 1977.

13. Ватугин В.А., Телевинова Т.М., Чистяков В.П. Вероятностные методы в физических исследованиях.-М.: Наука, 1985.
14. Видениекс П.О., Вентиньш Я.Я., Кривчиков А.А. Проблемно-ориентированные микропроцессорные системы в производстве РЭА.- М.: Радио и связь, 1987.
15. Вучков И., Бояджијева Л., Солаков Е. Прикладной линейный регрессионный анализ.- М.: Финансы и статистика, 1987.
16. Вычислительная математика/ Н. И. Данилина, Н.С. Дубровская, О.П.Кваша, Г.Л.Смирнов.- Высшая школа, 1985.
17. Гантмахер Ф.Р. Теория матриц. -М.: Наука, 1988.
18. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике.- М.: Высшая школа, 1975.-333с.
19. Гольцман Ф.М. Физический эксперимент и статистические выводы.- Л.: Изд-во Ленингр. гос. ун-та, 1982.
20. Горелик А.М., Ушаков В.А., Шура-Бура М.Р. Мобильность программ на фортране.- М.: Финансы и статистика, 1984.
21. Грановский В.А. Динамические измерения. : Основы метрологического обеспечения.. – Л.: Энергоатомиздат, 1984.
22. Гришин В.К., Живописцев Ф.А., Иванов В.А. Математическая обработка и интерпретация физического эксперимента.- М.: Изд-во Моск. ун-та, 1988.
23. Гришин Ю.П., Казаринов Ю.М. Динамические системы, устойчивые к отказам. – М.: Радио и связь, 1985.
24. Гроп Д. Методы идентификации систем. – М.: Мир, 1979
25. Губарев В.В. Алгоритмы статистических измерений. – М.: Энергоатомиздат, 1985.
26. Дайитбегов Д.М., Калмыкова О.В., Черепанов А.И. Программное обеспечение статистической обработки данных. – М.: Финансы и статистика, 1984.

27. Демиденко Е.З. Линейная и нелинейная регрессии.- М.: Финансы и статистика, 1981.
28. Демиденко Е.З. Линейная и нелинейная регрессия: Фортран-.IV- М.: Изд-во Ин-та мировой экономики и международных отношений , 1979.
29. Денисов В.И., Попов А.А. Пакет программ оптимального планирования эксперимента. – М.: Финансы и статистика , 1986.
30. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке : Методы обработки данных.-М.: Мир, 1980.
31. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке : Методы планирования эксперимента.-М.: Мир, 1981.
32. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. В 2-х книгах.-2-е изд. –М.: Финансы и статистика, 1987.
33. Дьяконов В.П. Справочник по алгоритмам и программам на языке бейсик для персональных ЭВМ.-М.: Наука, 1989.-240 с.
34. Дьяконов В.П. Справочник по расчётам на микрокалькуляторах. .-М.: Наука, 1989.-464с.
35. Единая система электронных вычислительных машин. Пакет научных подпрограмм. Техническое описание: Руководство программиста. Кн.2.- Минск, 1976.
36. Елисеева И.И., Рукавишников В.О. Логика прикладного статистического анализа.- М.: Финансы и статистика, 1982.
37. Енюков И.С. Методы, алгоритмы, программы многомерного статистического анализа: Пакет ППСА. .- М.: Финансы и статистика, 1986.
38. Загоруйко Н.Г., Елкина В.Н., Емельянов С.В. Пакет прикладных программ ОТЭКС (для анализа данных) .- М.: Финансы и статистика, 1986.
39. Иванов В.В. Методы вычислений на ЭВМ: Справочное

- пособие.- Киев: Наукова думка, 1986.
40. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным.- М.: Радио и связь, 1987.
41. Казаков И.Е., Гладков Д.И. Методы оптимизации стохастических систем. –М.: Наука, 1987.
42. Кендалл М., Стьюарт А. Статистические выводы и связи./Пер.с англ.-М.: Наука, 1973.
43. Кендалл М., Стьюарт А. Теория распределений./Пер.с англ.-М.: Наука, 1966.
44. Костылев А.А. Идентификация радиолокационных целей при использовании сверхширокополосных сигналов: Методы и приложения//Зарубежная радиоэлектроника.- 1984.-№4.-с.75-104.
45. Костылев А.А. Обработка сигналов при экспериментальном исследовании рассеяния коротких радиоимпульсов на проводящей сфере//Радиотехника.-1984.-№8.-с.64-66.
46. Котюков В.И. Многофакторные кусочно-линейные модели. .- М.: Финансы и статистика, 1984.
47. Крутько П. Д., Максимов А.И., Скворцов Л.М. Алгоритмы и программы проектирования автоматических систем/ Под ред. П. Д. Крутько. -М.: Радио и связь, 1988.
48. Кудрявцев Е.М. Исследование операций в задачах, алгоритмах и программах. –М.: Радио и связь, 1984.
49. Кузьмин С.З. Основы проектирования систем цифровой обработки радиолокационной информации-М.: Радио и связь , 1986.50.
50. Кузьмичев Д.А., Радкевич И.А., Смирнов А.Д. Автоматизация экспериментальных исследований: Учебное пособие. – М.: Наука, 1983.
51. Куликов Е.И. Методы измерения случайных процессов.-М.: Радио и связь, 1986.
52. Літнарівич Р.М. Дослідження впливу корупції і інфляції на купівельну спроможність громадян України.

- МЕРУ, Рівне, 2011.- 114 с.  
<http://essuir.sumdu.edu.ua/handle/123456789/14860>.
53. Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов/Пер.с англ.-М.: Наука, 1986.
  54. Микро-ЭВМ: Практическое пособие/Под ред.Л.Н.Преснухина. Кн.1. Семейство ЭВМ «Электроника-60»/И.Л.Талов, А.Н.Сохина, В.Д.Борисенков.- М.: Высшая школа, 1988.
  55. Мини- и микро-ЭВМ семейства «Электроника»/Б.Л. Толстых, И.Л.Талов, В.Г.Цивинский и др.- М.: Радио и связь, 1987.
  56. Мюллер П., Нойман П., Шторм Р. Таблицы по математической статистике/ Пер.с нем.. - М.: Финансы и статистика, 1982.
  57. Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений.- Л.: Энергоатомиздат, 1985.
  58. Обработка геологической информации на микрокалькуляторах /В.В.Бабенко, В.П.Афанасьев, Н.Н.Зинчук и др.- М.: Недра, 1988.
  59. Операционная система СМ ЭВМ РАФОС: Справочник/Л.И. Валикова, Г.В.Вигдорчик, А.Ю.Воробьев, А.А.Лукин; Под общ.ред. В.П. Семика.- М.: Финансы и статистика, 1984.
  60. Певчев Ю.Ф., Финогенов К.Г. Автоматизация физического эксперимента: Учебное пособие для вузов.- М.: Энергоатомиздат, 1986.
  61. Полард Дж. Справочник по вычислительным методам статистики. - М.: Финансы и статистика, 1982.
  62. Программирование, отладка и решение задач на ЭВМ единой серии. Язык фортран: Учебное пособие для вузов/И.А.Кудряшов, Н.Х.Кушнер, Л.В.Петрова, Н.А.Силов; Под ред. И.А.Кудряшова.- Энергоатомиздат, 1988.
  63. Программное обеспечение СМ ЭВМ. Операционная

- система с разделением функций РАФОС. Библиотека для научно-технических расчётов БНТР/РАФОС. Описание применения., 1981.-Т.9.-Кн.1,-Ч.1.
64. Сборник научных программ на фортране: Руководство для программиста. Вып.1. -М.: Статистика, 1974.
  65. Сборник научных программ на фортране: Руководство для программиста. Вып.2. -М.: Статистика, 1976.
  66. Светозарова Г.И., Мельников А.А., Козловский А.В. Практикум по программированию на языке «бейсик».- М.: Наука, 1988.
  67. Себер Дж. Линейный регрессионный анализ. М.:Мир, 1980.
  68. Селиванов М.Н., Фридман А.Э., Кудряшова Ж.Ф. Качество измерений: Метрологическая справочная книга.- Л.: Лениздат, 1987.
  69. Сильвестров Д.С. Программное обеспечение прикладной статистики: Обзор состояния. Тенденция развития.- М.: Финансы и статистика, 1988.
  70. Справочник по вероятностным расчетам. -М.: Воениздат, 1970.
  71. Справочник по специальным функциям с формулами, графиками и математическими таблицами/Под ред. М. Абрамовица и И. Стиган /Пер.с англ.-М.: Наука, 1979.
  72. Статистическая обработка результатов экспериментов на микро-ЭВМ и программируемых калькуляторах /А.А. Костылев, П.В. Миляев, Ю.Д. Дорский и др.: -Л.: Энергоатомиздат, 1991.-304 с.
  73. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач.-3-е изд.,перераб.-М.: Наука, 1986.
  74. Тихонов А.Н., Уфимцев М.В. Статистическая обработка результатов экспериментов.-М.: Изд-во Моск.ун-та, 1988.
  75. Трохименко Я.К. Программирование микрокалькуляторов «Электроника МК-52» и «Электроника МК-61».-

Киев: Техника, 1987.

76. Трохименко Я.К., Любич Ф.Д. Радиотехнические расчеты на микрокалькуляторах.- М.: Радиосвязь, 1988.
77. Уолш Б. Программирование на бейсике.-М.: Радио и связь, 1987.
78. Ферстер Э., Ренц Б. Методы корреляционного и регрессионного анализа.- М.: Финансы и статистика, 1983.
79. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений.- М.: Мир , 1980.
80. Хастингс Н., Пикок Дж. Справочник по статистическим распределениям/Пер.с англ.-М.:Статистика, 1980.
81. Цветков А.Н., Епанечников В.А. Прикладные программы для микро-ЭВМ «Электроника БЗ-34», «Электроника МК-56».- М.: Финансы и статистика, 1984.
82. Чуа Л.О., Лин Пен-Мин. Машинный анализ электронных схем: Алгоритмы и вычислительные методы/Пер. с англ.-М.: Энергия, 1980.
83. Шалыгин А.С., Палагин Ю.И. Прикладные методы статистического моделирования. – Л.: Машиностроение, 1986.
84. Шелест А.Е. Микрокалькуляторы в физике: Справочное пособие.- М.: Наука, 1988.
85. Шуп Т. Решение инженерных задач на ЭВМ.-М.: Мир, 1982.

**Руслан Миколайович Літнарівч,**  
**доцент, кандидат технічних наук**

**ПОБУДОВА І ДОСЛІДЖЕННЯ МАТЕМАТИЧНОЇ  
МОДЕЛІ ЗА ДЖЕРЕЛАМИ ЕКСПЕРИМЕНТАЛЬНИХ  
ДАНИХ МЕТОДАМИ РЕГРЕСІЙНОГО АНАЛІЗУ**

**НАВЧАЛЬНИЙ ПОСІБНИК**

**Наукове видання**

*Комп'ютерний набір, верстка – дизайн у редакторі  
Microsoft® Office 2003® Р.М.Літнарівч*

Книга написана за матеріалами роботи наукової фізико-математичної школи МЕГУ

*Міжнародний економіко-гуманітарний університет ім.  
академіка С. Дем'янчука*

**Кафедра математичного моделювання  
33027, м.Рівне, Україна  
Вул.акад. С.Дем'янчука,4, корпус 1  
Телефон:(+00380) 362 23-73-09  
Факс:(+00380) 362 23-01-86  
E-mail:mail@regi.rovno.ua**