Єрьоменко В. Метод умовних найменших квадратів дослідження поліноміальних регресійних моделей без вільного члена / В. Єрьоменко, О. Кочан // Вісник ТНТУ — Тернопіль : ТНТУ, 2014. — Том 74. — № 2. — С. 244-259. — (математичне моделювання. математика. фізика).

УДК 519.2

В. Єрьоменко канд. фіз.-мат. наук, О. Кочан канд. техн. наук

Тернопільський національний економічний університет

МЕТОД УМОВНИХ НАЙМЕНШИХ КВАДРАТІВ ДОСЛІДЖЕННЯ ПОЛІНОМІАЛЬНИХ РЕГРЕСІЙНИХ МОДЕЛЕЙ БЕЗ ВІЛЬНОГО ЧЛЕНА

Резюме. *Для поліноміальної регресійної моделі без вільного члена виявлені негативні наслідки безпосереднього використання МНК: сума емпіричних залишків, як правило, не дорівнює нулю, що суперечить стандартному припущенню про рівність нулю математичного сподівання збурення, а також унеможливлює так звану декомпозицію дисперсій і як наслідок – коректно визначити коефіцієнт детермінації. Встановлено, що пакети прикладних програм, зокрема всі версії Excel, не враховують вказані недоліки. На підставі модифікації МНК побудовані статистичні оцінки невідомих параметрів та досліджені їх властивості. З'ясовано, що математичне сподівання суми квадратів емпіричних залишків є дробовим числом, ціла частина якого є відповідним показником (числом ступенів вільності) для випадку МНК. Показано, що у класі лінійних незміщених оцінок не можна досягти мінімальних теоретичних дисперсій по всіх коефіцієнтах регресії одночасно, тобто відсутня рівномірна ефективність оцінок, характерна для МНК. На прикладі фіксованої вибірки вказано підхід до побудови емпіричного рівняння регресії, всі стандартні помилки коефіцієнтів регресії якого менші від відповідних показників рівняння регресії, отриманого на підставі МНК.*

Ключові слова*: МНК, регресійна модель, МНК-оцінки.*

V. Yeromenko, O. Kochan

LEAST SQUARES METHOD FOR INVESTIGATION OF POLYNOMIAL REGRESSION MODELS WITHOUT FREE TERM

Summary*. Study on the parameters of classical linear regression models is traditionally carried out using the least-squares method. The intercept (free) term is assumed to be a nonzero value. However, it is necessary to study models related to a number of important practical problems (determination using experimental data additive errors that are considered as the changes of the conversional function of devices and systems relative to some initial values under the impact of certain conditions) where free members are zeroes. The study revealed negative effects of direct using the method of least squares for polynomial regression models without an intercept term of the equation, i.e. total of empirical residuals of a model is usually not zero, which contradicts to the standard assumption of equality to zero of the residuals and makes impossible the variance decomposition and consequently resulting in the incorrect determination of the coefficient of determination. It is proved that software application packages, all versions of Excel in particular, do not take into account these deficiencies. Authors have built a statistical estimation of unknown parameters and explored their properties using the modified least-squares method, so-called conditional least-squares method. It is shown that the mathematical expectation of the sum of squares of the empirical residuals is a fractional number, the whole of which is an appropriate indicator (number of degrees of freedom) for the case of the classical least-squares method. It is shown that it is impossible to get minimal theoretical variances for all the coefficients of a linear regression simultaneously, so that there is no uniform efficiency of estimations, proper for classical least-squares method.*

There is shown an approach to constructing an empirical regression equation having all the standard errors of the regression coefficients, which are less than the corresponding parameters of the regression equation obtained using the classical least squares method for a fixed sample as an example.

Key words: *Least squares method, regression model, LSM-estimations.*

Вступ. З точки зору побудови математичних моделей технічних пристроїв і систем регресійний аналіз має ряд безумовних переваг [1]. До них відносяться, в першу чергу, добре обґрунтовані методи оцінювання якості моделей у цілому, а також індивідуального оцінювання статистичної значущості всіх коефіцієнтів моделей. Крім того, оцінюється довірчий інтервал регресійної кривої. Ці переваги зумовили широке використання регресійного аналізу, зокрема для побудови математичних моделей широкої номенклатури компонентів інформаційно-вимірювальних систем.

Одними із найпростіших є поліноміальні регресійні моделі. При цьому у ряді практично важливих випадків у таких моделях згідно з фізичним змістом вільний член дорівнює нулю [2]. На перший погляд такий випадок не заслуговує на увагу, тим більше, що у ряді пакетів прикладних програм, зокрема в Excel, реалізується дослідження як загального, так і цього випадку. Однак виявляється [2], що випадок відсутності вільного члена вимагає окремого дослідження й обережності у статистичних висновках. Зауважимо, що автори посібників, у яких наведені результати дослідження лінійних за параметрами регресійних моделей з допомогою класичного методу найменших квадратів (МНК), явно або неявно постулюють відмінність від нуля вільного члена [3-6].

1. **Об'єктом дослідження** є поліноміальна регресійна модель

$$\tilde{Y} = \alpha_1 t + \ldots + \alpha_j t^j + \ldots + \alpha_k t^k + \tilde{U}, \tag{1}$$

де $\alpha_1, \ldots, \alpha_j, \ldots, \alpha_k$ – невідомі параметри моделі; t – детермінована незалежна змінна; $\tilde{U}$ – невідома випадкова величина (збурення); $\tilde{Y}$ – результуюча (залежна) випадкова величина. Вхідною інформацією є статистичні дані вибірки обсягом $n > k$.

Метою роботи є вивчення питань:

1) негативні наслідки безпосереднього використання МНК;
2) здійснення модифікації МНК, з допомогою якої будуються статистичні оцінки невідомих параметрів та дослідження їх властивостей;
3) побудова емпіричних рівнянь регресії з мінімальними стандартними помилками коефіцієнтів регресії.

Нехай t набув значення $t_1, t_2, \ldots, t_n$ (n – обсяг вибірки). Тоді з (1) отримуються n рівнянь

$$Y_i = \sum_{j=1}^{k} \alpha_j t_i^j + U_i, \; i = \overline{1,n}, \tag{2}$$

де $U_1, U_2, \ldots, U_n$ – неспостережувані випадкові величини.

Повна специфікація моделі (2) передбачає виконання певних умов стосовно випадкових складових правих частин.

Передумова 1. Математичне сподівання всіх збурень дорівнює нулю:

$$M(U_i) = 0, \; i = \overline{1,n}. \tag{3}$$

Передумова 2. Збурення мають однакову дисперсію:

$$D(U_i) = \sigma^2, \; i = \overline{1,n}, \tag{4}$$

де σ^2 – невідоме число, яке підлягає оцінюванню.

Передумова 3. Збурення U_i та U_j при $i \neq j$ не корелюють між собою:

$$cov(U_i, U_j) = 0 \;\; \forall \, i, j = \overline{1,n}, \; i \neq j. \tag{5}$$

За конкретною вибіркою

$$(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n) \tag{6}$$

отримаємо оцінку моделі (2)

$$y_i = \sum_{j=1}^{k} a_j t_i^j + u_i, \; i = \overline{1,n}, \tag{7}$$

де $a_1, a_2, \ldots, a_k$ – оцінки невідомих параметрів $\alpha_1, \alpha_2, \ldots, \alpha_k$ відповідно; u_i – емпіричний залишок регресії або відхилення. Позначимо

$$\hat{y}_i = \sum_{j=1}^{k} a_j t_i^j, \; i = \overline{1,n}, \tag{8}$$

тоді

$$u_i = y_i - \hat{y}_i, \;\; i = \overline{1,n}, \tag{9}$$

при цьому $\hat{y}_i$ можна назвати розрахунковим або оцінювальним значенням Y_i для $t = t_i$, оскільки, згідно з (3) $M(Y_i) = \sum_{j=1}^{k} \alpha_j t_i^j$, а a_j – оцінка параметра α_j для $j = \overline{1,k}$. Емпіричні залишки визначаються неоднозначно і залежать від методу знаходження оцінок параметрів. Однак у кожному випадку повинна виконуватися рівність

$$\sum_{i=1}^{n} u_i = 0, \tag{10}$$

що випливає з (3).

Виявляється, що відсутність у моделі (1) вільного члена зумовлює некоректність використання МНК для оцінювання невідомих параметрів. Справді, згідно з МНК у відповідності з рівностями (8) та (9) мінімізується функція k змінних $Q(a_1, a_2, \ldots, a_k) = \sum_{i=1}^{n} u_i^2$, а оцінки $a_1, a_2, \ldots, a_k$ знаходяться як розв'язок системи k нормальних рівнянь

$$\left\{ \sum_{i=1}^{n} u_i \, t_i^j = 0, \;\; j = \overline{1,k}. \right. \tag{11}$$

Очевидно, що отримані МНК-оцінки в загальному випадку не задовольнятимуть рівність (10), тому ставиться під сумнів виконання передумови 1, яка використовується при доведенні незміщеності оцінок. Зауважимо, що наявність у поліноміальній моделі вільного члена $a_0 \neq 0$ зумовлює приєднання до системи (11) рівняння (10), тобто у цьому випадку МНК-оцінки вже стають незміщеними.

Наступний наслідок невиконання рівності (10) призводить до неможливості так званої декомпозиції дисперсій. Якщо виконується рівність (10), тоді правильною є рівність

$$\sigma^2_{заг} = \sigma^2_{регр} + \sigma^2_{пом}, \tag{12}$$

де

$$\sigma^2_{заг} = \sigma^2_y = n^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \text{ ; } \sigma^2_{регр} = n^{-1}\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \text{ ; }$$

$$\sigma^2_{пом} = n^{-1}\sum_{i=1}^{n}u_i^2 = n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 .$$

Якщо ж МНК-оцінки знайдені на підставі системи нормальних рівнянь (11), тоді рівність (12) вже не виконується. Справді, згідно з (9)

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_t - \bar{y})^2 + 2\sum_{i=1}^{n}(\hat{y}_1 - \bar{y})u_i + \sum_{i=1}^{n}u_i^2 ,$$

а у відповідності з (8) та (11)

$$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})u_i = \sum_{i=1}^{n}\sum_{j=1}^{k}a_j t_i^j u_i - \bar{y}\sum_{i=1}^{n}u_i = \sum_{j=1}^{k}a_j\sum_{i=1}^{n}u_i t_i^j - \bar{y}\sum_{i=1}^{n}u_i = -\bar{y}\sum_{i=1}^{n}u_i .$$

У підсумку отримаємо рівність

$$\sigma^2_{заг} = \sigma^2_{регр} - 2\,\bar{y}\,\bar{u} + \sigma^2_{пом}, \tag{13}$$

яка не дозволяє коректно визначити коефіцієнт детермінації R^2.

В якості ілюстрації викладеного розглянемо модель (1) для випадку $k = 2$ на підставі вибірки [7]

t_i	1	2	3	4	5	6	7	8
y_i	45,92	100	127,9	132,8	146	157,4	154,2	146

. (14)

Система (11) має такий вигляд:

$$\begin{cases} \left(\sum_{i=1}^{8}t_i^2\right)a_1 + \left(\sum_{i=1}^{8}t_i^3\right)a_2 = \sum_{i=1}^{8}t_i y_i , \\ \left(\sum_{i=1}^{8}t_i^3\right)a_1 + \left(\sum_{i=1}^{8}t_i^4\right)a_2 = \sum_{i=1}^{8}t_i^2 y_i . \end{cases}$$

Звідси отримаємо МНК-оцінки: $a_1 = 52{,}652803, a_2 = -4{,}366166$, на підставі яких $\sum_{i=1}^{8}u_i = 5{,}417063,$ $\sum_{i=1}^{8}u_i^2 = 387{,}487608,$ $\sigma^2_{регр} = 1337{,}482733.$ Разом з тим $\sigma^2_{заг} = \sigma^2_y = 1214{,}905 < \sigma^2_{регр}$.

Зауважимо, що згідно з пакетом Excel коефіцієнт детермінації, обчислений для вибірки (14), визначається рівністю $R^2 = 0{,}960132$, що свідчить про ігнорування рівності (13).

2. Для подолання розглянутих недоліків безпосереднього використання МНК стосовно дослідження моделі (1), розглянемо задачу знаходження умовного екстремуму

$$Q(a_1, a_2, \ldots, a_k) = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{k} a_j t_i^j \right)^2 \to min$$

за умови

$$\left(\sum_{i=1}^{n} t_i \right) a_1 + \left(\sum_{i=1}^{n} t_i^2 \right) a_2 + \ldots + \left(\sum_{i=1}^{n} t_i^k \right) a_k = \sum_{i=1}^{n} y_i .$$

Функція Лагранжа

$$L(a_1, a_2, \ldots a_k) = Q(a_1, a_2, \ldots, a_k) + \lambda \left[\left(\sum_{i=1}^{n} t_i \right) a_1 + \cdots + \left(\sum_{i=1}^{n} t_i^k \right) a_k - \sum_{i=1}^{n} y_i \right]$$

дозволяє звести цю задачу до задачі про безумовний екстремум.

Необхідні умови екстремуму функції L мають такий вигляд:

$$\begin{cases} \dfrac{\partial L}{\partial a_1} = -2 \sum\limits_{i=1}^{n} \left(y_i - \sum\limits_{j=1}^{k} a_j t_i^i \right) t + \lambda \sum\limits_{i=1}^{n} t_i = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial L}{\partial a_k} = -2 \sum\limits_{i=1}^{n} \left(y_i - \sum\limits_{j=1}^{k} a_j t_i^j \right) t_i^k + \lambda \sum\limits_{i=1}^{n} t_i^k = 0, \\ \dfrac{\partial L}{\partial \lambda} = \left(\sum\limits_{i=1}^{n} t_i \right) a_i + \cdots + \left(\sum\limits_{i=1}^{n} t_i^k \right) a_k - \sum\limits_{i=1}^{n} y_i = 0. \end{cases} \tag{15}$$

Для компактності наступних перетворень введемо позначення

$$T_m = \sum_{i=1}^{n} t_i^m, m = \overline{1,2k}; \; T_{y,m} = \sum_{i=1}^{n} t_i^m \, y_i, m = \overline{0,k}. \tag{16}$$

Тоді система рівнянь (15) набуде такого вигляду:

$$\begin{cases} T_1 a_1 + T_2 a_2 + \cdots + T_k a_k = T_{y,0}, \\ T_2 a_1 + T_3 a_2 + \cdots + T_{k+1} a_k + T_1 \dfrac{\lambda}{2} = T_{y,1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T_{k+1} a_1 + T_{k+2} a_2 + \cdots + T_{2k} a_k + T_k \dfrac{\lambda}{2} = T_{y,k}, \end{cases} \tag{17}$$

або

$$Ta^* = T_y, \tag{18}$$

де

$$T = \begin{pmatrix} T_1 & T_2 & \ldots & T_k & 0 \\ T_2 & T_3 & \ldots & T_{k+1} & T_1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ T_{k+1} & T_{k+2} & \ldots & T_{2k} & T_k \end{pmatrix}, \; T_y = \begin{pmatrix} T_{y,0} \\ T_{y,1} \\ \ldots \\ T_{y,k} \end{pmatrix}, \; a^* = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_k \\ \lambda/2 \end{pmatrix}. \tag{19}$$

Оскільки система функції $\left\{t, t^2, \ldots, t^k\right\}$ є лінійно незалежною, то матриця T є невиродженою, при цьому

$$T^{-1} = \Delta^{-1} \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1,k+1} \\ B_{21} & B_{22} & \ldots & B_{2,k+1} \\ \ldots & \ldots & \ldots & \ldots \\ B_{k+1,1} & B_{k+1,2} & \ldots & B_{k+1,k+1} \end{pmatrix}, \tag{20}$$

де

$$\Delta = \det T, \tag{21}$$

B_{ij} – алгебраїчне доповнення елемента $(.)_{ij}$ матриці

$$T' = \begin{pmatrix} T_1 & T_2 & \cdots & T_k & T_{k+1} \\ T_2 & T_3 & \cdots & T_{k+1} & T_{k+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ T_k & T_{k+1} & \cdots & T_{2k-1} & T_{2k} \\ 0 & T_1 & \cdots & T_{k-1} & T_k \end{pmatrix}, \tag{22}$$

транспонованої до T. Тому з (18) з урахуванням (19) – (20) отримаємо

$$a = BT_y \tag{23}$$

де

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_к \end{pmatrix}, \ B = \Delta^{-1} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1,k+1} \\ B_{21} & B_{22} & \cdots & B_{2,k+1} \\ \cdots & \cdots & \cdots & \cdots \\ B_{k,1} & B_{k,2} & \cdots & B_{k,k+1} \end{pmatrix}. \qquad (24)$$

Отже, точка $(a_1, a_2, \ldots, a_k)$, визначена рівністю (23), є стаціонарною. Використання критерію Рауса-Гурвіца до матриці Гесса дозволяє зробити висновок, що функція Лагранжа *L* у цій точці досягає мінімального значення.

Відзначимо, що з (20), (22) і (24) випливає рівність

$$B \begin{pmatrix} T_1 & T_2 & \cdots & T_k \\ T_2 & T_3 & \cdots & T_{k+1} \\ \cdots & \cdots & \cdots & \cdots \\ T_{k+1} & T_{k+2} & \cdots & T_{2k} \end{pmatrix} = I_k , \qquad (25)$$

де I_k – одинична матриця *k*-го порядку.

Викладений метод природно назвати методом умовних найменших квадратів (МУНК), а отримані оцінки невідомих параметрів – МУНК-оцінками.

3. Для дослідження властивостей МУНК-оцінок розглянемо випадок нефіксованої вибірки. Тоді результуюча ознака набирає значення $Y_1, Y_2, \ldots, Y_n$, які є випадковими величинами, а тому згідно з (23) і другим співвідношенням (19)

$$a = BT_Y , \qquad (26)$$

тобто МУНК-оцінки $a_1, a_2, \ldots, a_k$ невідомих параметрів $\alpha_1, \alpha_2, \ldots, \alpha_k$ є також випадковими величинами.

Позначимо

$$\tau = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \\ \cdots & \cdots & \cdots & \cdots \\ t_1^k & t_2^k & \cdots & t_n^k \end{pmatrix}, \tau_1 = \begin{pmatrix} t_1 & t_1^2 & \cdots & t_1^k \\ t_2 & t_2^2 & \cdots & t_2^k \\ \cdots & \cdots & \cdots & \cdots \\ t_n & t_n^2 & \cdots & t_n^k \end{pmatrix}, Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{pmatrix}, \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_k \end{pmatrix}.$$

Тоді з урахуванням (16) і (25)

$$\tau\tau_1 = \begin{pmatrix} T_1 & T_2 & \cdots & T_k \\ T_2 & T_3 & \cdots & T_{k+1} \\ \cdots & \cdots & \cdots & \cdots \\ T_{k+1} & T_{k+1} & \cdots & T_{2k} \end{pmatrix}, T_Y = \tau Y, \ B\tau\tau_1 = I_k , \qquad (27)$$

а модель (2) у векторно-матричній формі запишемо в такому вигляді:

$$Y = \tau_1 \alpha + U, \tag{28}$$

де $U = (U_1, U_2, \ldots, U_n)'$.

Властивість 1. МУНК-оцінки є лінійними комбінаціями спостережених значень $Y_1, Y_2, \ldots, Y_n$:

$$a_j = \sum_{i=1}^{n} \mu_i^{(j)} Y_i, \; j = \overline{1,k}, \qquad , \tag{29}$$

де вагові коефіцієнти визначаються за формулами

$$\mu_i^{(j)} = \Delta^{-1} \left(B_{j1} + B_{j2} t_i + \cdots + B_{j,k+1} t_i^k \right), \; i = \overline{1,n}, \; j = \overline{1,k}, \tag{30}$$

і задовольняють співвідношення

$$\sum_{i=1}^{n} \mu_i^{(j)} t_i^l = \delta_{jl}, \; j,l = \overline{1,k}, \tag{31}$$

$$\sum_{j=1}^{k} T_j \mu_i^{(j)} = 1, \; i = \overline{1,n}, \tag{32}$$

δ_{ij} – символ Кронекера.

Рівності (29) і (30) безпосередньо отримуються з векторної рівності (26) з урахуванням (24) та (19), а співвідношення (31) випливають з (16) і властивостей алгебраїчних доповнень B_{jl} та визначників. З першого рівняння системи (17) і співвідношень (16) та (29) отримаємо рівність відносно випадкових величин $Y_1, Y_2, \ldots, Y_n$

$$\sum_{i=1}^{n} (T_1 \mu_i^{(1)} + T_2 \mu_i^{(2)} + \ldots + T_k \mu_i^{(k)} - 1) Y_i = 0,$$

з якої з необхідністю випливають рівності (32).

Властивість 2. Якщо виконується передумова 1, тоді МУНК-оцінки $a_1, a_2 \ldots, a_k$ є незміщеними оцінками відповідних параметрів $\alpha_1, \alpha_2 \ldots, \alpha_k$, тобто

$$M(a) = \alpha. \tag{33}$$

Для доведення використаємо рівності (27) і (28)

$$a = BT_Y = B\tau Y = B\tau(\tau_1 \alpha + U) = B\tau\tau_1 \alpha + B\tau U = I_k \alpha + B\tau U = \alpha + B\tau U \tag{34}$$

З урахуванням детермінованості матриць B і τ, властивостей математичного сподівання і передумови 1

$$M(a) = M(a) + B\tau M(U) = \alpha.$$

Знайдемо дисперсійно-коваріаційну матрицю МУНК-оцінок, використавши передумови 1–3, співвідношення (33), (34) і властивості транспонування добутку матриць

$$\sum_a = M\left[(a-\alpha)(a-\alpha)'\right] = M\left[(\alpha + B\tau U - \alpha)(\alpha + B\tau U - \alpha)'\right] =$$
$$= M(B\tau UU'\tau' B') = B\tau M(UU)'\tau' B' = \sigma^2 B\tau\tau' B'.$$

Із означення матриці τ

$$\tau\tau' = \begin{pmatrix} n & T_1 & T_2 & \cdots & T_k \\ T_1 & T_2 & T_3 & \cdots & T_{k+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ T_k & T_{k+1} & T_{k+2} & \cdots & T_k \end{pmatrix},$$

а з урахуванням (24)

$$B\tau\tau' = \Delta^{-1}\begin{pmatrix} B_{11}n + \sum_{i=1}^{k} B_{1,i+1}T_i & \Delta & 0 & \cdots & 0 \\ B_{21}n + \sum_{i=1}^{k} B_{2,i+1}T_i & 0 & \Delta & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ B_{k1}n + \sum_{i=1}^{k} B_{k,i+1}T_i & 0 & 0 & \cdots & \Delta \end{pmatrix} = \Delta^{-1}\begin{pmatrix} B_{11}n & \Delta & 0 & \cdots & 0 \\ B_{21}n & 0 & \Delta & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ B_{k1}n & 0 & 0 & \cdots & \Delta \end{pmatrix}, \quad (35)$$

Оскільки згідно зі змістом B_{ij} як алгебраїчних доповнень, теоремою Лапласа про розклад визначника по рядках і властивістю визначника

$$\sum_{i=1}^{k} B_{1,i+1}T_i = \begin{vmatrix} 0 & T_1 & T_2 & \cdots & T_k \\ T_2 & T_3 & T_4 & \cdots & T_{k+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ T_k & T_{k+1} & T_{k+2} & \cdots & T_{2k} \\ 0 & T_1 & T_2 & \cdots & T_k \end{vmatrix} = 0.$$

Остаточно

$$K = B\tau\tau' B' = \begin{pmatrix} n\left(\frac{B_{11}}{\Delta}\right)^2 + \frac{B_{12}}{\Delta} & n\frac{B_{11} B_{21}}{\Delta^2} + \frac{B_{22}}{\Delta} & \cdots & n\frac{B_{11}B_{k,1}}{\Delta^2} + \frac{B_{k,2}}{\Delta} \\ & n\left(\frac{B_{21}}{\Delta}\right)^2 + \frac{B_{23}}{\Delta} & \cdots & n\frac{B_{21}B_{k,1}}{\Delta^2} + \frac{B_{k,3}}{\Delta} \\ & & \cdots & n\left(\frac{B_{k,1}}{\Delta}\right)^2 + \frac{B_{k,k+1}}{\Delta} \end{pmatrix}, \quad (36)$$

при цьому елементи під основною діагоналлю рівні відповідним елементам над нею, оскільки матриця K – симетрична.

Отже,

$$\sum\nolimits_a = \sigma^2 K, \quad (37)$$

а тому

$$\sigma_{a_1}^2 = D(a_1) = \sigma^2\left[n\left(\frac{B_{11}}{\Delta}\right)^2 + \frac{B_{12}}{\Delta}\right],\ \sigma_{a_2}^2 = D(a_2) = \sigma^2\left[n\left(\frac{B_{21}}{\Delta}\right)^2 + \frac{B_{23}}{\Delta}\right],\ldots,$$

$$\sigma_{a_k}^2 = D(a_k) = \sigma^2\left[n\left(\frac{B_{k,1}}{\Delta}\right)^2 + \frac{B_{k,k+1}}{\Delta}\right], \quad (38)$$

$cov(a_i, a_j)$ дорівнює елементу матриці K, що знаходиться в i-му рядку та j-му стовпцю, помноженому на σ^2, тобто

$$cov(a_i, a_j) = \sigma^2\left(\frac{nB_{i1} B_{j1}}{\Delta^2} + \frac{B_{j,i+1}}{\Delta}\right),\ i, j = \overline{1,k},\ i \neq j. \quad (39)$$

У співвідношеннях (37)–(39) невідомим є параметр σ^2 (дисперсія збурень), який необхідно оцінити. Позначимо $u = (u_1, u_2, \ldots u_n)'$. Тоді згідно з (7), (28) і (34) для нефіксованої вибірки

$$u = Y - \tau_1 a = \tau_1\alpha + U - \tau_1(\alpha + B\tau U) = U - \tau_1 B\tau U,\ u' = U' - U'\tau' B'\tau_1',$$
$$u'u = U'U - U'(\tau_1 B\tau + \tau' B'\tau_1')U + U'\tau' B'\tau_1'\tau_1 B\tau U. \quad (40)$$

У відповідності з передумовами 1, 2

$$M(U'U) = M\left(\sum_{i=1}^{n} U_i^2\right) = \sum_{i=1}^{n} M\left(U_i^2\right) = \sum_{i=1}^{n} M\left(U_i - 0\right)^2 = \sum_{i=1}^{n} D\left(U_i\right) = n\sigma^2. \quad (41)$$

Матриця $C = (c_{ij}) = \tau_1 B\tau + \tau' B'\tau_1'$ симетрична і детермінована, тому

$$U'CU = \sum_{i,j=1}^{n} c_{ij} U_i U_j, \; M(U'CU) = \sum_{i,j=1}^{n} c_{ij} M(U_i U_j).$$

Останню суму розкладемо на дві складові: суми елементів на головній діагоналі матриці C і поза нею:

$$M(U'CU) = \sum_{i=1}^{n} c_{ii} M(U_i^2) + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} c_{ij} M(U_i U_j).$$

Другий доданок дорівнює нулю згідно з передумовою 3. Сума діагональних елементів матриці C утворює слід матриці $tr(C)$. Але діагональні елементи матриць $\tau_1 B\tau$ і $\tau' B' \tau_1'$ відповідно рівні. Тому використавши властивості сліду матриці і останню рівність (27), отримаємо

$$\sum_{i=1}^{n} c_{ii} = tr(\tau_1 B\tau + \tau' B' \tau_1') = tr(\tau_1 B\tau) + tr(\tau' B' \tau_1') = 2tr(\tau_1 B\tau) =$$
$$= 2tr(B\tau\tau_1) = 2tr(I_k) = 2k.$$

Тобто

$$M[U'(\tau_1 B\tau + \tau' B' \tau_1')U] = 2k\sigma^2. \tag{42}$$

Аналогічно

$$M(U'\tau' B' \tau_1' \tau_1 B\tau U) = tr(\tau' B' \tau_1' \tau_1 B\tau)\sigma^2. \tag{43}$$

Використавши властивість сліду матриці, отримаємо

$$tr(\tau' B' \tau_1' \tau_1 B\tau) = tr(B' \tau_1' \tau_1 B\tau\tau') = tr(\tau\tau' B' \tau_1' \tau_1 B). \tag{44}$$

Згідно з означенням матриці τ_1 $\tau_1' \tau_1 = \begin{pmatrix} T_2 & T_3 & \cdots & T_{k+1} \\ T_2 & T_4 & \cdots & T_{k+2} \\ \cdots & \cdots & \cdots & \cdots \\ T_{k+1} & T_{k+2} & \cdots & T_{2k} \end{pmatrix}$,

тому з урахуванням (21) $T = \begin{pmatrix} T_1 \quad \cdots \quad T_k & 0 \\ & T_1 \\ \tau_1' \tau_1 & \cdots \\ & T_k \end{pmatrix}$,

а згідно з (20) і (25) $T^{-1} = \begin{pmatrix} & B & \\ \Delta^{-1} B_{k+1,1} & \cdots & \Delta^{-1} B_{k+1,k+1} \end{pmatrix}$.

Тоді рівність $TT^{-1} = I_{k+1}$ набуде такого вигляду

$$\begin{pmatrix} T_1 & \cdots & T_k & 0 \\ & & & T_1 \\ & \tau_1'\tau_1 & & \cdots \\ & & & T_k \end{pmatrix} \begin{pmatrix} & B & \\ \Delta^{-1}B_{k+1,1} & \cdots & \Delta^{-1}B_{k+1,k+1} \end{pmatrix} = I_{k+1},$$

звідки

$$\tau_1'\tau B = \Delta^{-1}\begin{pmatrix} -T_1B_{k+1,1} & 1-T_1B_{k+1,2} & -T_1B_{k+1,3} & \cdots & -T_1B_{k+1,k+1} \\ -T_2B_{k+1,1} & -T_2B_{k+1,2} & 1-T_2B_{k+1,3} & \cdots & -T_2B_{k+1,k+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -T_kB_{k+1,1} & -T_kB_{k+1,2} & -T_kB_{k+1,3} & \cdots & 1-T_kB_{k+1,k+1} \end{pmatrix}.$$

Згідно з (35) $\tau\tau' B' = \begin{pmatrix} \Delta^{-1}B_{11}n & \cdots & \Delta^{-1}B_{k1}n \\ & I_k & \end{pmatrix}$. Тому безпосередньо отримаємо

$$tr(\tau\tau' B'\tau_1'\tau_1 B) = -\frac{n}{\Delta}B_{k+1,1}(T_1B_{11} + T_2B_{21} + \cdots + T_kB_{k,1})\Delta^{-1} + 1 - \Delta^{-1}T_1B_{k+1,2} +$$
$$+1 - \Delta^{-1}T_2B_{k+1,3} + \cdots + 1 - \Delta^{-1}T_kB_{k+1,k+1} = -\Delta^{-1}B_{k+1,1}n + k - 1,$$

де враховано розклад визначника матриці T' по першому стовпцю і останньому рядку.

У підсумку із (40) з використанням (41)–(43) і детермінованості матриць τ, τ_1 і B остаточно отримаємо

$$M\left(\sum_i^n u_i^2\right) = \left[\left(1 - \frac{B_{k+1,1}}{\Delta}\right)n - k - 1\right]\sigma^2, \qquad (44)$$

де Δ визначено (21). Рівність (44) означає, що незміщена оцінка $\hat{\sigma}^2$ або вибіркова залишкова дисперсія параметра σ^2 визначається за формулою

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n u_i^2}{\left(1 - B_{k+1,1}\Delta^{-1}\right)n - k - 1}. \qquad (45)$$

Властивість 3. Якщо виконуються передумови 1–3, тоді оцінена дисперсійно-коваріаційна матриця для МУНК-оцінок $a_1, a_2, \ldots, a_k$ має такий вигляд

$$\hat{\Sigma}_a = \hat{\sigma}^2 K, \qquad (46)$$

де $\hat{\sigma}^2$ та K визначені рівностями (45) та (36) відповідно.

4. Дослідимо питання про ефективність отриманих МУНК-оцінок.

Згідно з теоремою Гаусса-Маркова всі МНК-оцінки невідомих параметрів лінійної моделі є ефективними. У випадку використання МУНК крім рівностей (31), характерних дня МНК, додатково використовуються співвідношення (32), які є наслідком примусового виконання рівності (10). Ці додаткові співвідношення стають перешкодою на шляху отримання рівномірно ефективних оцінок.

Обмежуючись випадком $k=2$, розглянемо довільні інші лінійні оцінки

$$\hat{a}_1=\sum_{i=1}^{n}x_iY_i \qquad \hat{a}_2=\sum_{i=1}^{n}z_iY_i \tag{46}$$

параметрів α_1,α_2 відповідно. Вимога незміщеності цих оцінок призводить до виконання рівностей

$$\sum_{i=1}^{n}x_it_i=1, \qquad \sum_{i=1}^{n}x_it_i^2=0, \tag{47}$$

$$\sum_{i=1}^{n}z_it_i=0, \qquad \sum_{i=1}^{n}z_it_i^2=1, \tag{48}$$

а умови (32) –

$$T_1x_i+T_2z_i=1,\ \ i=\overline{1,n}. \tag{49}$$

Зауважимо, що з (47) та (49) випливає виконання (48), а з (48) та (49) – виконання (47). Однак виконання (47) і (48) не гарантує правильність співвідношень (49).

Використавши рівності (2) і (4), отримаємо

$$\sigma_{\hat{a}_1}^2=\sigma^2\sum_{i=1}^{n}x_i^2, \quad \sigma_{\hat{a}_2}^2=\sigma^2\sum_{i=1}^{n}z_i^2. \tag{50}$$

Знайдемо незміщену точкову оцінку S^2 невідомого параметра σ^2. Позначивши $\hat{a}=(\hat{a}_1,\hat{a}_2)'$, $\alpha=(\alpha_1,\alpha_2)'$, $x=(x_1,x_2,...,x_n)$, $z=(z_1,z_2,...,z_n)$, за аналогією із виведенням рівності (45) для нефіксованої вибірки з урахуванням (46), отримаємо

$$\hat{u}=Y-\tau_1\hat{a}=\left[I_n-\tau_1\begin{pmatrix}x\\ z\end{pmatrix}\right]U, \quad \hat{u}'=U'\left[I_n-(x',z')\tau_1'\right],$$

$$M(\hat{u}'\hat{u})=M\left\{U'\left[I_n-(x',z')\tau_1'\right]\left[I_n-\tau_1\begin{pmatrix}x\\ z\end{pmatrix}\right]U\right\}=\sigma^2 tr\left\{U'\left[I_n-(x',z')\tau_1'\right]\left[I_n-\tau_1\begin{pmatrix}x\\ z\end{pmatrix}\right]U\right\}.$$

Згідно з властивостями сліду матриць і (47), (48)

$$tr[(x',z')\tau_1']=tr\left[\tau_1\begin{pmatrix}x\\ z\end{pmatrix}\right]=\sum_{i=1}^{n}\left(t_ix_i+t_i^2z_i\right)=2,$$

$$tr\left[(x',z')\tau_1'\tau_1\begin{pmatrix}x\\z\end{pmatrix}\right]=tr\left[\tau_1'\tau_1\begin{pmatrix}x\\z\end{pmatrix}(x',z')\right]=tr\left[\begin{pmatrix}T_2 & T_3\\T_3 & T_4\end{pmatrix}\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_iz_i\\ \sum_{i=1}^{n}x_iz_i & \sum_{i=1}^{n}z_i^2\end{pmatrix}\right]=$$

$$=T_2\sum_{i=1}^{n}x_i^2+2T_3\sum_{i=1}^{n}x_iz_i+T_4\sum_{i=1}^{n}z_i^2.$$

А тому остаточно

$$M(\hat{u}'\hat{u})=n-4+T_2\sum x_i^2+2T_3\sum x_iz_i+T_4\sum z_i^2=m \qquad (51)$$

і незміщена оцінка σ^2 має такий вигляд

$$S^2=\frac{\sum_{i=1}^{n}\hat{u}_i^2}{m}, \qquad (52)$$

де число m визначене рівністю (51).

Розглянемо наступні три варіанти знаходження векторів вагових коефіцієнтів x та z, які згідно з (46) визначають $\hat{a}_1$ та $\hat{a}_2$ відповідно:

I. x – розв'язок задачі $\sum_{i=1}^{n}x_i^2 \Rightarrow min$ при виконанні (47), z визначається з (49);

II. z – розв'язок задачі $\sum_{i=1}^{n}z_i^2 \Rightarrow min$ при виконанні (48), x знаходиться з (49);

III. x та z – розв'язок задачі $\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}z_i^2 \Rightarrow min$ при виконанні (47), (49).

Отримані результати на підставі вибірки (14), доповнені інформацією про МУНК-оцінки, зведені в табл.1.

Таблиця 1

Варіант	x_i	z_i	$\sum\hat{u}_i^2$	m	S_u^2
I	$\Delta_1^{-1}(T_4t_i-T_3t_i^2)$	$T_2^{-1}(1-T_1x_1)$	393,673014	6,108293	64,448943
II	$T_1^{-1}(1-T_2z_i)$	$\Delta_1^{-1}(T_3t_i-T_2t_i^2)$	392,106691	6,08087	64,482007

Закінчення **таблиці 1**

III	$\frac{T_2}{T_1^2+T_2^2}\left\{\frac{T_1}{T_2}+\Delta_1^{-1}\left[\left(T_2T_4+T_1T_3\right)t_i-\right.\right.$ $\left.\left.-T_2\left(T_1+T_3\right)\right]t_i^2\right\}$	$T_2^{-1}(1-T_1x_1)$	393,61 2601	6,057 869	64,97 5428
IV	$\mu_i^{(1)}$	$\mu_i^{(2)}$	391,40 744	6,068 627	64,49 6867

де $\Delta_1 = B_{31} = T_2T_4 - T_3^2$, $\mu_i^{(1)}$ та $\mu_i^{(2)}$ визначені (30).

Теоретичні дисперсії оцінок параметрів наведені у табл.2.

Таблиця 2

Теоретичні дисперсії оцінок	Варіанти			
	I	II	III	МУНК
$\sigma^2_{\hat{a}_1}$	$0{,}079838\sigma^2$	$0{,}080235\sigma^2$	$0{,}079839\sigma^2$	$0{,}081845\sigma^2$
$\sigma^2_{\hat{a}_2}$	$0{,}001869\sigma^2$	$0{,}001857\sigma^2$	$0{,}001863\sigma^2$	$0{,}001876\sigma^2$

У підсумку отримаємо такі емпіричні рівняння:

I: $\hat{y} = 52{,}652803t - 4{,}339612t^2$ II: $\hat{y} = 52{,}803277t - 4{,}366166t^2$,

(2,268369) (0,347071) (2,274577) (0,346012)

III: $\hat{y} = 52{,}657347t - 4{,}340414t^2$, МУНК:

$\hat{y} = 52{,}99137t - 4{,}399359t^2$,

(2,277621) (0,347892) (2,297555) (0,347845),

де в дужках вказані емпіричні середні квадратичні відхилення (стандартні помилки) коефіцієнтів регресії $\hat{a}_1$, $\hat{a}_2$, обчислені за формулою $S_{a_i} = S\sigma_{a_i}$, $i = 1{,}2$.

Згідно з табл.2 жоден із варіантів не дає найменших теоретичних дисперсій по обох параметрах, тобто відсутня рівномірна ефективність. При цьому незначне перевищення теоретичних дисперсій a_1 та a_2 для МУНК (табл.2) певною мірою компенсується мінімальністю суми квадратів емпіричних залишків у порівнянні із варіантами I-III (табл. 1). Крім того, на значення S^2 впливають числа *m* (узагальнені числа ступенів вільності).

Зменшення цього впливу слід очікувати за рахунок зменшення дробової частини чисел *m* на підставі таких значень $-\Delta^{-1}B_{k+1,1}n$ у випадку МУНК: 1,021021(*k*=3), 1,005719 (*k*=4), 1,001178 (*k*=5).

Нарешті, цікаво було б порівняти отримані результати із випадком використання МНК, ігноруючи виявлену для нього зміщеність оцінок параметрів. Згідно з МНК x_i визначається варіантом I, а z_i – варіантом II, тобто емпіричне рівняння має такий вигляд:

МНК: $\hat{y} = 52{,}652803t - 4{,}366166t^2$,

(2,270691) (0,347423),

де при обчисленні стандартних помилок враховано $m=6$, $\sum_{i=1}^{8}\hat{u}_i^2 = 387{,}4876$.

Очевидно, що останнє рівняння поступається за точністю рівнянню варіанта I.

Висновки. При дослідженні компонент інформаційно-вимірювальних систем часто використовуються поліноміальні регресійні моделі без вільного члена. Безпосереднє використання МНК для таких моделей призводить до зміщеності оцінок коефіцієнтів регресії, заниження дисперсій цих оцінок, а також недостовірність коефіцієнта детермінації як показника якості моделі. Запропонований метод умовних найменших квадратів забезпечує мінімальність суми квадратів емпіричних відхилень і рівність нулю суми цих відхилень. Однак побудовані МУНК-оцінки коефіцієнтів регресії не володіють властивістю ефективності. При цьому кожна спроба мінімізувати дисперсії оцінки призводить до збільшення суми квадратів емпіричних відхилень. Тим не менше доцільно розробити алгоритм знаходження незміщених оцінок, який би мінімізував емпіричні стандартні помилки коефіцієнтів регресії.

Наступним актуальним кроком дослідження вказаних моделей є формування статистичних висновків про значущість параметрів та моделі в цілому, виборі оптимального степеня полінома, побудова довірчих інтервалів параметрів моделі та регресійної кривої.

Conclusions. Polynomial regression models without a free member are often used for investigating of components of measurement systems. Direct use of least square method leads to bias in estimations of coefficients, understatement of estimations' deviations and inauthenticity of the coefficient of determination as an indicator of quality of a model. The proposed method of conditional least squares provides minimum sum of squares of empirical deviations and total of these deviations equals zero. But obtained estimations for this method are not effective i.e. their variances are not minimum among estimations obtained using another methods. All attempts directed to minimize variances of estimations of the proposed method lead to increase of total of squared empirical deviations. Nevertheless it is worthy developing the algorithm of finding unbiased estimations with simultaneous minimization of standard deviations of regression coefficients.

The next logical step of investigation of models mentioned above is to develop statistical tests about significance of each coefficient and a model in general, to choose the best fit polynomial, confidence intervals design for each coefficient and a regression line in general.

Список використаної літератури

1. Rawlings, J.O. Applied Regression Analysis: A Research Tool, Second Edition / J.O. Rawlings, S.G. Pantula, D.A. Dickey. – Springer-Verlag. – 1998. –678 p.
2. Yeromenko, V. The Conditional Least Squares Method for Thermocouples Error Modeling / V. Yeromenko O. Kochan // Proceedings of the 2013 IEEE 7 International Conference on Intelligent Data Acquisition and Advanced Computing Systems IDAACS'2013, September 12-14, 2013. – Berlin, Germany. – 2013. – P.157–163.
3. Назаренко, О.М. Основи економетрики. Вид. 2-ге, перероб: підручник [Текст] / О.М. Назаренко. – К: Центр навчальної літератури, 2005. – 392 с.
4. Грубер, Й. Эконометрия. – Том 1. Введение в эконометрию [Текст] / Й. Грубер. – К.: Астарта,1996. – 397 с.
5. Економетрія (економетрика) [Текст] / В.О. Єрьоменко, А.М. Алілуйко, О.М. Мартинюк, С.Ю. Попіна. – Тернопіль: Підручники і посібники, 2011. – 114 с.
6. Себер, Дж. Линейный регрессионный анализ: пер. с анг. [Текст] /Дж. Себер; пер. с англ. В.П. Носко ;под ред. М.Б. Малютова. – М.: Мир, 1980. – 456 с.
7. Рогельберг, И.Л. Изменения термоэлектрической силы проволок из хромеля и алюмеля при нагреве на воздухе при 800°С продолжительностью до 10000 ч. Том III [Текст] /И.Л. Рогельберг, Н.А. Пигидина, Э.Н. Покровская Г.Н. и др. // Сб. Исследование сплавов для термопар. – Труды института Гипроцветметобработка. – Москва: Металлургия, 1969.

Отримано 17.03.2014