Теорія ймовірностей
та математична статистика
Вип. 2(99)/2018, с. 152–167

Teoriya Ĭmovirnosteĭ
ta Matematychna Statystyka
No. 2(99)/2018, pp. 152–167

УДК 519.234

МЕТОД ОРТОГОНАЛЬНОЇ РЕГРЕСІЇ ДЛЯ СПОСТЕРЕЖЕНЬ ІЗ СУМІШІ

Р. Є. МАЙБОРОДА, Г. В. НАВАРА, О. В. СУГАКОВА

Анотація. Розглянуто узагальнення методу ортогональної регресії для оцінювання параметрів у моделі простої лінійної регресії з похибками у змінних за спостереженнями із суміші зі змінними концентраціями. Доведено консистентність та асимптотичну нормальність отриманих оцінок, знайдено матрицю розсіювання.

Ключові слова і фрази. Модель суміші зі змінними концентраціями, ортогональна регресія, метод узагальнених оцінюючих рівнянь.

2010 *Mathematics Subject Classification.* Primary 62G05, 62G20; Secondary 62J05.

1. Вступ

Моделі регресії з похибками у змінних часто застосовуються до аналізу статистичних даних астрономічних [10], радіоепідеміологічних [3] та інших прикладних досліджень. У випадку простої лінійної регресії вважається, що у кожному спостереженні вимірюються дві змінні (характеристики) об'єкта: ξ та η. При цьому між їхніми справжніми значеннями існує строга лінійна залежність

$$\eta = b_0 + b_1\xi. \tag{1}$$

Але вимірювання вносять у спостережувані значення адитивні похибки, тобто насправді спостерігаються

$$X = \xi + \delta,$$
$$Y = \eta + \varepsilon, \tag{2}$$

де ε та δ – відповідні похибки вимірювання. У класичній структурній регресійній моделі [1, розділ 1] вважається, що $\mathsf{E}\,\varepsilon = \mathsf{E}\,\delta = 0$, причому ξ, ε, δ є незалежними в сукупності випадковими величинами, а спостережувані дані $(X_j, Y_j); j = \overline{1,n}$, є незалежними однаково розподіленими копіями (2). Без додаткових припущень така модель не є ідентифіковною. Часто можна вважати, що розкид похибок вимірювання обох змінних однаковий:

$$\mathsf{D}\,\varepsilon = \mathsf{D}\,\delta. \tag{3}$$

У цьому випадку оцінки методу ортогональної регресії (методу повних квадратів) для b_0 і b_1 за незалежними однаково розподіленими спостереженнями (X_j, Y_j) є консистентними [1].

У цій роботі розглядається випадок, коли спостережувані об'єкти належать різним популяціям (компонентам), причому коефіцієнти регресії (1) можуть бути різними для різних компонентів. Справжній номер популяції, якій належать спостереження, невідомий, тому розподіл спостережуваних характеристик є сумішшю розподілів компонентів. Для класичних моделей регресії, в яких похибка регресора $\delta = 0$, параметричне оцінювання в моделі суміші регресій розглядалось у [4–6]. Непараметричний підхід до оцінювання в таких моделях на основі навантаженого методу найменших квадратів запропоновано у [8].

Ця робота присвячена узагальненню методу ортогональної регресії для моделі суміші простих лінійних регресій із похибками у змінних, за умови, що концентрації компонентів у суміші (змішуючі ймовірності) є відомими і можуть бути різними для різних спостережень (модель суміші зі змінними концентраціями [2, 7]).

Далі у розділі 2 описано формальну постановку задачі. У розділі 3 запропоновано оцінки для коефіцієнтів регресії з використанням методу навантажених моментів і доведено їх консистентність. У розділі 4 показано, що ці оцінки можна також трактувати як узагальнені оцінки навантаженого методу повних квадратів. У розділі 5 доведено асимптотичну нормальність оцінок і знайдено матрицю розсіювання. Розділ 6 присвячено результатам імітаційного моделювання. Висновки вміщено у розділі 7.

2. Постановка задачі

Отже, нехай спостережувані дані $\zeta_j = (X_j, Y_j), j = 1, \dots, n$, описуються моделлю суміші простих лінійних регресій із похибками у змінних. Опишемо формально цю модель. Ми будемо вважати, що у суміші є M компонентів і κ_j — це не спостережуваний номер компонента, якому належить j-те спостереження.

Тоді

$$X_j = \xi_j + \delta_j;$$

$$Y_j = b_0^{\kappa_j} + b_1^{\kappa_j} \xi_j + \varepsilon_j,$$

де b_0^m, b_1^m — невідомі невипадкові коефіцієнти регресії для m-го компонента суміші. Надалі будемо позначати $(\xi^m, \delta^m, \varepsilon^m)$ — вектор випадкових величин, що мають такий розподіл ξ, δ і ε, який відповідає m-му компоненту. Вважаємо, що $\xi^m, \delta^m, \varepsilon^m$ незалежні в сукупності при фіксованому m,

$$\mathsf{E}\,\delta^m = \mathsf{E}\,\varepsilon^m = 0;$$

$$\mathsf{D}\,\delta^m = \mathsf{D}\,\varepsilon^m = \sigma_m^2 > 0;\ m = \overline{1, M}. \tag{4}$$

Крім того, вважається, що $\mathsf{E}(\xi^m)^2 < \infty$.

Таким чином, спостерігаються вектори $\zeta_j, j = 1, \dots, n$, де

$$\zeta_j = \begin{pmatrix} X_j \\ Y_j \end{pmatrix} \stackrel{\mathrm{d}}{=} \zeta^m = \begin{pmatrix} \xi^m + \delta^m \\ b_0^m + b_1^m \xi^m + \varepsilon^m \end{pmatrix} \text{ за умови, що } \kappa_j = m.$$

Вектори спостережень $(X_j, Y_j), j = 1, \dots, n$, вважаються незалежними при різних j. Позначимо $p_{j;n}^m = P\{\kappa_j = m\}$ концентрацію m-го компонента у суміші під час j-го спостереження. Ці концентрації вважаються відомими.

Таким чином, розподіл спостережуваних даних являє собою суміш розподілів компонентів:

$$\mathsf{P}\{\zeta_j \in A\} = \sum_{m=1}^{M} p_{j,n}^m \,\mathsf{P}\{\zeta^m \in A\}, \quad A \in B(R).$$

Задача полягає в тому, щоб за спостереженнями $\zeta_j, j = 1, \dots, n$, оцінити коефіцієнти регресії b_0^i, b_1^i для i-го компонента суміші.

3. Оцінювання методом моментів

Для оцінювання коефіцієнтів регресії скористаємося підходом на основі навантажених моментів [2]. Для цього введемо такі позначення.

Для масивів $(p_{j;n}^m,\ j = 1, \dots, n,\ m = 1, \dots, M,\ n = 1, 2, \dots)$ концентрацій та інших вагових коефіцієнтів використовуються наступні позначення: кутовими дужками позначається операція усереднення за індексом j, тобто за всіма елементами вибірки:

$$\langle p^m \rangle_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{j=1}^{n} p_{j;n}^m.$$

При записі операцій із масивами всередині кутових дужок додавання, множення і піднесення до степеня трактують як поелементні дії:

$$\langle a^i p^m\rangle_n = \frac{1}{n}\sum_{j=1}^{n} a^i_{j;n} p^m_{j;n},\ \langle \left(a^i\right)^2\rangle_n = \frac{1}{n}\sum_{j=1}^{n}\left(a^i_{j;n}\right)^2.$$

Позначимо $\langle p^m\rangle = \lim_{n\to\infty}\langle p^m\rangle_n$, якщо ця границя існує. Помітимо, що $\langle a^i, p^m\rangle_n = \langle a^i p^m\rangle_n$ можна трактувати як скалярний добуток у R^n.

Нехай Γ_n – матриця Грама векторів концентрацій p^m_j: $\Gamma_n = \left(\langle p^i, p^m\rangle_n\right)_{i,m=1}^{M}$. Уведемо в розгляд вагові коефіцієнти a^i_j:

$$a^i_j = \frac{1}{\det\Gamma_n}\sum_{m=1}^{M}(-1)^{m+i}\gamma_{im} p^k_j, \tag{5}$$

де γ_{im} – i,m-й мінор матриці Γ_n. Для a^i_j виконується тотожність

$$\langle a^i, p^m\rangle = \frac{1}{n}\sum_{j=1}^{n} a^i_j p^m_j = I\{i=m\},\quad m=1,\dots,M. \tag{6}$$

Тут $I\{A\}$ — індикатор події A. Підсумовуючи рівність (6) по m, отримуємо, що

$$\langle a^i\rangle = \frac{1}{n}\sum_{j=1}^{n} a^i_j = 1. \tag{7}$$

Для довільної борелевої функції g позначимо

$$\hat{g}^i_n = \frac{1}{n}\sum_{j=1}^{n} a^i_j g(\zeta_j)$$

– навантажений функціональний емпіричний момент.

Теорема 3.1. [2, с. 70] або [7]. *Нехай виконані такі умови.*

1. Для всіх $m=\overline{1,M}$ $\quad \mathsf{E}\,|g(\zeta^m)| < \infty$*;*

2. Для деякого $C>0$, $\det\Gamma_n > C$ *для всіх* n.

Тоді

$$\hat{g}^i_n = \frac{1}{n}\sum_{j=1}^{n} a^i_j g(\zeta_j) \to \mathsf{E}\,g(\zeta^i);\quad n\to\infty \text{ за ймовірністю.}$$

Будемо позначати E_i, D_i, а також Cov_i умовні математичні сподівання, дисперсії і коваріації за умови, що $\kappa_j = i$. Згідно з методом моментів, прирівнюємо теоретичні моменти та їх оцінки за допомогою відповідних навантажених моментів. Використаємо такі теоретичні моменти:

$$\begin{gathered}
\mathsf{E}_i\, X_j = \mathsf{E}(\xi^i + \delta^i) = m_i;\\
\mathsf{E}_i\, Y_j = \mathsf{E}(b^i_0 + b^i_1\xi^i + \varepsilon^i) = b^i_0 + b^i_1 m_i;\\
\mathsf{D}_i\, X_j = \mathsf{D}(\xi^i + \delta^i) = \mathsf{D}\,\xi^i + \sigma^2_i;\\
\mathsf{D}_i\, Y_j = \mathsf{D}(b^i_0 + b^i_1\xi^i + \varepsilon^i) = (b^i_1)^2\,\mathsf{D}\,\xi^i + \sigma^2_i;\\
\mathrm{Cov}_i(X_j, Y_j) = \mathsf{E}(\xi^i - m_i + \delta^i)(b^i_1(\xi^i - m_i) + \varepsilon^i) = b^i_1\,\mathsf{D}\,\xi^i.
\end{gathered} \tag{8}$$

Позначимо

$$\overline{X}_i = \frac{1}{n}\sum_{j=1}^{n} a^i_j X_j, \qquad \overline{Y}_i = \frac{1}{n}\sum_{j=1}^{n} a^i_j Y_j,$$

$$\hat{D}^X_i = \frac{1}{n}\sum_{j=1}^{n} a^i_j (X_j - \overline{X}_i)^2, \qquad \hat{D}^Y_i = \frac{1}{n}\sum_{j=1}^{n} a^i_j (Y_j - \overline{Y}_i)^2,$$

$$\hat{S}^i_{XY} = \frac{1}{n}\sum_{j=1}^{n} a^i_j (X_j - \overline{X}_i)(Y_j - \overline{Y}_i).$$

Відповідно, оцінка методу моментів є розв'язком системи 5 рівнянь відносно 5 невідомих: $b^i_0, b^i_1, m_i, \mathsf{D}\,\xi^i, \sigma^2_i$.

$$\begin{cases} m_i = \overline{X}_i; \\ b^i_0 + b^i_1 m_i = \overline{Y}_i; \\ \mathsf{D}\,\xi^i + \sigma^2_i = \hat{D}^X_i; \\ (b^i_1)^2\,\mathsf{D}\,\xi^i + \sigma^2_i = \hat{D}^Y_i; \\ b^i_1\,\mathsf{D}\,\xi^i = \hat{S}^i_{XY}. \end{cases} \tag{9}$$

Віднімаючи від 4-го рівняння системи 3-тє і вилучаючи з нього $\mathsf{D}\,\xi^i$ за допомогою 5-го, одержуємо квадратне рівняння для оцінки b^i_1:

$$\hat{S}^i_{XY}(b^i_1)^2 - b^i_1(\hat{D}^Y_i - \hat{D}^X_i) - \hat{S}^i_{XY} = 0. \tag{10}$$

Коренями цього рівняння є

$$\hat{b^i_1} = \frac{\hat{D}^Y_i - \hat{D}^X_i \pm \sqrt{(\hat{D}^Y_i - \hat{D}^X_i)^2 + 4(\hat{S}^i_{XY})^2}}{2\hat{S}^i_{XY}}, \tag{11}$$

якщо $\hat{S}^i_{XY} \neq 0$.

Як ми бачимо, один із коренів має додатний чисельник, інший – від'ємний. Підставляючи (11) у 5-те рівняння системи (9), бачимо, що чисельник правої частини (11) дорівнює $\frac{2(\hat{S}^i_{XY})^2}{D\xi^i}$ – отже, він додатний. Один із коренів можна відкинути. Таким чином, маємо оцінки методу моментів для параметрів регресії:

$$\hat{b^i_1} = \frac{\hat{D}^Y_i - \hat{D}^X_i + \sqrt{(\hat{D}^Y_i - \hat{D}^X_i)^2 + 4(\hat{S}^i_{XY})^2}}{2\hat{S}^i_{XY}}; \quad \hat{b^i_0} = \overline{Y}_i - \hat{b^i_1}\overline{X}_i \tag{12}$$

Теорема 3.2. *Нехай виконуються такі умови.*

1. Для всіх $i = \overline{1, M}$: $\mathsf{E}(\xi^i)^2 < \infty$.

2. Для деякого $C > 0$, $\det \Gamma_n > C$ *для всіх* n.

3. $b^i_1 \neq 0$; $\mathsf{D}\,\xi^i \neq 0$.

Тоді оцінки параметрів регресії (12) є консистентними.

Доведення теореми випливає з теореми 3.1, оскільки при $n \to \infty$ має місце збіжність $\hat{S}^i_{XY} \to b^i_1\,\mathsf{D}\,\xi^i \neq 0$ за ймовірністю за умовами 1 і 3.

4. Узагальнений навантажений метод повних квадратів

Можливий інший підхід до задачі оцінки параметрів b^i_k, що узагальнює техніку ортогональної регресії (метод повних квадратів). Розглянемо навантажений функціонал методу повних квадратів:

$$J(b_0, b_1) = \sum_{j=1}^{n} a^i_j d^2(X_j, Y_j, b_0, b_1), \quad b_0, b_1 \in R, \tag{13}$$

де $d(X_j, Y_j, b_0, b_1)$ — довжина перпендикуляра, опущеного з точки спостереження (X_j, Y_j) на регресійну пряму $y = b_0 + b_1 x$. Оскільки

$$d = \frac{|b_1 X_j + b_0 - Y_j|}{\sqrt{1 + (b_1)^2}},$$

то функціонал набуває вигляду

$$J(b_0, b_1) = \sum_{j=1}^{n} a_j^i \frac{(b_1 X_j + b_0 - Y_j)^2}{1 + (b_1)^2}. \tag{14}$$

На роль оцінки для b_0^i, b_1^i природно обрати точку мінімуму $J(b_0, b_1)$. Дійсно, за теоремою 3.1 при $n \to \infty$

$$\frac{1}{n} \sum_{j=1}^{n} a_j^i \frac{(b_1 X_j + b_0 - Y_j)^2}{1 + (b_1)^2} \xrightarrow{\mathrm{P}} \mathsf{E}_i \left(\frac{(b_1 \xi^i + b_0 - \eta^i)^2}{1 + (b_1)^2} \right) =$$

$$= \sigma_i^2 + \frac{\mathsf{E}((b_1 - b_1^i)\xi^i + (b_0 - b_0^i))^2}{1 + (b_1)^2}$$

і мінімум правої частини досягається при $b_1 = b_1^i;\ b_0 = b_0^i$.

Узявши похідні від $J(b_0^i, b_1^i)$ по b_0^i, b_1^i, отримуємо, що в точці мінімуму J

$$\begin{cases} \sum_{j=1}^n a_j^i (b_1^i X_j^n - Y_j^n + b_0^i) = 0; \\ \sum_{j=1}^n a_j^i \frac{(b_1^i X_j^n - Y_j^n + b_0^i) X_j^n (1 + (b_1^i)^2) - b_1^i (b_1^i X_j^n - Y_j^n + b_0^i)^2}{(1 + (b_1^i)^2)^2} = 0. \end{cases} \tag{15}$$

$$\begin{cases} b_1^i \sum_{j=1}^n a_j^i X_j^n - \sum_{j=1}^n a_j^i Y_j^n + b_0^i \sum_{j=1}^n a_j^i = 0; \\ \sum_{j=1}^n a_j^i (b_1^i X_j^n - Y_j^n + b_0^i)(X_j^n + b_1^i Y_j^n) = 0. \end{cases} \tag{16}$$

Враховуючи рівність (7), із першого рівняння отримуємо

$$b_0^i = \overline{Y}_i - \hat{b_1^i} \overline{X}_i.$$

Вилучаючи b_0^i із другого рівняння, приходимо до того самого квадратного рівняння (10) відносно b_1^i:

$$((b_1^i)^2 - 1) S_{XY}^i - b_1^i (D_i^Y - D_i^X) = 0.$$

Таким чином, формальне застосування техніки методу повних квадратів приводить до тих самих оцінок, які ми отримали у розділі 3. Зрозуміло, що у загальному випадку ці оцінки не обов'язково будуть точкою мінімуму (13). Більше того, (13) може взагалі не мати точок мінімуму, оскільки вагові коефіцієнти a_j^i можуть набувати від'ємних значень. Тим не менше, оцінки $\hat{b}_k^i$ є розв'язком нормального рівняння навантаженого методу повних квадратів (15), тому їх природно назвати узагальненими оцінками повних квадратів.

5. Асимптотична нормальність оцінок коефіцієнтів регресії

Перейдемо до дослідження асимптотичної нормальності оцінок (12). Для цього скористаємось загальною теоремою про асимптотичну нормальність оцінок методу узагальнених оціночних рівнянь(УОР-оцінок) із [9].

Нехай $\{Z_j, j \geq 1\} \subset R^d$ – послідовність випадкових векторів; $\theta \in \Theta \subset R^k$ – багатовимірний параметр, що підлягає оцінюванню; $\{\psi_j(x, \gamma), j \geq 1\}$ – послідовність борелевих векторних функцій на $R^d \times \Theta$.

Позначимо

$$S_n(x, \gamma) = \sum_{j=1}^{n} \psi_j(x, \gamma); \quad \gamma \in \Theta. \tag{17}$$

Нехай $\mathsf{E}\, S_n(\theta) = \mathsf{E} \sum_{j=1}^n \psi_j(Z_j, \theta) = 0$.

Означення 5.1. Розв'язок $\hat{\theta}$ системи

$$S_n(\theta) = \sum_{j=1}^{n} \psi_j(Z_j, \theta) = 0$$

називається оцінкою методу узагальнених оціночних рівнянь(УОР-оцінкою) параметра θ, а ψ_j – оціночними функціями.

Теорема 5.1 ([9, с. 323]). *Нехай виконуються такі умови.*

1. Існує $\varphi_j(x,\gamma) = \frac{\partial \psi_j}{\partial \gamma}(x,\gamma)$.

Позначимо $\varphi_{j,m}$ *–* m*-й рядок матриці* φ_j*;*

$$h_j(Z_j) = \sup_{\gamma\in\Theta} \|\varphi_{jm}(Z_j,\gamma)\|; \quad j \geq 1; \quad \Theta - \textit{деякий замкнений окіл } \theta.$$

2. Існує такий замкнений окіл Θ *параметра* θ*, що*

$$\sup_j \mathsf{E}\,|h_j(Z_j)|^{\delta+1} < \infty \qquad i \qquad \sup_j \mathsf{E}\,\|Z_j\|^{\delta} < \infty$$

для деякого $\delta > 0$.

3. Для деякого $C > 0$ *й обмеженої ним послідовності* $\{x_j\} : \ \|x_j\| \leq C$ *послідовність функцій* $g_j(\gamma) = \varphi_{jm}(x_j,\gamma)$ *є одностайно неперервною на будь-якій відкритій підмножині* Θ.

4. $\sup_j \mathsf{E}\,\|\psi_j(Z_j,\theta)\|^{\delta+2} < \infty$ *для деякого* $\delta > 0$.

5. Існують

$$S_\infty = \lim_{n\to\infty} \frac{1}{n} \operatorname{Var} S_n(\vartheta), \ \ M = \lim_{n\to\infty} \frac{1}{n} M_n(\vartheta),$$

де $M_n(\theta) = -\mathsf{E}(\nabla S_n(\theta))$*; при цьому* $\det M \neq 0$, $\det S_\infty \neq 0$.

Якщо $\hat\theta_n$ *– консистентна послідовність УОР-оцінок, тоді*

$$\sqrt{n}(\hat\theta_n - \theta) \xrightarrow{\mathrm{W}} N(0,V); \qquad n\to\infty,$$

де

$$V = M^{-1} S_\infty (M^{-1})^T,$$

$\xrightarrow{\mathrm{W}}$ *означає слабку збіжність.*

Для застосування теореми 5.1 перетворимо (9) до вигляду

$$\begin{cases} S_1 = \sum_{j=1}^n a_j^i (X_j - m_i) = 0; \\ S_2 = \sum_{j=1}^n a_j^i (Y_j - b_0^i - b_1^i m_i) = 0; \\ S_3 = \sum_{j=1}^n a_j^i \big(((b_1^i)^2 - 1)(X_j - m_i)(Y_j - b_0^i - b_1^i m_i) - \\ \qquad\qquad - b_1^i (Y_j - b_0^i - b_1^i m_i)^2 + b_1^i (X_j - m_i)^2\big) = 0. \end{cases} \tag{18}$$

Перші два рівняння (18) – це перші два рівняння системи (9), а третє отримано з (10) заміною $\hat S^i_{XY}$, $\hat D^Y_i$, $\hat D^X_i$ на відповідні суми. Очевидно, що (12) — розв'язок системи (18), якщо з неї виключити m_i.

Отже, розглянемо невідомий тривимірний параметр $\theta_i = (m_i, b_0^i, b_1^i)^T$. В умовах теореми 3.2 розв'язок системи (18) $\hat\theta_i = (\hat m_i, \hat b_0^i, \hat b_1^i)^T$ є його консистентною оцінкою, а оціночні функції мають вигляд

$$\psi_j(x,\gamma) = \begin{pmatrix} a_j^i (x_1 - \gamma_1) \\ a_j^i (x_2 - \gamma_2 - \gamma_3\gamma_1) \\ \Psi_j^3 \end{pmatrix}, \tag{19}$$

де

$$\Psi_j^3 = a_j^i \big(((\gamma_3)^2 - 1)(x_1 - \gamma_1)(x_2 - \gamma_2 - \gamma_3\gamma_1)) - \gamma_3 (x_2 - \gamma_2 - \gamma_3\gamma_1)^2 + \gamma_3 (x_1 - \gamma_1)^2\big)$$

Уведемо деякі позначення

$$S_\infty = \lim_{n\to\infty} \frac{1}{n} \operatorname{Var} S_n(\theta);$$

$$M = \begin{pmatrix} 1 & 0 & 0 \\ b_1^i & 1 & m_i \\ 0 & 0 & -((b_1^i)^2+1)\,\mathsf{D}\,\xi^i \end{pmatrix}; \tag{20}$$

$$u_{ki} = m_k - m_i; \quad r_{ki} = b_0^k + b_1^k m_k - b_0^i - b_1^i m_i;$$

$$l_{ki} = ((b_1^i)^2 - 1)(u_{ki} b_1^k + r_{ki}) + 2b_1^i (u_{ki} - b_1^k r_{ki});$$

$$s_{ki} = (b_1^i r_{ki} + u_{ki})(b_1^i u_{ki} - r_{ki}); \quad q_{ki} = (b_1^i - b_1^k)(b_1^i b_1^k + 1);$$

$$c_{ki} = q_{ki}\,\mathsf{D}\,\xi^k + s_{ki}; \quad R_{kl} = \mathsf{E}(\xi^k - m_k)^l.$$

Також уведемо в розгляд матрицю

$$V = (M)^{-1} S_\infty \big(M^{-1}\big)^T =$$

$$= \begin{pmatrix} \mathrm{Disp}_1 & -b_1^i D_1 + C_{12} + \frac{m_i C_{13}}{K_i} & -\frac{C_{13}}{K_i} \\ -b_1^i D_1 + C_{12} + \frac{m_i C_{13}}{K_i} & \mathrm{Disp}_2 & \frac{b_1^i C_{13} - C_{23}}{K_i} - \frac{m_i D_3}{K_i^2} \\ -\frac{C_{13}}{K_i} & \frac{b_1^i C_{13} - C_{23}}{K_i} - \frac{m_i D_3}{K_i^2} & \mathrm{Disp}_3 \end{pmatrix};$$

тут $\mathrm{Disp}_1 = D_1$; $\mathrm{Disp}_2 = (b_1^i)^2 D_1 + D_2 - 2b_1^i C_{12} + \dfrac{m_i^2 D_3 + 2m_i(C_{23} - b_1^i C_{13})K_i}{K_i^2}$;

$$\mathrm{Disp}_3 = \frac{D_3}{K_i^2}; \quad K_i = ((b_1^i)^2 + 1)\,\mathsf{D}\,\xi^i,$$

а коефіцієнти $D_1, D_2, D_3, C_{12}, C_{23}, C_{13}$ задаються формулами

$$D_1 = \sum_{k=1}^{M} \big(\langle (a^i)^2, p^k \rangle_\infty (\mathsf{E}(\xi^k)^2 + \sigma_k^2) - \langle (a^i)^2, (p^k)^2 \rangle_\infty m_k^2\big) -$$

$$-2 \sum_{k,l=1;k\neq l}^{M} \langle (a^i)^2, p^k p^l \rangle_\infty m_k m_l;$$

$$D_2 = \sum_{k=1}^{M} \big(\langle (a^i)^2, p^k \rangle_\infty ((b_1^k)^2\,\mathsf{D}\,\xi^k + \sigma_k^2 + (b_1^k m_k + b_0^k)^2)) -$$

$$-\langle (a^i)^2, (p^k)^2 \rangle_\infty (b_1^k m_k + b_0^k)^2\big) - 2 \sum_{k,l=1;k\neq l}^{M} \langle (a^i)^2, p^k p^l \rangle_\infty (b_1^k m_k + b_0^k)(b_1^l m_l + b_0^l);$$

$$D_3 = \sum_{k=1}^{M} \big(\langle (a^i)^2, p^k \rangle_\infty A - \langle (a^i)^2, (p^k)^2 \rangle_\infty B\big) -$$

$$-\sum_{k=1}^{M} \langle (a^i)^2, (p^k)^2 \rangle_\infty c_{ki}^2 - \frac{2}{n} \sum_{k,l=1;k\neq l}^{M} \langle (a^i)^2, p^k p^l \rangle_\infty c_{ki} c_{li};$$

$$A = ((b_1^i)^2 + 1)\,\mathsf{D}\,\xi^i \sigma_i^2 + ((b_1^i)^4 - 4(b_1^i)^2 + 1)\sigma_i^4 + (b_1^i)^2(\mathsf{E}(\delta^i)^4 + \mathsf{E}(\varepsilon^i)^4);$$

$$B = R_{k4} q_{ki}^2 + s_{ki}^2 (l_{ki}^2 + 2q_{ki} s_{ki}) R_{k2} + ((b_1^i)^2 + 1)^2 ((b_1^k)^2 + 1)\sigma_k^2 R_{k2} + (b_1^i)^2(\mathsf{E}(\delta^i)^2 +$$
$$+ \mathsf{E}(\varepsilon^i)^2) + \sigma_k^4((b_1^i)^4 - 4(b_1^i)^2 + 1) + \sigma_k^2(r_{ki}^2((b_1^i)^2 - 1)^2 + (((b_1^i)^2 - 1)u_{ki} - 2r_{ki} b_1^i)^2) +$$
$$+2q_{ki} l_{ki} R_{k3} + 2b_1^i((b_1^i)^2 - 1) r_{ki}\,\mathsf{E}(\delta^k)^3 - 2b_1^i(((b_1^i)^2 - 1)u_{ki} - 2r_{ki} b_1^i)\,\mathsf{E}(\varepsilon^k)^3\big);$$

$$C_{12} = \sum_{k=1}^{M} \langle (a^i)^2, p^k \rangle_\infty (b_1^k R_{k2} + u_{ki} r_{ki}) - \sum_{k,m=1}^{M} \langle (a^i)^2, p^k p^m \rangle_\infty b_1^m u_{ki} r_{mi};$$

$$C_{13} = \sum_{k=1}^{M} \langle (a^i)^2, p^k \rangle_\infty \big(q_{ki} R_{k3} + (((b_1^i)^2 - 1)(2b_1^k u_{ki} + r_{ki}) +$$
$$+b_1^i(-2b_1^k r_{ki} - (b_1^k)^2 u_{ki} + 3u_{ki}))R_{k2} + u_{ki} s_{ki} +$$
$$+\sigma_k^2(((b_1^i)^2 - 1) r_{ki} + 2b_1^i u_{ki}) - b_1^i\,\mathsf{E}(\delta^k)^3\big) -$$

$$-\sum_{k,m=1}^{M}\langle (a^i)^2, p^k p^m\rangle_\infty u_{ki} c_{mi};$$

$$C_{23}=\sum_{k=1}^{M}\langle (a^i)^2,p^k\rangle_\infty \big(b_1^k q_{ki}R_{k3}+R_{k2}(((b_1^i)^2-1)(2b_1^k r_{ki}+u_{ki}(b_1^k)^2)+$$
$$+b_1^i(2b_1^k u_{ki}-3(b_1^k)^2 r_{ki}+r_{ki}))+r_{ki}s_{ki}+(((b_1^i)^2-1)u_{ki}-2b_1^i r_{ki})\sigma_k^2+b_1^i\,\mathsf{E}(\delta^k)^3\big)-$$
$$-\sum_{k,m=1}^{M}\langle (a^i)^2, p^k p^m\rangle_\infty r_{ki} c_{mi}. \tag{21}$$

Теорема 5.2. *Нехай виконуються такі умови:*

1. Для деякого $C>0$, $\det\Gamma_n > C$ *для всіх* n.

2. Існують $\mathsf{E}(\xi^i)^{12}<\infty$; $\mathsf{E}(\delta^i)^{12}<\infty$. $\mathsf{E}(\varepsilon^i)^{12}<\infty$; $\forall i=\overline{1,M}$.

3. Для вагових коефіцієнтів a_j^i *і концентрацій* p_j^k *існують границі:*

$$\lim_{n\to\infty}\langle (a^i)^2,p^k p^l\rangle_n<\infty \quad \forall k,l=\overline{1,M}.$$

4. $\det S_\infty\neq 0$.

5. $\mathsf{D}\,\xi^i\neq 0$, $b_1^i\neq 0$.

Тоді

$$\sqrt{n}(\hat\theta_i-\theta_i)\xrightarrow{\mathrm{W}} N(0,V);\quad n\to\infty,$$

де матриця V *задається формулою (21).*

Доведення. Як ми показали, оцінка $\hat\theta_i=(\hat m_i,\hat b_0^i,\hat b_1^i)^T$ є розв'язком системи

$$S_n(\gamma)=\sum_{j=1}^{n}\psi_j(\zeta_j,\gamma)=0;$$

де ψ_j задаються (19).

Перевіримо виконання умов теореми 5.1. Умова 1, вочевидь, виконується:

$$\varphi_j(x,\gamma)=\begin{pmatrix}-1 & 0 & 0\\ -\gamma_3 & -1 & -\gamma_1\\ a & b & c\end{pmatrix}, \tag{22}$$

$$\text{де}\qquad a=a_j^i((\gamma_3)^2+1)(Y_j-\gamma_2-\gamma_3X_j);$$
$$b=a_j^i(2\gamma_3(Y_j-\gamma_2)-((\gamma_3)^2-1)(X_j+\gamma_1));$$
$$c=a_j^i\big(2\gamma_3X_j(Y_j-\gamma_2-\gamma_3\gamma_1)-\gamma_1((\gamma_3)^2-1)(X_j-\gamma_1)-$$
$$-(Y_j-\gamma_2-\gamma_3\gamma_1)^2+(X_j-\gamma_1)^2\big).$$

Перейдемо до умови 2, поклавши $\delta=4$. Тоді

$$\sup_j \mathsf{E}\,\|\zeta_j\|^4=\sup_j \mathsf{E}(X_j^2+Y_j^2)^2=\sup_j\sum_{k=1}^{M}a_j^k\,\mathsf{E}\big((\xi^k+\delta^k)^2+(b_1^k\xi^k+b_0^k+\varepsilon^k)^2\big)^2.$$

За нерівністю

$$|a+b|^p\le 2^{p-1}(|a|^p+|b|^p),\quad p\ge 2, \tag{23}$$

бачимо, що для існування $\sup_j \mathsf{E}\,\|\zeta_j\|^4<\infty$ достатньо існування математичних сподівань $\mathsf{E}(\xi^i)^4<\infty$; $\mathsf{E}(\delta^i)^4<\infty$; $\mathsf{E}(\varepsilon^i)^4<\infty$ та виконання умови $\sup_j|a_j^k|<\infty$, а це випливає з умов 1, 2 теореми.

Тепер візьмемо довільний скінченний окіл параметра θ вигляду

$$\Theta=\{r_1\le m_i\le r_2; p_1\le b_1^i\le p_2; q_1\le b_0^i\le q_2\}.$$

Розглянемо
$$\sup_j \mathsf{E} \,|\sup_{\gamma\in\Theta} \|\varphi_{jm}(\zeta_j,\gamma)\||^5, \tag{24}$$
де φ_{jm} – рядки матриці $\varphi_j(x,\gamma)$, що задається (22). Оскільки перші два рядки взагалі є сталими, вираз (24) у цьому випадку є скінченним. Розглянемо третій рядок. Надалі будемо позначати $c_1, c_2, \dots$ різні додатні сталі.
$$A = \sup_{\gamma\in\Theta} \|\varphi_{j3}(\zeta_j,\gamma)\| = \sup_{\gamma\in\Theta}\sup_{j,i} |a_j^i| \big((\gamma_3)^2+1)^2(Y_j-\gamma_2-\gamma_3X_j)^2 + (2\gamma_3(Y_j-\gamma_2) -$$
$$-((\gamma_3)^2-1)(X_j+\gamma_1))^2 + (2\gamma_3X_j(Y_j-\gamma_2-\gamma_3\gamma_1) - \gamma_1((\gamma_3)^2-1)(X_j-\gamma_1) -$$
$$-(Y_j-\gamma_2-\gamma_3\gamma_1)^2 + (X_j-\gamma_1)^2)^2\big)^{1/2} \le$$
$$\le c_1\big(c_2 + c_3\xi_j^2 + c_4\delta_j^2 + c_5\varepsilon_j^2 + c_6\xi_j^4 + c_7\delta_j^4 + c_8\varepsilon_j^4\big)^{1/2}.$$
Остання нерівність випливає із (23); сталі залежать від $r_1,\dots,q_2$. Далі скористаємось нерівністю Йєнсена
$$(\mathsf{E}\,|\xi|^v)^{\frac{1}{v}} \le (\mathsf{E}\,|\xi|^u)^{\frac{1}{u}}, \text{ якщо } v<u. \tag{25}$$
$$\sup_j \mathsf{E}\, A^5 \le \sup_j c_1\, \mathsf{E}\big(c_2 + c_3\xi_j^2 + c_4\delta_j^2 + c_5\varepsilon_j^2 + c_6\xi_j^4 + c_7\delta_j^4 + c_8\varepsilon_j^4\big)^{5/2} \le$$
$$\le c_9 \sup_j \sum_{i=1}^M p_j^i \big(\mathsf{E}(c_2 + c_3(\xi^i)^2 + c_4(\delta^i)^2 + c_5(\varepsilon^i)^2 + c_6(\xi^i)^4 + c_7(\delta^i)^4 + c_8(\varepsilon^i)^4)^3\big)^{\frac{2}{15}}.$$
Бачимо, що цей вираз буде скінченний при виконанні умови 2 теореми.

Умова 3 теореми 5.1 виконана, оскільки у нашому випадку $\varphi_{jm}(x_j,\gamma)$ є поліномами відносно аргументів γ та x_j, отже, ця система функцій є одностайно неперервною на компакті.

Перевіримо умову 4. Покладемо $\delta = 2$.
$$\sup_j \mathsf{E}\,\|\psi_j(\zeta_j,\theta)\|^4 = \sup_j (a_j^i)^2\, \mathsf{E}\big((X_j-m_i)^2 + (Y_j-b_0^i-b_1^im_i)^2 +$$
$$+(((b_1^i)^2-1)(X_j-m_i)(Y_j-b_0^i-b_1^im_i) - b_1^i(Y_j-b_0^i-b_1^im_i)^2 + b_1^i(X_j-m_i)^2)^2\big)^2.$$
Очевидно, що при існуванні 12-го моменту у $\xi^i, \delta^i, \varepsilon^i$ цей вираз буде скінченним.

Перейдемо до умови 5 теореми 5.1.

При знаходженні матриці $\frac{M_n(\theta_i)}{n} = -\frac{1}{n}\mathsf{E}(\nabla S_n(\theta_i))$, де S_n задаються формулами (18), обчислення скрізь тривіальні, питання можуть виникнути лише при знаходженні $\frac{1}{n}\mathsf{E}\frac{\partial S_3}{\partial b_1^i}$. Використовуючи (6), маємо
$$\frac{1}{n}\mathsf{E}\frac{\partial S_3}{\partial b_1^i} = \frac{1}{n}\sum_{j=1}^n a_j^i \sum_{k=1}^M p_j^k\, \mathsf{E}\big(2b_1^i(\xi^k+\delta^k-m_i)(b_1^k\xi^k + b_0^k + \varepsilon^k - b_0^i - b_1^im_i) -$$
$$-(b_1^k\eta_k + b_0^k + \varepsilon^k - b_0^i - b_1^im_i)^2 + (\eta_k+\delta^k-m_i)^2\big) =$$
$$= \mathsf{E}\big(2b_1^i(\xi^i-m_i+\delta^i)(b_1^i(\xi^i-m_i)+\varepsilon^i) - (b_1^i(\xi^i-m_i)+\varepsilon^i)^2 + (\xi^i-m_i+\delta^i)^2\big) =$$
$$= ((b_1^i)^2+1)D\xi^i.$$
Отже ми довели, що матриця M визначається (20). Очевидно, що з умови 5 теореми випливає $\det M = -((b_1^i)^2+1)\,\mathsf{D}\,\xi^i \ne 0$.

Тепер випишемо граничну матрицю V з формулювання теореми 5.2. Знаходження матриці $\frac{1}{n}\operatorname{Var}(S_n(\theta_i))$ вимагає більш громіздких обчислень, ніж $M_n(\theta_i)$.

Позначимо елементи матриці:
$$\frac{1}{n}\operatorname{Var}(S_n(\theta_i)) = \begin{pmatrix} \widetilde{D}_1 & \widetilde{C}_{12} & \widetilde{C}_{13} \\ \widetilde{C}_{12} & \widetilde{D}_2 & \widetilde{C}_{23} \\ \widetilde{C}_{13} & \widetilde{C}_{23} & \widetilde{D}_3 \end{pmatrix}, \tag{26}$$

де

$$\widetilde{D}_1 = \frac{\mathsf{D}\, S_1}{n} = \frac{1}{n}\,\mathsf{D}\sum_{j=1}^{n} a_j^i (X_j - m_i) =$$

$$= \frac{1}{n}\sum_{j=1}^{n} (a_j^i)^2 \big(\mathsf{E}(\xi_j + \delta_j)^2 - (\mathsf{E}(\xi_j+\delta_j))^2\big) = A - B,$$

$$A = \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{M} (a_j^i)^2 p_j^k (\mathsf{E}(\xi^k)^2 + \sigma_k^2) = \sum_{k=1}^{M} \langle (a^i)^2, p^k\rangle_n (\mathsf{E}(\xi^k)^2 + \sigma_k^2),$$

$$B = \frac{1}{n}\sum_{j=1}^{n} (a_j^i)^2 \left(\sum_{k=1}^{M} p_j^k m_k\right)^2 = \sum_{k=1}^{M} \langle (a^i)^2, (p^k)^2\rangle_n m_k^2 + 2\sum_{k,l=1,k\neq l}^{M} \langle (a^i)^2, p^k p^l\rangle_n m_k m_l.$$

$$\widetilde{D}_1 = \sum_{k=1}^{M}\big(\langle (a^i)^2, p^k\rangle_n (\mathsf{E}(\xi^k)^2 + \sigma_k^2) - \langle (a^i)^2, (p^k)^2\rangle_n m_k^2\big) -$$

$$-2\sum_{k,l=1;k\neq l}^{M} \langle (a^i)^2, p^k p^l\rangle_n m_k m_l. \tag{27}$$

Аналогічно

$$\widetilde{D}_2 = \frac{\mathsf{D}\, S_2}{n} = \frac{1}{n^2}\sum_{j=1}^{n} (a_j^i)^2 \big(\mathsf{E}\, Y_j^2 - (\mathsf{E}\, Y_j)^2\big) = A - B,$$

де

$$A = \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{M} (a_j^i)^2 p_j^k\, \mathsf{E}(b_1^k \xi^k + b_0^k + \varepsilon^k)^2 =$$

$$= \sum_{k=1}^{M} \langle (a^i)^2, p^k\rangle_n ((b_i^k)^2\, \mathsf{E}(\xi^k)^2 + (b_0^k)^2 + \sigma_k^2 + 2 b_1^k b_0^k m_k);$$

$$B = \sum_{k=1}^{M} \langle (a^i)^2, (p^k)^2\rangle_n (b_1^k m_k + b_0^k)^2 + 2\sum_{k,l=1;k\neq l}^{M} \langle (a^i)^2, p^k p^l\rangle_n (b_1^k m_k + b_0^k)(b_1^l m_l + b_0^l).$$

$$\widetilde{D}_2 = \sum_{k=1}^{M}\big(\langle (a^i)^2, p^k\rangle_n ((b_1^k)^2\, \mathsf{D}\,\xi^k + \sigma_k^2 + (b_1^k m_k + b_0^k)^2)) -$$

$$-\langle (a^i)^2, (p^k)^2\rangle_n (b_1^k m_k + b_0^k)^2\big) - 2\sum_{k,l=1;k\neq l}^{M} \langle (a^i)^2, p^k p^l\rangle_n (b_1^k m_k + b_0^k)(b_1^l m_l + b_0^l). \tag{28}$$

Тепер знайдемо $\widetilde{D}_3$.

$$\widetilde{D}_3 = \frac{\mathsf{D}\, S_3}{n} = \frac{1}{n}\sum_{j=1}^{n} (a_j^i)^2\, \mathsf{D}\big((b_1^i)^2 - 1)(X_j - m_i)(Y_j - b_0^i - b_1^i m_i) +$$

$$+ b_i^i (X_j - m_i)^2 - b_1^i (Y_j - b_0^i - b_1^i m_i)^2\big) = \frac{1}{n}\sum_{j=1}^{n} (a_j^i)^2 \big(\mathsf{E}\,\mu^2 - (\mathsf{E}\,\mu)^2\big). \tag{29}$$

Знайдемо

$$\mathsf{E}\,\mu^2 = p_j^i\, \mathsf{E}\big((b_1^i)^2 - 1)(\xi^i - m_i + \delta^i)(b_1^i(\xi^i - m_i) + \varepsilon^i) + b_1^i(\xi^i - m_i + \delta^i)^2 -$$

$$-b_1^i (b_1^i(\xi^i - m_i) + \varepsilon^i)^2\big)^2 + \sum_{k=1;k\neq i}^{M} p_j^k\, \mathsf{E}\big((b_1^i)^2 - 1)(\xi^k - m_i + \delta^k)(b_1^k \xi^k + b_0^k + \varepsilon^k - b_0^i -$$

$$-b_1^i m_i) + b_1^i(\xi^k - m_i + \delta^k)^2 - b_1^i(b_1^k\xi^k + b_0^k + \varepsilon^k - b_0^i - b_1^i m_i)^2)\big)^2 = p_j^i A + \sum_{k=1;k\neq i}^{M} p_j^k B_k, \tag{30}$$

де

$$A = \mathsf{E}\big(b_1^i((b_1^i)^2+1)\delta^i(\xi^i - m_i) - ((b_1^i)^2+1)(\xi^i - m_i)\varepsilon^i + ((b_1^i)^2-1)\delta^i\varepsilon^i + b_1^i(\delta^i)^2 -$$
$$-b_1^i(\varepsilon^i)^2\big)^2 = ((b_1^i)^2+1)\,\mathsf{D}\,\xi^i\sigma_i^2 + ((b_1^i)^4 - 4(b_1^i)^2+1)\sigma_i^4 + (b_1^i)^2(\mathsf{E}(\delta^i)^4 + \mathsf{E}(\varepsilon^i)^4), \tag{31}$$
$$B_k = \mathsf{E}\big(((b_1^i)^2-1)(\xi^k - m_k + m_k - m_i + \delta^k)(b_1^k(\xi^k - m_k) + b_0^k + b_1^k m_k -$$
$$-b_0^i - b_1^i m_i + \varepsilon^k) + b_1^i(\xi^k - m_k + m_k - m_i + \delta^k)^2 -$$
$$-b_1^i(b_1^k(\xi^k - m_k) + b_0^k + b_1^k m_k - b_0^i - b_1^i m_i + \varepsilon^k)^2\big)^2.$$

Уводячи позначення:

$$u_{ki} = m_k - m_i; \quad r_{ki} = b_0^k + b_1^k m_k - b_0^i - b_1^i m_i;$$
$$l_{ki} = ((b_1^i)^2 - 1)(u_{ki}b_1^k + r_{ki}) + 2b_1^i(u_{ki} - b_1^k r_{ki});$$
$$s_{ki} = (b_1^i r_{ki} + u_{ki})(b_1^i u_{ki} - r_{ki}); \quad q_{ki} = (b_1^i - b_1^k)(b_1^i b_1^k + 1);$$
$$c_{ki} = q_{ki}\,\mathsf{D}\,\xi^k + s_{ki};$$
$$R_{kl} = \mathsf{E}(\xi^k - m_k)^l, \tag{32}$$

маємо такий вираз для B_k:

$$B_k = R_{k4}q_{ki}^2 + s_{ki}^2(l_{ki}^2 + 2q_{ki}s_{ki})R_{k2} + ((b_1^i)^2+1)^2((b_1^k)^2+1)\sigma_k^2 R_{k2} + (b_1^i)^2(\mathsf{E}(\delta^i)^2 + \mathsf{E}(\varepsilon^i)^2) +$$
$$+\sigma_k^4((b_1^i)^4 - 4(b_1^i)^2 + 1) + \sigma_k^2(r_{ki}^2((b_1^i)^2-1)^2 + (((b_1^i)^2-1)u_{ki} - 2r_{ki}b_1^i)^2) +$$
$$+2q_{ki}l_{ki}R_{k3} + 2b_1^i((b_1^i)^2-1)r_{ki}\,\mathsf{E}(\delta^k)^3 - 2b_1^i(((b_1^i)^2-1)u_{ki} - 2r_{ki}b_1^i)\,\mathsf{E}(\varepsilon^k)^3. \tag{33}$$

Тепер розглянемо другу частину (29):

$$(\mathsf{E}\,\mu)^2 = \left(\sum_{k=1;k\neq i}^{M} p_j^k\big((b_1^i - b_1^k)(b_1^i b_1^k + 1)\,\mathsf{D}\,\xi^k + (b_1^i r_{ki} + u_{ki})(b_1^i u_{ki} - r_{ki})\big)\right)^2 =$$
$$= \sum_{k=1}^{M}(p_j^k)^2(q_{ki}\,\mathsf{D}\,\xi^k + s_{ki})^2 + 2\sum_{k,l=1;k\neq l}^{M} p_j^k p_j^l(q_{ki}\,\mathsf{D}\,\xi^k + s_{ki})(q_{li}\,\mathsf{D}\,\xi^l + s_{li}). \tag{34}$$

Об'єднуючи (29)–(34), отримуємо

$$\widetilde{D}_3 = \sum_{k=1}^{M}\big(\langle(a^i)^2, p^k\rangle_n A - \langle(a^i)^2, (p^k)^2\rangle_n B\big) -$$
$$-\sum_{k=1}^{M}\langle(a^i)^2, (p^k)^2\rangle_n c_{ki}^2 - 2\sum_{k,l=1;k\neq l}^{M}\langle(a^i)^2, p^k p^l\rangle_n c_{ki}c_{li}; \tag{35}$$

A, B задаються (31), (33). Тепер випишемо $\widetilde{C}_{12}$.

$$\widetilde{C}_{12} = \mathrm{Cov}\left(\frac{S_1}{\sqrt{n}}, \frac{S_2}{\sqrt{n}}\right) = \frac{1}{n}\sum_{j=1}^{n}(a_j^i)^2\,\mathsf{E}(X_j - m_i)(Y_j - b_0^i - b_1^i m_i) +$$
$$+\frac{1}{n}\sum_{j,l=1;j\neq l}^{n} a_j^i a_l^i\,\mathsf{E}(X_j - m_i)\,\mathsf{E}(Y_j - b_0^i - b_1^i m_i) = \frac{1}{n}\sum_{j=1}^{n}(a_j^i)^2\sum_{k=1}^{M} p_j^k(b_1^k\,\mathsf{D}\,\xi^k + u_{ki}r_{ki}) +$$
$$+\frac{1}{n}\sum_{j,l=1;j\neq l}^{n} a_j^i a_l^i\sum_{k=1}^{M} p_j^k u_{ki}\sum_{m=1}^{M} p_j^m b_1^m r_{mi} = \sum_{k=1}^{M}\langle(a^i)^2, p^k\rangle_n(b_1^k\,\mathsf{D}\,\xi^k + u_{ki}r_{ki}) +$$

$$+\frac{1}{n}\sum_{k=1}^{M}\sum_{j=1}^{n} a_j^i p_j^k u_{ki} \sum_{m=1}^{M}\sum_{l=1}^{n} a_l^i p_l^m b_1^m r_{mi} - \frac{1}{n}\sum_{j=1}^{n}(a_j^i)^2 \sum_{k=1}^{M} p_j^k u_{ki} \sum_{m=1}^{M} p_j^m b_1^m r_{mi}.$$

Тут другий доданок дорівнює нулю внаслідок (6) і оскільки $u_{ii} = 0$.

Тому

$$\widetilde{C}_{12} = \sum_{k=1}^{M}\langle (a^i)^2, p^k\rangle_n (b_1^k R_{k2} + u_{ki} r_{ki}) - \sum_{k,m=1}^{M}\langle (a^i)^2, p^k p^m\rangle_n b_1^m u_{ki} r_{mi}. \tag{36}$$

Міркуючи аналогічно, виписуємо решту елементів коваріаційної матриці.

$$\widetilde{C}_{13} = \mathrm{Cov}\left(\frac{S_1}{\sqrt{n}}, \frac{S_3}{\sqrt{n}}\right) = \sum_{k=1}^{M}\langle (a^i)^2, p^k\rangle_n \big(q_{ki} R_{k3} + (((b_1^i)^2 - 1)(2b_1^k u_{ki} + r_{ki}) +$$

$$+ b_1^i(-2b_1^k r_{ki} - (b_1^k)^2 u_{ki} + 3u_{ki}))R_{k2} + u_{ki} s_{ki} + \sigma_k^2(((b_1^i)^2 - 1) r_{ki} + 2b_1^i u_{ki}) - b_1^i\,\mathsf{E}(\delta^k)^3\big)$$

$$- \sum_{k,m=1}^{M}\langle (a^i)^2, p^k p^m\rangle_n u_{ki} c_{mi}. \tag{37}$$

$$\widetilde{C}_{23} = \mathrm{Cov}\left(\frac{S_2}{\sqrt{n}}, \frac{S_3}{\sqrt{n}}\right) =$$

$$= \sum_{k=1}^{M}\langle (a^i)^2, p^k\rangle_n \big(b_1^k q_{ki} R_{k3} + R_{k2}(((b_1^i)^2 - 1)(2b_1^k r_{ki} + u_{ki}(b_1^k)^2) +$$

$$+ b_1^i(2b_1^k u_{ki} - 3(b_1^k)^2 r_{ki} + r_{ki})) + r_{ki} s_{ki} + (((b_1^i)^2 - 1)u_{ki} - 2b_1^i r_{ki})\sigma_k^2 + b_1^i\,\mathsf{E}(\delta^k)^3\big) -$$

$$- \sum_{k,m=1}^{M}\langle (a^i)^2, p^k p^m\rangle_n r_{ki} c_{mi}. \tag{38}$$

Випишемо тепер матрицю V з формулювання теореми 5.2.

$$V = \lim_{n\to\infty}\frac{V_n}{n} = \lim_{n\to\infty}\frac{1}{n}\big(M_n(\overline{\theta})\big)^{-1}\mathrm{Var}(S_n(\theta))\big((M_n(\theta))^{-1}\big)^T.$$

Переходячи до границі при $n \to \infty$ в матриці $\frac{1}{n}\mathrm{Var}(S_n(\theta_i))$, елементи якої задаються формулами (27), (28), (35), (36), (37), (38), маємо

$$S_\infty = \begin{pmatrix} D_1 & C_{12} & C_{13} \\ C_{12} & D_2 & C_{23} \\ C_{13} & C_{23} & D_3 \end{pmatrix}.$$

Використовуючи вигляд матриці M із формулювання теореми 5.2 і вигляд S_∞, отримуємо

$$V = \begin{pmatrix} \mathrm{Disp}_1 & -b_1^i D_1 + C_{12} + \frac{m_i C_{13}}{K_i} & -\frac{C_{13}}{K_i} \\ -b_1^i D_1 + C_{12} + \frac{m_i C_{13}}{K_i} & \mathrm{Disp}_2 & \frac{b_1^i C_{13} - C_{23}}{K_i} - \frac{m_i D_3}{K_i^2} \\ -\frac{C_{13}}{K_i} & \frac{b_1^i C_{13} - C_{23}}{K_i} - \frac{m_i D_3}{K_i^2} & \mathrm{Disp}_3 \end{pmatrix};$$

$$\mathrm{Disp}_1 = D_1;$$

$$\mathrm{Disp}_2 = (b_1^i)^2 D_1 + D_2 - 2b_1^i C_{12} + \frac{m_i^2 D_3 + 2m_i(C_{23} - b_1^i C_{13})K_i}{K_i^2};$$

$$\mathrm{Disp}_3 = \frac{D_3}{K_i^2}. \tag{39}$$

Як ми бачимо з побудови матриці розсіювання, границя $S_\infty = \lim_{n\to\infty}\frac{1}{n}\mathrm{Var}\, S_n(\theta)$ буде існувати при існуванні границь

$$\lim_{n\to\infty}\langle (a^i)^2, p^k\rangle_n < \infty; \qquad \lim_{n\to\infty}\langle (a^i)^2, (p^k)^2\rangle_n < \infty;$$

$$\lim_{n\to\infty}\langle (a^i)^2, p^k p^l\rangle_n < \infty \quad \forall k,l=\overline{1,M}.$$

Існування третьої із цих границь забезпечується умовою 3) теореми 5.2, а існування першої і другої випливає із третьої – наприклад, для першої це видно з тотожності: $\sum_{l=1}^{M}\langle (a^i)^2, p^k p^l\rangle_n = \langle (a^i)^2, p^k\rangle_n$.

Бачимо, що при виконанні умов 3, 4 теореми 5.2 виконується умова 5 теореми 5.1. Отже, усі умови теореми 5.1 виконані для нашої системи. Застосовуючи цю теорему, отримуємо твердження теореми 5.2. □

Зауваження 5.1. Отже, при виконанні умов теореми 5.2 оцінки (12) задовольняють

$$\sqrt{n}(\hat{b}_1^i - b_1^i) \xrightarrow{\mathrm{W}} N(0,\sigma^2_{b_1^i}); \qquad \sqrt{n}(\hat{b}_0^i - b_0^i) \xrightarrow{\mathrm{W}} N(0,\sigma^2_{b_0^i}); \quad n\to\infty,$$

де

$$\sigma^2_{b_1^i} = \frac{D_3}{K_i^2};$$

$$\sigma^2_{b_0^i} = ((b_1^i)^2 D_1 + D_2 + \frac{m_i^2}{K_i^2} D_3 - 2b_1^i C_{12} - 2b_1^i \frac{m_i}{K_i} C_{13} + 2\frac{m_i}{K_i} C_{23}),$$

де коефіцієнти $K_i, D_1, D_2, D_3, C_{12}, C_{13}, C_{23}$ задаються формулою (21).

6. Результати імітаційного моделювання

Моделювання проводилось на вибірках обсягу n від 100 до 10000, що являли собою суміш двох компонентів; концентрації задавались таким чином:

$$p_j^1 = \frac{j}{n}; \quad p_j^2 = 1 - \frac{j}{n}; \quad j = \overline{1,n}.$$

Оскільки обчислення коефіцієнта розсіювання для оцінки коефіцієнта b_0 вимагає на кілька порядків більше обчислень, ніж для b_1, ми обмежились розглядом лише коефіцієнта розсіювання для b_1, для b_0 рахували лише оцінки для обох компонентів.

У першому експерименті компоненти і похибки були нормально розподіленими з параметрами:

$$\eta_1 \sim N(0,2); \quad \eta_2 \sim N(1,2);$$

$$\delta^1 \sim \delta^2 \sim \varepsilon^1 \sim \varepsilon^2 \sim N(0, 0.25);$$

$$b_0^1 = 0.5; \quad b_1^1 = 2;$$

$$b_0^2 = -0.5; \quad b_1^2 = -\frac{1}{3}.$$

Безпосередня перевірка показує, що для цієї та наступних моделей умови теореми 5.2 виконані, зокрема $\det S_\infty \neq 0$.

Таблиця 1. Результати експерименту 1

n	Bias $\hat{b}_1^1$	$nD\hat{b}_1^1$	Bias $\hat{b}_1^2$	$nD\hat{b}_1^2$	Bias $\hat{b}_0^1$	Bias $\hat{b}_0^2$
100	0,155	3,576	3,208	77490,9	−0,168	−4,378
200	0,111	2,903	−0,487	8357,42	−0,153	0,874
500	0,046	2,7	−0,084	131,72	−0,032	0,253
1000	0,043	2,575	−0,095	7,926	0,0162	0,195
2500	0,002	2,6	−0,066	6,559	−0,055	0,116
5000	0,017	2,559	−0,046	6,184	−0,077	0,073
10000	0,013	2,569	−0,035	6,22	0,046	0,085
∞		2,674		6159		

Таблиця 2. Результати експерименту 1

n	Median $\hat{b}_1^1$	IQ($\hat{b}_1^1$)	Median $\hat{b}_1^2$	IQ($\hat{b}_1^2$)
100	$-0{,}007$	1,648	0,073	2,498
200	0,025	1,652	$-0{,}022$	2,522
500	0,004	1,593	$-0{,}034$	2,561
1000	0,002	1,595	$-0{,}035$	2,442
2500	$-0{,}025$	1,606	$-0{,}048$	2,476
5000	$-0{,}011$	1,6	$-0{,}012$	2,484
10000	0,037	1,608	$-0{,}001$	2,492
∞		1,635		2,482

Таблиця 3. Результати експерименту 2

n	Bias $\hat{b}_1^1$	$nD\hat{b}_1^1$	Bias $\hat{b}_1^2$	$nD\hat{b}_1^2$	Bias $\hat{b}_0^1$	Bias $\hat{b}_0^2$
100	1,356	72,56	3,378	396799	$-0{,}252$	$-3{,}459$
200	0,611	27,24	0,545	7985	$-0{,}07$	$-0{,}17$
500	0,496	24,34	$-0{,}217$	5039	$-0{,}029$	0,108
1000	0,308	22,72	$-0{,}321$	76,51	$-0{,}032$	0,356
2500	0,084	21,99	$-0{,}081$	20,96	$-0{,}005$	0,072
5000	0,096	21,83	$-0{,}098$	19,45	$-0{,}023$	0,135
10000	0,136	21,15	$-0{,}075$	18,81	0,105	0,133
∞		19,85		23,002		

Результати наведено у табл. 1, 2. У табл. 1 Bias $\hat{b}_i^k$ позначає підраховане за набором оцінок зміщення відповідної оцінки, помножене на $\sqrt{n}$, $nD\hat{b}_i^k$ — вибіркову дисперсію оцінки, помножену на n.

В останньому рядочку таблиць (позначеному ∞) вміщено відповідні значення коефіцієнта розсіювання (в табл. 1) і кореня з нього (в табл. 2).

Як ми можемо спостерігати, результати експерименту досить непогано узгоджуються з теорією.

Слід враховувати, що при невеликих обсягах модельованих вибірок, у наборах побудованих за ними оцінок можуть зустрічатись викиди. Тому для дослідження точності асимптотичних формул доцільно використовувати робастні характеристики, а саме – медіану для характеризації зміщення та інтерквартильний розмах для характеризації розкиду. У табл. 2 у стовпчику Median $\hat{b}_1^1$ указано $\sqrt{n}\times\times(\mathrm{Median}(\hat{b}_1^1)-b_1^1))$, а у стовпчику IQ($\hat{b}_1^1$) указано інтерквартильний розмах $\hat{b}_1^1$, ділений на інтерквартильний розмах стандартного нормального розподілу і помножений на $\sqrt{n}$. (Асимптотично він має прямувати до кореня квадратного з коефіцієнта розсіювання)

У другому експерименті ми збільшили дисперсії похибок:

$$\sigma_1^2 = \sigma_2^2 = 2.$$

Результати вміщено у табл. 3, 4.

Як бачимо, коефіцієнти розсіювання і відповідні дисперсії оцінок зросли, але точність наближення асимптотичними значеннями залишилась приблизно такою ж, як і у попередньому експерименті.

У третьому експерименті досліджувався випадок важких хвостів похибок: δ^1, δ^2, $\varepsilon^1, \varepsilon^2$ розподілені за законом Стьюдента з 14 ступенями вільності. (Це найменша кількість ступенів вільності для розподілу похибок, при якій виконується теорема 5.2)

Таблиця 4. Результати експерименту 2

n	Median $\hat{b}_1^1$	IQ($\hat{b}_1^1$)	Median $\hat{b}_1^2$	IQ($\hat{b}_1^2$)
100	−0,021	4,942	0,214	4,341
200	−0,023	4,713	0,068	4,4
500	0,121	4,765	0,054	4,353
1000	0,086	4,662	0,021	4,427
2500	−0,081	4,609	−0,019	4,361
5000	−0,069	4,613	−0,058	4,384
10000	0,061	4,577	−0,018	4,262
∞		4,455		4,796

Таблиця 5. Результати експерименту 3

n	Bias $\hat{b}_1^1$	$nD\hat{b}_1^1$	Bias $\hat{b}_1^2$	$nD\hat{b}_1^2$	Bias $\hat{b}_0^1$	Bias $\hat{b}_0^2$
100	0,744	29,26	−0,949	29144	−0,114	1,203
200	0,38	14,06	−0,225	16549	−0,108	0,356
500	0,212	13,09	0,036	25051	-0,122	0,182
1000	0,197	12,7	−0,158	21,94	0,	0,287
2500	0,063	12,25	−0,14	13,42	0,016	0,177
5000	0,1	12,27	−0,039	12,37	0,023	−0,033
10000	0,012	12,22	−0,014	12,29	−0,003	−0,108
∞		12,09		14,88		

Таблиця 6. Результати експерименту 3

n	Median $\hat{b}_1^1$	IQ($\hat{b}_1^1$)	Median $\hat{b}_1^2$	IQ($\hat{b}_1^2$)
100	−0,002	3,714	0,133	3,571
200	0,006	3,623	0,022	3,66
500	0,002	3,543	−0,089	3,552
1000	0,035	3,495	0,044	3,517
2500	−0,019	3,453	−0,074	3,452
5000	−0,006	3,504	−0,004	3,451
10000	−0,087	3,481	0,026	3,481
∞		3,477		3,858

Результати наведено у табл. 5, 6.

Ці результати свідчать про наявність викидів, що, можливо, відповідають важким хвостам розподілів оцінок при малих та помірних обсягах вибірки. Але точність наближення асимптотичними формулами при $n > 2500$ є достатньою для практичних потреб.

7. Висновки

Таким чином, запрононовані оцінки для коефіцієнтів регресії за спостереженнями із суміші у моделі з похибками у змінних можна трактувати як оцінки навантаженого методу моментів або узагальнені оцінки навантаженого методу повних квадратів. Вони є консистентними та асимптотичнио нормальними за природних умов. Результати моделювання показують, що асимптотичні формули для розкиду оцінок дають хороше наближення при помірних обсягах вибірок.

Подяка. Автори вдячні рецензенту за увагу до роботи і корисні зауваження.

Список літератури

1. C.-L. Cheng, J. Van Ness, *Statistical Regression with Measurement Error*, Kendall's Library of Statistics 6, Arnold, London, 1999.
2. Р. Майборода, О. Сугакова, *Оцінювання та класифікація за спостереженнями із суміші*, ВПЦ "Київський університет", Київ, 2008.
3. S. Masiuk, A. Kukush, S. Shklyar, M. Chepurny, I. Likhtarov (ed.), *Radiation Risk Estimation: Based on Measurement Error Models*, 2nd ed. (de Gruyter series in Mathematics and Life Sciences, vol. 5). de Gruyter, 2017.
4. T. Benaglia, D. Chauveau, D. Hunter, D. Young, *mixtools: An R Package for Analyzing Finite Mixture Models*, Journal of Statistical Software, **32** (2009), no 6, 1–29 .
5. S. Faria, G. Soromenhob, *Fitting mixtures of linear regressions*, Journal of Statistical Computation and Simulation **80**, (2010), no 2, 201–225.
6. B. Grün, F. Leisch, *Fitting finite mixtures of linear regression models with varying & fixed effects in R*, Proceedings in Computational Statistics, Physica Verlag, Heidelberg, Germany, (2006), 853–860.
7. R. Maiboroda, O. Sugakova, *Statistics of mixtures with varying concentrations with application to DNA microarray data analysis*, Journal of Nonparametric Statistics, **24** (2012), №1, 201–205.
8. D. Liubashenko, R. Maiboroda, *Linear regression by observations from mixtures with varying concentrations*, Modern Stochastics: Theory and Applications, **2** (2015), 343–353.
9. J. Shao, *Mathematical statistics*, Springer-Verlag, New York, 1998.
10. R. Branham, *Total Least Squares in Astronomy*, Total Least Squares and Errors-in-Variables Modeling. Springer, Dordrecht, (2002), 375–384.

Кафедра теорії ймовірностей та математичної статистики, механіко-математичний факультет, Національний університет імені Тараса Шевченка, проспект Глушкова, 6, Київ, 03127, Україна

Адреса електронної пошти: mre@univ.kiev.ua

Кафедра теорії ймовірностей та математичної статистики, механіко-математичний факультет, Національний університет імені Тараса Шевченка, проспект Глушкова, 6, Київ, 03127, Україна

Адреса електронної пошти: mrswade111017@gmail.com

Кафедра математики і теоретичної радіофізики, факультет радіофізики, електроніки і комп'ютерних систем, Національний університет імені Тараса Шевченка, проспект Глушкова, 4г, Київ, 03187, Україна

Адреса електронної пошти: sugak@univ.kiev.ua

Стаття надійшла до редколегії 4.03.2018

ORTHOGONAL REGRESSION FOR OBSERVATIONS FROM MIXTURES

R. MAIBORODA, H. NAVARA, O. SUGAKOVA

Abstract. A generalization of orthogonal regression estimators is considered for estimation of simple regression coefficients in the error-in-variables model with observations from a mixture with varying concentrations. Consistency and asymptotic normality of the estimators are shown. The dispersion matrix is evaluated.

МЕТОД ОРТОГОНАЛЬНОЙ РЕГРЕССИИ ДЛЯ НАБЛЮДЕНИЙ ИЗ СМЕСИ

Р. Е. МАЙБОРОДА, А. В. НАВАРА, Е. В. СУГАКОВА

Аннотация. Рассматривается обобщение метода ортогональной регрессии для оценки параметров в модели простой линейной регрессии с ошибками в переменных по наблюдениям из смеси с переменными концентрациями. Доказаны состоятельность и асимптотическая нормальность полученных оценок, найдена матрица рассеивания.