УДК 685.115

В. О. Кутовий

ПРО ТЕОРЕМУ ГАУССА-МАРКОВА У ВИПАДКУ ВИРОДЖЕНОЇ МАТРИЦІ СПОСТЕРЕЖЕНЬ

V. O. Kutovyi. *On Gauss-Markov's theorem in the case of degenerate observation matrix*, Matematychni Studii, **14** (2000) 213–216.

In this paper sufficient conditions for validity of Gauss-Markov's theorem in the case of degenerate observation matrix are established.

В. А. Кутовый. *О теореме Гаусса-Маркова в случае вырожденной матрицы наблюдений* // Математичні Студії. – 2000. – Т.14, №2. – С.213–216.

В данной работе устанавливаются достаточные условия при которых теорема Гаусса-Маркова распространяется на случай вырожденной матрицы наблюдений.

Нехай існує лінійне співвідношення між змінною Y і m пояснюючими змінними $X_1, X_2, \ldots, X_m$ та збуренням ε.

Якщо ми маємо вибірку з n спостережень над змінними Y і X_j, $j = 1, 2, \ldots, m$, то можна записати

$$y_i = \sum_{j=1}^{m} x_{ij}\beta_j + \varepsilon_i, \quad i = 1, 2, \ldots, n. \tag{1}$$

Рівняння (1) можуть бути записані в матричній формі

$$Y = X\beta + \varepsilon, \tag{2}$$

де

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1m} \\ x_{21} & x_{22} & \ldots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nm} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Через X' і ε' позначимо матриці транспоновані до X і ε, відповідно. Нехай виконуються умови:

1. $M\varepsilon = 0;$ (3)
2. $M(\varepsilon\varepsilon') = \sigma^2 \cdot E, \quad E$— одинична матриця; (4)
3. X— матриця, елементи якої — детерміновані числа; (5)
4. $\operatorname{rang} X = m$ (матриця X— повного рангу.) (6)

2000 *Mathematics Subject Classification:* 62B15, 62F10.

Оцінки отримані методом найменших квадратів (МНК) $\hat{\beta}$ параметра β визначаються наступною формулою [1]

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{7}$$

Використовуючи (2), отримаємо

$$\hat{\beta}(X'X)^{-1}X'(X\beta + \varepsilon) = \beta + (X'X)^{-1}X'\varepsilon,$$

при цьому $M\hat{\beta} = \beta$. Матриця $(X'X)^{-1}X'$ є лівою зворотньою до матриці X. Справді $(X'X)^{-1}X'X = E$. Нехай матриця C є також лівою зворотньою до X, тобто $CX = E$ і $\hat{\hat{\beta}} = CY$. Тоді, використовуючи (2), отримаємо $\hat{\hat{\beta}} = CX\beta + C\varepsilon = \beta + C\varepsilon$. Отже, $M\hat{\hat{\beta}} = \beta$ і $\hat{\hat{\beta}}$ — незміщена лінійна по Y оцінка β. Нехай $\hat{Y} = X\hat{\beta}$, $\hat{\hat{Y}} = X\hat{\hat{\beta}}$. Тоді справедлива теорема Гаусса-Маркова, яку можна сформульовати в наступних еквівалентних формах.

Теорема 1. $D(\hat{y}_i) \le D(\hat{\hat{y}}_i)$; $i = 1, 2, \ldots, n$.

Теорема 2. $D(\hat{\beta}_j) \le D(\hat{\hat{\beta}}_j)$; $j = 1, 2, \ldots, m$.

Теорема 3. *Для загальної лінійної моделі (2) з припущеннями (3)–(6) МНК-оцінки* $\sum_{j=1}^{m} c_j\hat{\beta}_j$ *величини* $\sum_{j=1}^{m} c_j\beta_j$ *(c_j — довільні константи) є незміщеними, мають найменшу дисперсію серед усіх незміщених оцінок суми* $\sum_{j=1}^{m} c_j\beta_j$ *лінійних по* $y_1, y_2, \ldots, y_n$, *тобто*

$$D\left(\sum_{j=1}^{m} c_j\hat{\beta}_j\right) \le D\left(\sum_{j=1}^{m} c_j\hat{\hat{\beta}}_j\right)$$

У даній статті теорема Гаусса-Маркова поширюється на випадок, коли порушується умова (6), тобто $\operatorname{rang} X < m$.

Означення 1. Матриця X^+ з розмірами $m \times n$ називається псевдозворотньою Мура-Пенроуза для матриці X, якщо вона задовольняє наступні чотири умови:

$$\begin{array}{ll} 1. & X^+XX^+ = X^+; \\ 2. & XX^+X = X; \\ 3. & X^+X\text{— симетрична;} \\ 4. & XX^+\text{— симетрична.} \end{array} \tag{8}$$

Можна довести, що така матриця X^+ завжди існує та єдина [2]. Якщо X — невироджена квадратна матриця, то $X^+ = X^{-1}$, очевидно, задовольняє умови (8). Якщо X — прямокутна і має повний ранг, то $X^+ = (X'X)^{-1}X'$. Можна перевірити, що псевдозворотньою до діагональної $n \times m$ матриці

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \ldots & \sigma_m \\ 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

є діагональна $m \times n$ матриця

$$\Sigma^+ = \begin{pmatrix} \sigma_1^+ & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^+ & \dots & \cdot & \cdot & \dots & \cdot \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_m^+ & 0 & \dots & 0 \end{pmatrix}, \quad \text{де} \ \ \sigma_j^+ = \begin{cases} \frac{1}{\sigma_j}, & \text{якщо } \sigma_j \neq 0, \\ 0, & \text{якщо } \sigma_j = 0. \end{cases}$$

Далі використаємо сингулярний розклад матриці X [3].

$$X = U\Sigma V', \tag{9}$$

де U — ортогональна $n\times n$ матриця, V — ортогональна $m\times m$ матриця, а Σ — діагональна $n \times m$ матриця, у якої $\sigma_{ij} = 0$ при $i \neq j$ і $\sigma_{ii} = \sigma_i \geq 0$. Стовпчики матриці U є власними векторами матриці XX', а стовпчики матриці V є власними векторами матриці $X'X$. Використовуючи (9), отримаємо $X^+ = V\Sigma^+U'$.

Оцінки за методом найменших квадратів параметра β в (1) визначаються як значення $\hat\beta_1$, $\hat\beta_2, \dots, \hat\beta_m$, що мінімізують суму квадратів

$$S = \sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{m}\beta_j x_{ij}\right)^2 \to \min \tag{10}$$

по всіх наборах β_1, $\beta_2, \dots, \beta_m$. У випадку, коли виконується (6), точка мінімуму визначається однозначно формулою (7). Коли ж (6) не виконується, таких точок мінімуму виявляється багато. Будемо вимагати, щоб сума

$$L = \sum_{j=1}^{m}\beta_j^2 \to \min \tag{11}$$

була мінімальною. Тоді розв'язок (10), (11) однозначний і дорівнює [3] $\hat\beta = X^+Y$.

Означення 2. Оцінка $\hat\beta$ параметра β називається X-незміщеною, якщо $MX\hat\beta = X\beta$, тобто якщо $X\hat\beta$ — незміщена оцінка $X\beta$.

Лема 1. *Оцінка $\hat\beta = X^+Y$ є X-незміщеною оцінкою β.*

Доведення. Використаємо (2). Тоді $\hat\beta = X^+Y = X^+(X\beta + \varepsilon) = X^+X\beta + X^+\varepsilon$, отже з (8) $M(X\hat\beta) = M(XX^+X\beta + XX^+\varepsilon) = XX^+X\beta = X\beta$. □

Лема 2. *Матриця коваріацій $D(X\hat\beta)$ параметра $X\hat\beta$-оцінки $X\beta$ в моделі (2) дорівнює $D(X\hat\beta) = \sigma^2X(X'X)^+X'$, де $(X'X)^+$ — псевдозворотня до $X'X$.*

Доведення. Використаємо $X^+ = (X'X)^+X'$ [2]. Тоді $X\hat\beta = XX^+Y = XX^+X\beta + XX^+\varepsilon = X\beta + XX^+\varepsilon$. Отже,

$$\begin{aligned} D(X\hat\beta) &= M(X\hat\beta - X\beta)(X\hat\beta - X\beta)' = M(XX^+\varepsilon(XX_\varepsilon^+)') = \\ &= M\big(X(X'X)^+X'\varepsilon\varepsilon'X(X'X)^+X'\big) = \sigma^2X(X'X)^+(X'X)(X'X)^+X' = \sigma^2X(X'X)^+X'. \end{aligned}$$

□

Нехай C — матриця така, що $XCX = X$ і $\hat{\hat{\beta}} = CY$, тоді $\hat{\hat{\beta}}$ — X-незміщена оцінка β. Справді, $MX\hat{\hat{\beta}} = M(XCY) = M(XCX\beta + XC\varepsilon) = X\beta$.

Теорема 1′. *Якщо $\hat{\beta}$ — розв'язок (10)–(11), а $\hat{\hat{\beta}} = CY$ — інша лінійна X-незміщена оцінка β, де $XCX = X$ і $\hat{Y} = X\hat{\beta}$, $\hat{\hat{Y}} = X\hat{\hat{\beta}}$, то справджується $D(\hat{y}_i) \leq D(\hat{\hat{y}}_i)$, $i = 1, 2, \ldots, n$.*

Доведення. Нехай $C = X^+ + T$. Тоді $X = XCX = X(X^+ + T)X = XX^+X + XTX = X + XTX$. Отже маємо, $XTX = 0 \cdot E$, $X'T'X' = 0 \cdot E$. Використовуючи співвідношення Мура-Пенроуза (8), дослідимо матрицю коваріацій $D(X\hat{\hat{\beta}})$.

Маємо $M(X\hat{\hat{\beta}}) = MX(X^+ + T)(X\beta + \varepsilon) = X(X^+ + T)X\beta$ та

$$\begin{aligned} D(X\hat{\hat{\beta}}) &= M\left[X(X^+ + T)\varepsilon\varepsilon'(X^+ + T)'X'\right] = \sigma^2\big(X(X^+X^{+\prime} + X^+T' + TX^{+\prime} + TT')X')X'\big) \\ &= \sigma^2\big(X(X'X)^+X'X(X'X)^+X' + X(X'X)^+X'T'X' + XTX(X'X)^+X' + XTT'X'\big) \\ &= \sigma^2(X(X'X)^+X' + XTT'X') = \sigma^2(D(X\hat{\beta}) + XTT'X'). \end{aligned} \tag{12}$$

Але матриця $XTT'X'$ — напівдодатна, незалежно від того, яка матриця T. Тому мінімально можливе значення для $D(X\hat{\hat{\beta}})$ отримується, коли $XTT'X = 0$, тобто коли $D(X\hat{\hat{\beta}}) = D(X\hat{\beta})$. □

Нехай $c' = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$ належить до лінійної оболонки рядків $x_1, x_2, \ldots, x_n$ матриці X. Тоді $c = \gamma X$, де $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$. Помножимо (12) зліва на γ і справа на γ'. Тоді отримаємо $D(c\hat{\hat{\beta}}) = D(\gamma X\hat{\beta}) = \sigma^2(\gamma X(X'X^+)X'\gamma' + \gamma XTT'X'\gamma) = D(c\hat{\beta}) + \sigma^2(cTT'c')$. Отже маємо $D(c\hat{\hat{\beta}}) \leq D(c\hat{\beta})$, звідки випливає справедливість теореми 2 за додаткової умови, що $c = \{c_1, c_2, \ldots, c_m\}$ належить до лінійної оболонки рядків матриці X. Як наслідок отримаємо, якщо $e_j = \{0, \ldots, 0, 1, 0, \ldots, 0\}$ (де 1 на j-тій позиції), належить до лінійної оболонки рядків матриці X, що $D(\hat{\beta}_j) \leq D\hat{\hat{\beta}}_j)$.

Висновок. У випадку виродженої матриці X теорема 1 залишається справедливою. Теорема 3 справедлива лише за вказаних вище додаткових обмеженнях на c. Теорема 2, взагалі кажучи, для всіх β_j не справджується, тому що всі e_j не можуть належати до лінійної оболонки $x_1, x_2, \ldots, x_n$, оскільки $\operatorname{rang} X < m$.

ЛІТЕРАТУРА

1. Джонстон Дж. Эконометрические методы. – Москва, Статистика, 1980. – 445 с.
2. Лоусон Ч., Хенсон Р. Численное решение задач метода наименших квадратов. – Москва, Наука, 1986. – 230 с.
3. Воеводин В. В. Вычислительные основы линейной алгебры. – Москва, Наука, 1977. – 303 с.

Інститут математики НАН України, м. Київ

Надійшло 2.12.1999
Після переробки 4.06.2000