Журнал обчислювальної та прикладної математики | 2015, №1(118) | Journal of Computational & Applied Mathematics

УДК 519.71

КОВАРІАЦІЙНА МАТРИЦЯ ОЦІНКИ ПАРАМЕТРІВ ЛІНІЙНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ З ОБМЕЖЕННЯМИ НА ПАРАМЕТРИ

М. Ю. Савкіна

Резюме. В роботі знайдено в явному вигляді коваріаційну матрицю оцінки параметрів лінійної регресійної моделі з обмеженнями на параметри в вигляді рівностей у випадку, коли кількість обмежень на одиницю менше кількості параметрів. Невідомі параметри моделі оцінюються за методом найменших квадратів.

Abstract. Covarience matrix of equality constrained estimation of parameters of linear regression model is found in explicit form when the number of restrictions on the number one less parameters in the paper. The unknown model parameters are estimated by the method of least squares.

1. Вступ

Розглянемо модель лінійної регресії

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

де $\xi_0, ..., \xi_n$ — незалежні у сукупності випадкові величини з $E\xi_i = 0$ та $D\xi_i = \sigma^2$, а функцію регресії $f(t)$ можна подати у вигляді:

$$f(t) = a_0 x_0(t) + a_1 x_1(t) + ... + a_{k-1} x_{k-1}(t), \tag{2}$$

де $a_0,\ a_1, ...,\ a_{k-1}$ невідомі параметри, які підлягають оцінюванню; $x_0(t)$, $x_1(t), ..., x_{k-1}(t)$ — базисні функції.

Найпоширенішим 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 $\vec{a} = (a_0, a_1, ..., a_{k-1})^T$.

Позначимо

$$\vec{y} = \begin{pmatrix} y_0 \\ y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}; \qquad X = \begin{pmatrix} x_0(t_0) & x_1(t_0) & ... & x_{k-1}(t_0) \\ x_0(t_1) & x_1(t_1) & ... & x_{k-1}(t_1) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_0(t_n) & x_1(t_n) & ... & x_{k-1}(t_n) \end{pmatrix}.$$

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

$$\vec{a}^{(\wedge)} = (X^T X)^{-1} X^T \vec{y}, \tag{3}$$

де

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

а її матриця коваріацій

$$cov(\vec{a}^{(\wedge)}) = \sigma^2 (X^T X)^{-1}.$$

Теорема Гауса-Маркова [1] стверджує незміщеність та ефективність оцінки $\vec{a}^{(\wedge)}$ в класі незміщених оцінок, лінійних по $\vec{y}$.

В класичній регресії передбачається, що на вектор невідомих параметрів не накладено ніяких обмежень. Це значить, що $\Theta = R^k$, де Θ — множина апріорних значень вектора $\vec{a}$. Іноді деяка інформація все-таки існує. Апріорні обмеження можуть виступати в вигляді лінійних рівнянь або нерівностей з відомими коефіцієнтами відносно $a_0,\ a_1, ...,\ a_{k-1}$. В цьому випадку оцінка МНК буде відрізнятися від звичайної.

2. Регресія з обмеженнями на параметри в вигляді рівностей

Припустимо тепер, що апріорна множина Θ не збігається з усім простором R^k, а є лише його частиною. Розглянемо випадок, коли Θ являє собою гіперплощину в R^k розмірності $q \le k-1$. Це значить, що на вектор невідомих параметрів накладено q лінійно-незалежних обмежень у вигляді рівнянь з відомими коефіцієнтами, тобто вектор $\vec{a}$ задовільняє співвідношенню

$$R\vec{a} = \vec{r}, \tag{4}$$

де

$$R = \begin{pmatrix} r_{10} & r_{11} & \dots & r_{1,k-1} \\ r_{20} & r_{21} & \dots & r_{2,k-1} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r_{q0} & r_{q1} & \dots & r_{q,k-1} \end{pmatrix}, \quad rang\ R = q; \quad \vec{r} = \begin{pmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_q \end{pmatrix}.$$

Матриця R та вектор $\vec{r}$ задани.

Згідно з [1] оцінка, що мінімізує суму квадратів відхилень при обмеженнях на параметри в вигляді рівностей для регресії (1), (2) при обмеженнях (4) матиме вигляд

$$\vec{a}_R^{(\wedge)} = \vec{a}^{(\wedge)} + (X^T X)^{-1} R^T S (\vec{r} - R\vec{a}^{(\wedge)}), \tag{5}$$

де

$$S = \left(R (X^T X)^{-1} R^T\right)^{-1},$$

а її матриця коваріацій

$$cov(\vec{a}_R^{(\wedge)}) = \sigma^2 (X^T X)^{-1} \left[I_k - R^T S R (X^T X)^{-1}\right], \tag{6}$$

де I_k — одинична матриця $k-$ го порядку.

Незміщеність та лінійну ефективність оцінки $\vec{a}_R^{(\wedge)}$ доведено в [2].

Моделі з обмеженнями на параметри в вигляді рівностей використовуються для побудови важливих статистичних критеріїв перевірки гіпотез [2] про оцінки параметрів моделі та про модель в цілому.

В загальному випадку такий критерій перевіряє істотність обмежень на параметри вигляду (4). Гіпотеза H_0 полягає в тому, що обмеження (4) істотні. Для того, щоб прийняти або відхилити дану гіпотезу, знаходимо оцінки МНК (3) та (5) та суми квадратів відхилень

$$s_1 = (\vec{y} - X\vec{a}^{(\wedge)})^T(\vec{y} - X\vec{a}^{(\wedge)}), \quad s_2 = (\vec{y} - X\vec{a}_R^{(\wedge)})^T(\vec{y} - X\vec{a}_R^{(\wedge)}).$$

Далі, з таблиць [3] для заданого рівня значущості λ знаходимо критичне значення $F_\lambda(q, n-k-1)$ критерія Фішера, яке залежить від λ, кількості спостережень n, кількості обмежень q, кількості факторів в регресії k. Якщо

$$\frac{(s_2 - s_1)(n-k-1)}{s_1 q} < F_\lambda(q, n-k-1),$$

гіпотезу H_0 відхиляємо, інакше приймаємо.

Частинними випадками критерію перевірки істотності обмежень є:

1) перевірка сумісної значущості частини факторів в моделі (розглядається гіпотеза про те, що коефіцієнти при деяких факторах не відрізняються від нуля);

2) RESET-тест Рамсея [4] (перевіряється гіпотеза про те, що зв'язок між залежною та незалежною змінними дійсно лінійний);

3) тест Бреуша-Годфрі [4] (перевіряється гіпотеза про те, що відсутня автокореляція залишків);

4) тест Чоу [4],[5] (перевіряється гіпотеза про те, що параметри моделі сталі для всіх спостережень).

В роботі [6] знайдено в явному вигляді коваріаційну матрицю (6) у випадку $k = 3$ та $q = 1, 2$.

3. Коваріаційна матриця оцінки параметрів моделі в випадку $q = k - 1$.

Припустимо тепер, що $q = k - 1$, тобто надалі будемо вважати, що

$$R = \begin{pmatrix} r_{10} & r_{11} & \dots & r_{1,k-1} \\ r_{20} & r_{21} & \dots & r_{2,k-1} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ r_{k-1,0} & r_{k-1,1} & \dots & r_{k-1,k-1} \end{pmatrix}, \quad \vec{r} = \begin{pmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_{k-1} \end{pmatrix}.$$

Виходячи з формули (6), знайдемо матрицю коваріацій $cov(\vec{a}_R^{(\wedge)})$ для даного випадку в явному вигляді.

Подамо матрицю S^{-1} у вигляді

$$S^{-1} = \begin{pmatrix} \vec{r}_1(X^TX)^{-1}\vec{r}_0 & \vec{r}_1(X^TX)^{-1}\vec{r}_2 & ... & \vec{r}_1(X^TX)^{-1}\vec{r}_{k-1} \\ \vec{r}_2(X^TX)^{-1}\vec{r}_1 & \vec{r}_1(X^TX)^{-1}\vec{r}_2 & ... & \vec{r}_2(X^TX)^{-1}\vec{r}_{k-1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \vec{r}_{k-1}(X^TX)^{-1}\vec{r}_1 & \vec{r}_{k-1}(X^TX)^{-1}\vec{r}_2 & ... & \vec{r}_{k-1}(X^TX)^{-1}\vec{r}_{k-1} \end{pmatrix},$$

де

$$\vec{r}_i = \begin{pmatrix} r_{i0} \\ r_{i1} \\ \cdot \\ \cdot \\ \cdot \\ r_{i,k-1} \end{pmatrix}, i = 1, 2, ..., k-1.$$

Позначимо Δ та δ визначники матриць S^{-1} та X^TX відповідно.

Має місце наступна теорема.

Теорема 1. *Нехай*

$$R^{(0)} = \begin{pmatrix} r_{11} & ... & r_{1,k-1} \\ r_{21} & ... & r_{2,k-1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ r_{k-1,1} & ... & r_{k-1,k-1} \end{pmatrix}, \quad R^{(k-1)} = \begin{pmatrix} r_{10} & ... & r_{1,k-2} \\ r_{20} & ... & r_{2,k-2} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ r_{k-1,0} & ... & r_{k-1,k-2} \end{pmatrix},$$

$$R^{(m)} = \begin{pmatrix} r_{10} & ... & r_{1,m-1} & r_{1,m+1} & ... & r_{1,k-1} \\ r_{20} & ... & r_{2,m-1} & r_{2,m+1} & ... & r_{2,k-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{k-1,0} & ... & r_{k-1,m-1} & r_{k-1,m+1} & ... & r_{k-1,k-1} \end{pmatrix},$$

$$m = 1, 2, ..., k-2.$$

Тоді

$$cov(\vec{a}_R^{(\wedge)}) = (\Delta\delta)^{-1}\cdot$$

$$\cdot\begin{pmatrix} |R^{(0)}|^2 & -|R^{(0)}||R^{(1)}| & \dots & (-1)^{k-1}|R^{(0)}||R^{(k-1)}| \\ -|R^{(1)}||R^{(0)}| & |R^{(1)}|^2 & \dots & (-1)^{k}|R^{(1)}||R^{(k-1)}| \\ |R^{(2)}||R^{(0)}| & -|R^{(2)}||R^{(1)}| & \dots & (-1)^{k+1}|R^{(2)}||R^{(k-1)}| \\ \vdots & \vdots & \vdots & \vdots \\ (-1)^{k-1}|R^{(k-1)}||R^{(0)}| & (-1)^{k}|R^{(k-1)}||R^{(1)}| & \dots & |R^{(k-1)}|^2 \end{pmatrix}, \tag{7}$$

де $|R^{(i)}|,\ i = 0, 2, ..., k-1,-$ *визначник матриці* $R^{(i)}$.

Доведення. Позначимо $\vec{\gamma}^{(j)} = (\gamma_1^{(j)}, \gamma_2^{(j)}, ..., \gamma_{k-1}^{(j)}),\ j = 0, 1, ..., k-1, -$ розв'язок системи лінійних алгебраїчних рівнянь

$$S^{-1}\vec{\gamma} = \vec{r}_j^{(\perp)},$$

де

$$\vec{r}_j^{(\perp)} = \begin{pmatrix} r_{1j} \\ \vdots \\ r_{k-1,j} \end{pmatrix}, j = 0, 1, ..., k-1.$$

В цих позначеннях матрицю $R^T SR$ можна подати у вигляді

$$R^T SR = \begin{pmatrix} \sum_{i=1}^{k-1} r_{i0}\gamma_i^{(0)} & \sum_{i=1}^{k-1} r_{i0}\gamma_i^{(1)} & \dots & \sum_{i=1}^{k-1} r_{i0}\gamma_i^{(k-1)} \\ \sum_{i=1}^{k-1} r_{i1}\gamma_i^{(0)} & \sum_{i=1}^{k-1} r_{i1}\gamma_i^{(1)} & \dots & \sum_{i=1}^{k-1} r_{i1}\gamma_i^{(k-1)} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{k-1} r_{i,k-1}\gamma_i^{(0)} & \sum_{i=1}^{k-1} r_{i,k-1}\gamma_i^{(1)} & \dots & \sum_{i=1}^{k-1} r_{i,k-1}\gamma_i^{(k-1)} \end{pmatrix}.$$

Далі, елементи матриць $(X^T X)^{-1}$ та $R^T SR(X^T X)^{-1}$ позначимо $\bar{x}_{lm}$ та $\nu_{lm},\ l, m = 0, 1, ..., k-1$, відповідно.

Знайдемо $\nu_{lm},\ l, m = 0, 1, ..., k-1$. Маємо

$$\nu_{lm} = \left(\sum_{i=1}^{k-1} r_{il}\gamma_i^{(0)}\right)\bar{x}_{0m} + \left(\sum_{i=1}^{k-1} r_{il}\gamma_i^{(1)}\right)\bar{x}_{1m} + ... + \left(\sum_{i=1}^{k-1} r_{il}\gamma_i^{(k-1)}\right)\bar{x}_{k-1,m} =$$

$$= r_{1l}\left(\sum_{i=0}^{k-1}\gamma_1^{(i)}\bar{x}_{im}\right) + r_{2l}\left(\sum_{i=0}^{k-1}\gamma_2^{(i)}\bar{x}_{im}\right) + ... + r_{k-1,l}\left(\sum_{i=0}^{k-1}\gamma_{k-1}^{(i)}\bar{x}_{im}\right) =$$

$$= \Delta^{-1}\big(r_{1l}\Delta_1^{(m)} + r_{2l}\Delta_2^{(m)} + ... + r_{k-1,l}\Delta_{k-1}^{(m)}\big),$$

де $\Delta_p^{(m)},\ p = 1,2,...,k-1,-$ визначник матриці, яка утворена з матриці S^{-1} заміною $p-$го стовпця на стовпець

$$\vec{v}_m = \begin{pmatrix} v_{1m} \\ v_{2m} \\ . \\ . \\ . \\ v_{k-1,m} \end{pmatrix}, \quad v_{jm} = \sum_{i=0}^{k-1} r_{ji}\bar{x}_{im}, \quad j = 1,2,...,k-1.$$

Нехай

$$\Omega_0 = \begin{pmatrix} v_{11} & ... & r_{1,k-1} \\ v_{21} & ... & v_{2,k-1} \\ . & . & . \\ . & . & . \\ . & . & . \\ v_{k-1,1} & ... & v_{k-1,k-1} \end{pmatrix}, \quad \Omega_{k-1} = \begin{pmatrix} v_{10} & ... & v_{1,k-2} \\ v_{20} & ... & v_{2,k-2} \\ . & . & . \\ . & . & . \\ . & . & . \\ v_{k-1,0} & ... & v_{k-1,k-2} \end{pmatrix},$$

$$\Omega_p = \begin{pmatrix} v_{10} & ... & v_{1,p-1} & v_{1,p+1} & ... & v_{1,k-1} \\ v_{20} & ... & v_{2,p-1} & v_{2,p+1} & ... & v_{2,k-1} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ v_{k-1,0} & ... & v_{k-1,p-1} & r_{k-1,p+1} & ... & v_{k-1,k-1} \end{pmatrix}, \ p = 1,2,...,k-2.$$

Після низки перетворень отримаємо

$$r_{1l}\Delta_1^{(m)} + r_{2l}\Delta_2^{(m)} + ... + r_{k-1,l}\Delta_{k-1}^{(m)} = (-1)^{m+l+1}\sum_{p=0,\ p\neq m}^{k-1}\big|R_p^{(m,l)}\big| \cdot \big|\Omega_p\big|,$$

де $\big|R_p^{(m,l)}\big|$ — визначник матриці $R_p^{(m,l)}, p = 0,...,m-1,m+1,...,k-1$, яка утворена з матриці $R^{(m)}$ заміною стовпця $\vec{r}_p^{(\perp)}$ на стовпець $\vec{r}_l^{(\perp)}$, а $\big|\Omega_p\big|-$ визначник матриці $\Omega_p,\ p = 0,1,...,k-1$.

Розглянемо випадок $m \neq l$. Якщо $p \neq l$, то $\big|R_p^{(m,l)}\big| = 0$. Отже,

$$r_{1l}\Delta_1^{(m)} + r_{2l}\Delta_2^{(m)} + ... + r_{k-1,l}\Delta_{k-1}^{(m)} = (-1)^{m+l+1}\big|R_l^{(m,l)}\big| \cdot \big|\Omega_l\big|.$$

У випадку $m = l$ отримаємо

$$r_{1l}\Delta_1^{(l)} + r_{2l}\Delta_2^{(l)} + ... + r_{k-1,l}\Delta_{k-1}^{(l)} = \sum_{p=0,\ p\neq l}^{k-1} \left|R_p^{(l,l)}\right| \cdot \left|\Omega_p\right| = \Delta - \left|R_l^{(l,l)}\right| \cdot \left|\Omega_l\right|.$$

$$l = 0, 1, ..., k-1.$$

Помітимо, що $R_l^{(m,l)} = R^{(m)}$. Маємо

$$\left|\Omega_l\right| = \left|R^{(0)}\right| \cdot \bar{X}_{0l} + \left|R^{(1)}\right| \cdot \bar{X}_{1l} + ... + \left|R^{(k-1)}\right| \cdot \bar{X}_{k-1,l},$$

де $\bar{X}_{il}, i = 0, 1, ..., k-1,$ — алгебраїчне доповнення до елементу $\bar{x}_{il}$ матриці $(X^TX)^{-1}$. Позначимо через x_{lm} елементи матриці X^TX. Тоді

$$\left|\Omega_l\right| = \delta^{-1}\big((-1)^l\left|R^{(0)}\right|\cdot x_{0l} + (-1)^{l+1}\left|R^{(1)}\right|\cdot x_{1l} + ... + (-1)^{l+k-1}\left|R^{(k-1)}\right|\cdot x_{k-1,l}\big).$$

Таким чином,

$$\nu_{ll} = 1 - \Delta^{-1}\left|R^{(l)}\right| \cdot \left|\Omega_l\right|,\ l = 0, 1, ..., k-1;$$

$$\nu_{lm} = (-1)^{m+l+1}\Delta^{-1}\left|R^{(m)}\right| \cdot \left|\Omega_l\right|,\ l, m = 0, 1, ..., k-1,\ \text{якщо } m \neq l.$$

Насамкінець, знайлемо елементи матриці $(X^TX)^{-1}\big[I_k - R^TSR(X^TX)^{-1}\big]$. Позначимо їх через $\mu_{lm},\ l, m = 0, 1, ..., k-1$.

Маємо

$$\mu_{lm} = \sum_{i=0,\ i\neq m}^{k-1} \bar{x}_{li}(-\nu_{im}) + \bar{x}_{lm}(1 - \nu_{mm}) = \Delta^{-1}\left|R^{(m)}\right| \sum_{i=0}^{k-1}(-1)^{m+i}\bar{x}_{li}\left|\Omega_i\right|.$$

Далі,

$$\sum_{i=0}^{k-1}(-1)^{m+i}\bar{x}_{li}\left|\Omega_i\right| = \delta^{-1}\Bigg(\left|R^{(0)}\right| \sum_{i=0}^{k-1}(-1)^{m+2i}\bar{x}_{li}x_{0i} +$$

$$+\left|R^{(1)}\right| \sum_{i=0}^{k-1}(-1)^{m+2i+1}\bar{x}_{li}x_{1i} + ... + \left|R^{(k-1)}\right| \sum_{i=0}^{k-1}(-1)^{m+2i+k-1}\bar{x}_{li}x_{k-1,i}\Bigg) =$$

$$= \delta^{-1}(-1)^m\Bigg(\left|R^{(0)}\right| \sum_{i=0}^{k-1}\bar{x}_{li}x_{0i} - \left|R^{(1)}\right| \sum_{i=0}^{k-1}\bar{x}_{li}x_{1i} + ... - (-1)^k\left|R^{(k-1)}\right| \sum_{i=0}^{k-1}\bar{x}_{li}x_{k-1,i}\Bigg).$$

Помітимо, що сума $\sum_{i=0}^{k-1}\bar{x}_{li}x_{ji}$ є добуток l-го рядка матриці $(X^TX)^{-1}$ на j-й стовпець матриці X^TX. Оскільки $(X^TX)^{-1}X^TX = I_k$, то

$$\sum_{i=0}^{k-1}\bar{x}_{li}x_{ji} = 0,\ l, j = 0, 1, ..., k-1, j \neq l;$$

$$\sum_{i=0}^{k-1}\bar{x}_{li}x_{li} = 1,\ l = 0, 1, ..., k-1.$$

Отже,

$$\sum_{i=0}^{k-1}(-1)^{m+i}\bar{x}_{li}\big|\Omega_i\big| = \delta^{-1}(-1)^{m+l}\big|R^{(l)}\big|,$$

звідки випливає, що

$$\mu_{lm} = (-1)^{m+l}(\delta\Delta)^{-1}\big|R^{(m)}\big|\big|R^{(l)}\big|.$$

Таким чином, отримали формулу (7).

Теорему 1 доведено.

Висновки

В роботі доведено теорему, яка дає в явному вигляді коваріаційну матрицю оцінки параметрів лінійної регресійної моделі з обмеженнями на параметри в вигляді рівностей у випадку, коли кількість обмежень на одиницю менше кількості параметрів. Невідомі параметри моделі оцінюються за методом найменших квадратів. З теореми випливає, що шукана коваріаційна матриця є добуток числа $(\delta\Delta)^{-1}$ на матрицю, елементи якої залежать тільки від матриці обмежень на параметри R. Коефіцієнт $(\delta\Delta)^{-1}$ залежить від матриць R та X.

Література

1. Демиденко Е. З. Линейная и нелинейная регрессии. — Москва: Финансы и статистика, 1981. — 304 с.
2. Рао С. Р. Линейные статистические методы и их применение. — Москва: Наука, 1968.— 548 с.
3. Кендалл М. Дж., Стьюарт А. Теория распределений. — Москва: Наука, 1966.— 588 с.
4. Суслов В. И., Ибрагимов Н. М., Талышева Л. П., Цыплаков А. А. Эконометрия. — Новосибирск: Изд-во СО РАН, 2005. — 742 с.
5. Лукашин Ю. Л. Адаптивные методы краткосрочного прогнозирования временных рядов. — Москва: Финансы и статистика, 2003. — 416 с.
6. Савкіна М. Ю. Коваріаційна матриця оцінки параметрів моделі сплайнової регресії з обмеженнями на параметри в вигляді рівностей // Вісник Київського університету. — 2012. — Вип. 4. — С. 191–194.

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

Надійшла 30.08.14