Журнал обчислювальної та прикладної математики | 2014, № 1(115) | Journal of Computational & Applied Mathematics

УДК 519.71

УМОВИ СПІВПАДАННЯ ОЦІНОК МЕТОДА НАЙМЕНШИХ КВАДРАТІВ ТА ОРТОГОНАЛЬНОЇ РЕГРЕСІЇ ПАРАМЕТРІВ ЛІНІЙНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ

М. Ю. САВКІНА

Резюме. В роботі отримано необхідні та достатні умови співпадання оцінки метода найменших квадратів та оцінки ортогональної регресії невідомих параметрів лінійної регресійної моделі з вільним членом.

Abstract. The sufficient and necessary conditions for coincidence of least square estimation and estimation of ortogonal regression of parameters of linear regression model with absolute term were obtained in the paper.

1. Вступ

Розглянемо модель лінійної регресії

$$y_i = f(t_i) + \xi_i, \qquad i = 0, 1, ..., n,$$

де $\xi_0, ..., \xi_n$ – незалежні у сукупності нормально розподілені випадкові величини з $E\xi_i = 0$ та $D\xi_i = \sigma^2$, а функцію регресії можна подати у вигляді лінійної комбінації k базисних функцій $x_0(t),\ x_1(t), ..., x_{k-1}(t)$:

$$f(t) = a_0 x_0(t) + a_1 x_1(t) + ... + a_{k-1} x_{k-1}(t).$$

Таким чином, існують справжні значення незалежних змінних $x_{0i},\ x_{1i}, ...,$ $x_{k-1,i}$ та залежної змінної $f_i,\ i = 0, 1, ..., n$, (вони детерміновані), між якими існує лінійний звязок

$$f_i = a_0 x_{0i} + a_1 x_{1i} + ... + a_{k-1} x_{k-1,i}, \qquad i = 0, 1, ..., n.$$

В класичній регресії справжні значення f_i невідомі, замість них маємо y_i, а значення $x_{0i},\ x_{1i}, ...,\ x_{k-1,i}$ спостерігаємо без помилок; параметри $a_0,\ a_1, ...,\ a_{k-1}$ невідомі, вони підлягають оцінюванню.

Поширенішим методом оцінювання невідомих параметрів в регресійному аналізі є метод найменших квадратів (МНК), який полягає в мінімізації суми

$$\sum_{i=0}^{n} \xi_i^2$$

відносно вектора $(a_0, a_1, ..., a_{k-1})$.

Позначимо

$$Y = (y_0, ..., y_n)^T;$$

$$X = \begin{pmatrix} x_{00} & x_{10} & ... & x_{k-1,0} \\ x_{01} & x_{11} & ... & x_{k-1,1} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ x_{0n} & x_{1n} & ... & x_{k-1,n} \end{pmatrix}.$$

Тоді оцінка МНК $\hat{a}_0, \hat{a}_1, ..., \hat{a}_{k-1}$ параметрів $a_0, a_1, ..., a_{k-1}$ матиме вигляд [1]

$$(\hat{a}_0, \hat{a}_1, ..., \hat{a}_{k-1})^T = (X^T X)^{-1} X^T Y.$$

Припустимо тепер, що незалежні змінні вимірюються з похибками, тобто замість $x_{0i},\ x_{1i}, ...,\ x_{k-1,i}$ спостерігаємо

$$w_{0i} = x_{0i} + \epsilon_{0i},\ w_{1i} = x_{1i} + \epsilon_{1i}, ...,\ w_{k-1,i} = x_{k-1,i} + \epsilon_{k-1,i},$$

де $\epsilon_{0i}, \epsilon_{1i}, ..., \epsilon_{k-1,i},\ i = 0, 1, ..., n$, незалежні у сукупності нормально розподілені випадкові величини з $E\epsilon_{li} = 0$ та $D\epsilon_{li} = \sigma_l^2,\ l = 0, 1, ..., k-1$, $i = 0, 1, ..., n$.

В цьому випадку для невідомих параметрів $a_0,\ a_1, ...,\ a_{k-1}$ можна також побудувати оцінку МНК:

$$(\check{a}_0, \check{a}_1, ..., \check{a}_{k-1})^T = (W_k^T W_k)^{-1} W_k^T Y,$$

де

$$W_k = \begin{pmatrix} w_{00} & w_{10} & ... & w_{k-1,0} \\ w_{01} & w_{11} & ... & w_{k-1,1} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ w_{0n} & w_{1n} & ... & w_{k-1,n} \end{pmatrix}.$$

В класичній регресії оцінка МНК є незміщеною, лінійно ефективною в класі незміщених оцінок та при деяких умовах слушною[1]. В схемі з помилками в незалежних змінних ці властивості оцінки МНК зникають.

Тому природньо в схемі

$$y_i = a_0 w_{0i} + a_1 w_{1i} + ... + a_{k-1} w_{k-1,i} + \xi_i, \qquad i = 0, 1, ..., n, \tag{1}$$

шукати інші методи оцінювання параметрів $a_0,\ a_1, ...,\ a_{k-1}$.

2. Ортогональна регресія

Принцип ортогональної регресії[2] полягає в мінімізації суми квадратів відстаней від точок $(y_i, w_{0i}, w_{1i}, ..., w_{k-1,i}),\ i = 0, 1, ..., n$, до гіперплощини

$$\prod = \{r \in R^{k+1} : \beta_0 r_0 + \beta_1 r_1 + ... + \beta_k r_k = 0\},$$

де

$$(\beta_0, \beta_1, ..., \beta_k)^T = \left(-\frac{1}{A}, \frac{a_0}{A}, ..., \frac{a_{k-1}}{A}\right)^T,$$

$$A = \sqrt{1 + a_0^2 + a_1^2 + ... + a_{k-1}^2}.$$

Оцінку ортогональної регресії $\bar{a}_0, \bar{a}_1, ..., \bar{a}_{k-1}$ параметрів $a_0, a_1, ..., a_{k-1}$ знаходимо наступним чином.

Нехай $(b_0, b_1, ..., b_k)^T-$ власний вектор, який відповідає мінімальному власному числу λ_0 матриці $(W_k^y)^T W_k^y$, де

$$W_k^y = \begin{pmatrix} y_0 & w_{00} & w_{10} & ... & w_{k-1,0} \\ y_1 & w_{01} & w_{11} & ... & w_{k-1,1} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ y_n & w_{0n} & w_{1n} & ... & w_{k-1,n} \end{pmatrix}.$$

Тоді

$$\bar{a}_0 = -b_1/b_0,\ \bar{a}_1 = -b_2/b_0, ...,\ \bar{a}_{k-1} = -b_k/b_0.$$

У випадку, коли $w_{k-1,i} = x_{k-1,i} = 1,\ i = 0, 1, ..., n$ (регресія (1) з вільним членом), доцільно перейти до центрованих рядів, тобто замість матриці W_k^y розглядати матрицю

$$\grave{W}_k^y = \begin{pmatrix} \grave{y}_0 & \grave{w}_{00} & \grave{w}_{10} & ... & \grave{w}_{k-2,0} \\ \grave{y}_1 & \grave{w}_{01} & \grave{w}_{11} & ... & \grave{w}_{k-2,1} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \grave{y}_n & \grave{w}_{0n} & \grave{w}_{1n} & ... & \grave{w}_{k-2,n} \end{pmatrix},$$

де

$$\grave{y}_i = y_i - \bar{y},\ \bar{y} = \frac{1}{n+1}\sum_{i=0}^{n} y_i,\ \grave{w}_{li} = w_{li} - \bar{w}_l,\ \bar{w}_l = \frac{1}{n+1}\sum_{i=0}^{n} w_{li},$$

$$l = 0, 1, ..., k-2,\ i = 0, 1, ..., n.$$

Позначимо $(\grave{b}_0, \grave{b}_1, ..., \grave{b}_{k-1})^T-$ власний вектор, який відповідає мінімальному власному числу $\grave{\lambda}_0$ матриці $(\grave{W}_k^y)^T \grave{W}_k^y$.

Тоді

$$\bar{a}_0 = -\grave{b}_1/\grave{b}_0,\ \bar{a}_1 = -\grave{b}_2/\grave{b}_0, ...,\ \bar{a}_{k-2} = -\grave{b}_{k-1}/\grave{b}_0, \tag{2}$$

$$\bar{a}_{k-1} = \bar{y} - \bar{a}_0\bar{w}_0 - \bar{a}_1\bar{w}_1 - ... - \bar{a}_{k-2}\bar{w}_{k-2}. \tag{3}$$

Оцінка ортогональної регресії є зміщеною, але при певних умовах слушною[1].

3. Умови співпадання оцінки МНК та оцінки ортогональної регресії в рівнянні з вільним членом.

Розглянемо регресію

$$y_i = a_0 w_{0i} + ... + a_{k-2} w_{k-2,i} + a_{k-1} + \xi_i, \qquad i = 0, 1, ..., n.$$

Теорема 1. *Оцінка МНК та оцінка ортогональної регресії параметрів $a_0,\ a_1, ...,\ a_{k-1}$ співпадають тоді і тільки тоді, коли виконується принаймні одна з двох умов:*

1)точки $(w_{0i}, w_{1i}, ..., w_{k-2,i}, y_i),\ i = 0, 1, ..., n,$ *належать одній гіперплощині;*

або

2)$\sum_{i=0}^n y_i w_{0i} = (n+1)\bar{y}\bar{w}_0, \quad \sum_{i=0}^n y_i w_{1i} = (n+1)\bar{y}\bar{w}_1, \quad ..., \quad \sum_{i=0}^n y_i w_{k-2,i} = (n+1)\bar{y}\bar{w}_{k-2}$ *одночасно.*

Доведення. Необхідність. Нехай оцінки МНК та ортогональної регресії параметрів $a_0,\ a_1, ..., a_{k-1}$ співпадають.

Оцінка МНК $\breve{a}_0,\ \breve{a}_1, ..., \breve{a}_{k-1}$ параметрів $a_0,\ a_1, ..., a_{k-1}$ задовільняє наступну систему лінійних алгебраїчних рівнянь[1]

$$\begin{pmatrix} \sum_{i=0}^n w_{0i}^2 & ... & \sum_{i=0}^n w_{0i}w_{k-2,i} & \sum_{i=0}^n w_{0i} \\ . & ... & . & . \\ . & ... & . & . \\ . & ... & . & . \\ \sum_{i=0}^n w_{k-2,i}w_{0i} & ... & \sum_{i=0}^n w_{k-2,i}^2 & n+1 \end{pmatrix} * \begin{pmatrix} \breve{a}_0 \\ . \\ . \\ . \\ \breve{a}_{k-2} \\ \breve{a}_{k-1} \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{i=0}^n y_i w_{0i} \\ . \\ . \\ . \\ \sum_{i=0}^n y_i w_{k-2,i} \\ \sum_{i=0}^n y_i \end{pmatrix}. \tag{4}$$

З останнього рівняння системи (4) маємо

$$\breve{a}_{k-1} = \bar{y} - \breve{a}_0\bar{w}_0 - ... - \breve{a}_{k-2}\bar{w}_{k-2}. \tag{5}$$

Підставляємо (5) в перші $k-1$ рівняння системи (4); отримаємо систему $(k-1)$-го порядку відносно $\breve{a}_0,\ \breve{a}_1, ..., \breve{a}_{k-2}$:

$$\begin{pmatrix} \sum_{i=0}^n w_{0i}^2 - (n+1)\bar{w}_0^2 & ... & \sum_{i=0}^n w_{0i}w_{k-2,i} - (n+1)\bar{w}_0\bar{w}_{k-2} \\ . & ... & . \\ . & ... & . \\ . & ... & . \\ \sum_{i=0}^n w_{k-2,i}w_{0i} - (n+1)\bar{w}_0\bar{w}_{k-2} & ... & \sum_{i=0}^n w_{k-2,i}^2 - (n+1)\bar{w}_{k-2}^2 \end{pmatrix} *$$

$$* \begin{pmatrix} \breve{a}_0 \\ . \\ . \\ . \\ \breve{a}_{k-2} \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^n y_i w_{0i} - (n+1)\bar{w}_0\bar{y} \\ . \\ . \\ . \\ \sum_{i=0}^n y_i w_{k-2,i} - (n+1)\bar{w}_{k-2}\bar{y} \end{pmatrix} \tag{6}$$

Згідно (2) оцінка ортогональної регресії $\bar{a}_0,\ \bar{a}_1, ..., \bar{a}_{k-2}$ параметрів $a_0,\ a_1, ..., a_{k-2}$ є розв'язок такої системи лінійних алгебраїчних рівнянь

$$\begin{pmatrix} \sum_{i=0}^{n} \grave{w}_{0i}^2 - \grave{\lambda}_0 & ... & \sum_{i=0}^{n} \grave{w}_{0i}\grave{w}_{k-2,i} \\ . & ... & . \\ . & ... & . \\ . & ... & . \\ \sum_{i=0}^{n} \grave{w}_{k-2,i}\grave{w}_{0i} & ... & \sum_{i=0}^{n} \grave{w}_{k-2,i}^2 - \grave{\lambda}_0 \end{pmatrix} * \begin{pmatrix} \breve{a}_0 \\ . \\ . \\ . \\ \breve{a}_{k-2} \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{n} \grave{y}_i\grave{w}_{0i} \\ . \\ . \\ . \\ \sum_{i=0}^{n} \grave{y}_i\grave{w}_{k-2,i} \end{pmatrix}.$$

Враховуючи

$$\sum_{i=0}^{n} \grave{w}_{l_1 i}\grave{w}_{l_2 i} = \sum_{i=0}^{n} w_{l_1 i}w_{l_2 i} - (n+1)\bar{w}_{l_1}\bar{w}_{l_2},\ l_1, l_2 = 0, 1, ..., k-2,$$

маємо

$$\begin{pmatrix} \sum_{i=0}^{n} w_{0i}^2 - (n+1)\bar{w}_0^2 - \grave{\lambda}_0 & ... & \sum_{i=0}^{n} w_{0i}w_{k-2,i} - (n+1)\bar{w}_0\bar{w}_{k-2} \\ . & ... & . \\ . & ... & . \\ . & ... & . \\ \sum_{i=0}^{n} w_{k-2,i}w_{0i} - (n+1)\bar{w}_0\bar{w}_{k-2} & ... & \sum_{i=0}^{n} w_{k-2,i}^2 - (n+1)\bar{w}_{k-2}^2 - \grave{\lambda}_0 \end{pmatrix} *$$

$$* \begin{pmatrix} \bar{a}_0 \\ . \\ . \\ . \\ \bar{a}_{k-2} \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{n} y_i w_{0i} - (n+1)\bar{w}_0\bar{y} \\ . \\ . \\ . \\ \sum_{i=0}^{n} y_i w_{k-2,i} - (n+1)\bar{w}_{k-2}\bar{y} \end{pmatrix} \tag{7}$$

З умови $\bar{a}_0 = \breve{a}_0, ..., \bar{a}_{k-2} = \breve{a}_{k-2}$ та систем рівнянь (6) та (7) отримаємо

$$\grave{\lambda}_0\bar{a}_0 = 0\ , ...,\ \grave{\lambda}_0\bar{a}_{k-2} = 0,$$

звідки випливає, що

$$\grave{\lambda}_0 = 0 \tag{8}$$

або

$$\bar{a}_0 = \bar{a}_1 = ... = \bar{a}_{k-2} = 0. \tag{9}$$

Якщо має місце (8), то це означає, що стовпці матриці $\grave{W}_k^y$ лінійно залежні (інакше матриця $(\grave{W}_k^y)^T\grave{W}_k^y$ була б достатньо визначеною і мала б тільки достатні власні значення), тобто

$$y_i - \bar{y} = C_0(w_{0i} - \bar{w}_0) + ... + C_{k-2}(w_{k-2,i} - \bar{w}_{k-2}),\ i = 0, 1, ..., n,$$

де $C_0, ..., C_{k-2}-$ деякі сталі. Отже, точки $(w_{0i}, ..., w_{k-2,i}, y_i), i = 0, 1, ..., n,$ належать одній гіперплощині

$$C_0 w_0 + C_1 w_1 + ... + C_{k-2}w_{k-2} - y = \bar{y} - C_0\bar{w}_0 - C_0\bar{w}_0 ... - C_{k-2}\bar{w}_{k-2},$$

так що виконується умова 1).

Припустимо, має місце (9). Підставимо (9) в систему (6), отримаємо умову 2).

Достатність. Нехай виконується умова 1). Це означає, що стовпці матриці $\grave{W}_k^y$ лінійно залежні, і матриця $(\grave{W}_k^y)^T\grave{W}_k^y$ має нульове власне число. Тоді системи (6) та (7) співпадуть, а значить, співпадуть і їх розв'язки:

$$\bar{a}_0 = \check{a}_0\ , ...,\ \bar{a}_{k-2} = \check{a}_{k-2}. \tag{10}$$

З (3) та (9) ураховуючи (10) отримаємо $\bar{a}_{k-1} = \check{a}_{k-1}$, тобто оцінки МНК та ортогональної регресії параметрів $a_0,\ a_1, ...,\ a_{k-1}$ в цьому випадку співпадуть.

Далі, припустимо виконується умова 2). Це означає, що системи рівнянь (6) та (7) перетворюються на однорідні, тобто вони мають тільки нульовий розв'язок. Отже,

$$\bar{a}_0 = \check{a}_0 = 0, ..., \bar{a}_{k-2} = \check{a}_{k-2} = 0. \tag{11}$$

З (3) та (9) ураховуючи (11) отримаємо $\bar{a}_{k-1} = \bar{y}$ та $\check{a}_{k-1} = \bar{y}$ відповідно, тобто

$$\bar{a}_{k-1} = \check{a}_{k-1}.$$

Таким чином, оцінки МНК та ортогональної регресії параметрів $a_0,\ a_1, ...,$ a_{k-1} і в цьому випадку співпадають.

Теорему доведено.

Висновки

В роботі доведено теорему, яка дає необхідні та достатні умови співпадання оцінки МНК та оцінки ортогональної регресії невідомих параметрів лінійної регресійної моделі з вільним членом. Умови накладено на зв'язок між залежними та незалежними змінними. Завдяки цій теоремі в деяких випадках можна замість оцінки ортогональної регресії шукати оцінку МНК навіть тоді, коли незалежні змінні вимірюються з похибками. Знайти оцінку МНК набагато легше.

Література

1. Демиденко Е.З. Линейная и нелинейная регрессии. — Москва: Финансы и статистика, 1981.—304 с.
2. Грешилов А.А., Стакун В.А., Стакун А.А. Математические методы построения прогнозов. — Москва: Радио и связь, 1997.—112 с.

Інститут матаматики НАН України, вул. Терещенківська, 3, Київ, 01601, Україна.

Надійшла 10.10.13