Теорія ймовір. та матем. статист.
Вип. 88, 2013, стор. 157–170

Teor. Ĭmovir. ta Matem. Statyst.
No. 88, 2013, pp. 157–170

АСИМПТОТИЧНА НОРМАЛЬНІСТЬ ПОКРАЩЕНИХ ОЦІНОК НАЙМЕНШИХ КВАДРАТІВ ДЛЯ МНОЖИННОЇ ВЕКТОРНОЇ РЕГРЕСІЇ З ПОХИБКАМИ В ЗМІННИХ

УДК 519.21

І. О. СЕНЬКО

Анотація. У роботі розглядається покращена оцінка найменших квадратів для лінійної множинної векторної моделі з похибками у змінних. Наведено умови асимптотичної нормальності цієї оцінки. Побудовано модифікацію оцінки, яка асимптотично еквівалентна покращеній оцінці найменших квадратів, але більш стійка вже при невеликих обсягах вибірки.

Abstract. A paper deals with adjusted least squares estimator in linear multivariate vector error-in-variables model. Conditions of asymptotic normality for this estimator are stated. A modification of this estimator, which has same asymptotic properties and is more stable starting from small sizes of sample, is built.

Аннотация. В работе рассматривается улучшенная оценка наименьших квадратов для линейной множественной векторной модели с ошибками в переменных. Представлены условия асимптотической нормальности этой оценки. Построена асимптотически эквивалентная модификация оценки, которая является более устойчивой уже при небольших объемах выборки.

1. Вступ

Моделі з похибками в змінних є важливим узагальненням регресійних моделей, оскільки існує велика кількість практичних застосувань, де усі дані, присутні в моделі, отримуються в результаті вимірювань, а отже є неточними. Зокрема, такі моделі зустрічаються в задачах ідентифікації динамічних систем, обробки сигналів, поширення звуку, а також задачах геологічних та океанографічних досліджень. Загальний огляд моделей з похибками в змінних можна знайти в [3]. Для точного оцінювання у подібних моделях необхідно враховувати наявність похибок як у залежних, так і в незалежних змінних.

У даній роботі розглядається лінійна модель з похибками в змінних, що описується переозначеною системою лінійних рівнянь, яка у матричному виді виглядає як $AX = B$, де матриці A та B спостерігаються з похибками, X — невідома матриця, яку потрібно оцінити. Сукупна коваріаційна матриця похибок у спостереженнях матриці A вважається відомою, а у матриці B — невідомою. Розглядається покращена оцінка найменших квадратів матричного параметра X. Умови консистентності та строгої консистентності цієї оцінки для даної моделі встановлено у [7]. Також там можна знайти посилання на роботи, де вивчалися інші можливі оцінки для лінійних моделей з похибками у змінних.

У даній роботі встановлено умови асимптотичної нормальності цієї оцінки, коли кількість рядків у матриці A нескінченно зростає, та побудовано її модифікацію, яка є більш стійкою з обчислювальної точки зору і має такі ж асимптотичні властивості.

2000 *Mathematics Subject Classification.* Primary 62J12.
Ключові слова і фрази. Моделі з похибками в змінних, покращена оцінка найменших квадратів, асимптотична нормальність, малі вибірки.

Ідея побудови такої модифікації запозичена з [2], де така модифікація побудована для скалярної поліноміальної моделі з похибками вимірювання.

Робота побудована наступним чином. У розділі 2 описано модель спостережень та наведено попередні результати. У розділі 3 доведено асимптотичну нормальність, побудовано консистентну оцінки асимптотичної коваріаційної матриці та будується довірчий еліпсоїд для матриці X. У розділі 4 побудовано модифікацію для малої вибірки. Розділ 5 містить висновки. Технічні доведення містяться у Додатку.

Надалі будуть використовуватися наступні позначення. Усі вектори є стовпцями; $\|z\|$ — евклідова норма вектора $z \in \mathbb{R}^n$; I_n — одинична матриця розміру $n \times n$; для матриці $Z \in \mathbb{R}^{m\times n}$, $Z = (z_{ij})_{i=1,\ j=1}^{m\quad n}$ будемо позначати $\|Z\|$ — операторна норма матриці, що відповідає евклідовій нормі у $\mathbb{R}^n$ та $\mathbb{R}^m$, а $\|Z\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^m \sum_{j=1}^n z_{ij}^2}$ — норма Фробеніуса матриці Z. Через $Z^\dagger$ будемо позначати псевдообернену матрицю [6] до матриці Z. Математичне сподівання та дисперсія будуть позначатися символами E та Var, відповідно. Для послідовності випадкових матриць $\left\{X_m \in \mathbb{R}^{n\times d}, m \geqslant 0\right\}$ запис $X_m \xrightarrow{\mathsf{P}} X_0$, $m \to \infty$, означає, що $\|X_m - X_0\| \xrightarrow{\mathsf{P}} 0$, $m \to \infty$. Аналогічно визначається збіжність випадкових матриць майже напевно $X_m \xrightarrow{\mathsf{P1}} X_0$, $m \to \infty$. Позначення $X_m = o_{\mathsf{P}}(1)$, $m \to \infty$, означає, що $X_m \xrightarrow{\mathsf{P}} 0$, $m \to \infty$, а $X_m = O_{\mathsf{P}}(1)$, $m \to \infty$, означає, що послідовність випадкових величин $\{\|X_m\|\}$ стохастично обмежена. Для квадратної симетричної матриці V через $\lambda_{\min}(V)$ та $\lambda_{\max}(V)$ позначаються відповідно її найменше та найбільше власні числа. Для квадратних матриць U та V однакової розмірності запис $U > V$ означає, що матриця $U - V$ є додатно визначеною, а запис $U \geqslant V$ означає, що матриця $U - V$ є невід'ємно визначеною. Для прямокутної матриці $W = \begin{bmatrix} w_1 & w_2 & \dots & w_d \end{bmatrix} \in \mathbb{R}^{q\times d}$ будемо позначати $\operatorname{vec}(W) = \begin{bmatrix} w_1^T & w_2^T & \dots & w_d^T \end{bmatrix}^T \in \mathbb{R}^{qd}$. Будемо позначати $Z \otimes W = \begin{bmatrix} z_{11}W & z_{12}W & \dots & z_{1n}W \\ z_{21}W & z_{22}W & \dots & z_{2n}W \\ \dots & & & \\ z_{m1}W & z_{m2}W & \dots & z_{mn}W \end{bmatrix}$ — кронекерівський добуток прямокутних матриць Z та W. Через const будемо позначати довільну сталу, яка не залежить від кількості вимірювань m.

2. Модель спостережень

Нехай є деякий невідомий лінійний оператор $\mathfrak{X}$ з $\mathbb{R}^n$ в $\mathbb{R}^d$, $\mathfrak{X}z = X_0^T z$, $z \in \mathbb{R}^n$. Тут $X_0 = \left(x_{ij}^0\right)_{i=1,\ j=1}^{n\quad d}$. Для знаходження цього оператора спостерігаються набори його вхідних і вихідних значень; спостереження містять адитивні випадкові похибки. Позначимо через $a_i^0 \in \mathbb{R}^n$, $i \geqslant 1$, невипадкові вектори, що задають істинні значення вхідних даних; $b_i^0 \in \mathbb{R}^d$, $i \geqslant 1$, — це вектори, що задають істинні значення вихідних даних; елементи цих векторів позначатимемо відповідно a_{ij}^0 та b_{ij}^0. Для істинних значень виконується рівність

$$b_i^0 = X_0^T a_i^0, \qquad i \geqslant 1.$$

Спостерігаються вектори $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}^d$, $1 \leqslant i \leqslant m$, причому

$$a_i = a_i^0 + \tilde{a}_i, \qquad b_i = b_i^0 + \tilde{b}_i,$$

де $\tilde{a}_i$ та $\tilde{b}_i$ — центровані випадкові похибки вимірювань. Елементи цих векторів будуть позначатися аналогічно через a_{ij}, $\tilde{a}_{ij}$, b_{ij}, $\tilde{b}_{ij}$. Тоді векторна лінійна модель з похибками у змінних задається рівностями

$$b_i = X_0^T a_i^0 + \tilde{b}_i, \qquad a_i = a_i^0 + \tilde{a}_i, \qquad 1 \leqslant i \leqslant m.$$

Або, використовуючи позначення $A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}^T = (a_{ij})_{i=1,\, j=1}^{m,\ n}$ і так само задаючи A_0, $\tilde{A}$, B, B_0, $\tilde{B}$, де $A = A_0 + \tilde{A}$, $B = B_0 + \tilde{B}$, цю модель можна символічно записати у вигляді наближеної рівності

$$AX_0 \approx B$$

за умови, що

$$A_0 X_0 = B_0.$$

Будемо позначати $V_{\tilde{A}} = \mathsf{E}\,\tilde{A}^T\tilde{A}$ — кореляційна матриця сукупних похибок у регресорі. Вона вважається відомою.

Основним припущенням про похибки спостережень є наступне:

(i) : Набори похибок $\{\tilde{a}_{ij}, i \geqslant 1, 1 \leqslant j \leqslant n\}$ та $\{\tilde{b}_{il}, i \geqslant 1, 1 \leqslant l \leqslant d\}$ незалежні між собою, центровані, а також мають скінченні другі моменти.

Тоді за допомогою методу Corrected Score [1, розд. 7], може бути побудована покращена оцінка найменших квадратів матричного параметра X_0

$$\hat{X} = \left(A^T A - V_{\tilde{A}}\right)^{\dagger} A^T B. \tag{1}$$

Позначимо $V_{A_0} = A_0^T A_0$. У [7] встановлені наведені нижче умови, за яких ця оцінка буде консистентною та строго консистентною:

Теорема 2.1 (Консистентність)**.** *Нехай виконуються припущення* (i) *та наступні умови*

(ii) : *Рядки матриці $\tilde{A}$ незалежні, тобто незалежними є вектори $\{\tilde{a}_i, i \geqslant 1\}$, та рядки матриці $\tilde{B}$ незалежні, тобто незалежними є вектори $\{\tilde{b}_i, i \geqslant 1\}$.*

(iii) : $\mathsf{E}\,\tilde{a}_{ij}^4 \leqslant \mathrm{const}$; $\mathsf{E}\,\tilde{b}_{il}^2 \leqslant \mathrm{const}$.

(iv) :

$$\frac{\lambda_{\max}\left(V_{A_0}\right) + m}{\lambda_{\min}^2\left(V_{A_0}\right)} \to 0, \qquad m \to \infty.$$

Тоді $\hat{X} \xrightarrow{\mathsf{P}} X_0$, $m \to \infty$.

Теорема 2.2 (Строга консистентність)**.** *Нехай виконуються припущення* (i), (ii) *та наступні умови*

(v) : $\exists\, r \geqslant 2\colon\ \mathsf{E}\left|\tilde{a}_{ij}\right|^{2r} \leqslant \mathrm{const}$; $\mathsf{E}\left|\tilde{b}_{ij}\right|^{2r} \leqslant \mathrm{const}$.

(vi) : *Для числа r з умови* (v) *та деякого $m_0 \geqslant 1$ виконується:*

$$\sum_{m=m_0}^{\infty} \left(\frac{m^{r/2}}{\lambda_{\min}^r\left(V_{A_0}\right)} + \frac{\lambda_{\max}^r\left(V_{A_0}\right)}{\lambda_{\min}^{2r}\left(V_{A_0}\right)} \right) < \infty.$$

Тоді

$$\hat{X} \xrightarrow{\mathsf{P1}} X_0, \qquad m \to \infty.$$

3. Асимптотична нормальність

3.1. **Теорема про асимптотичну нормальність.** Для доведення асимптотичної нормальності будуть використовуватися припущення (i)–(iii), а також наступні додаткові припущення.

(vii) : $\frac{1}{m} V_{A_0} \to V_{A_\infty}$, $m \to \infty$, де $V_{A_\infty} > 0$.

З цієї умови випливає: $m^{-1}\lambda_{\max}\left(V_{A_0}\right) \to \lambda_{\max}\left(V_{A_0}\right)$, $m \to \infty$,

$$m^{-1}\lambda_{\min}\left(V_{A_0}\right) \to \lambda_{\min}\left(V_{A_0}\right), \qquad m \to \infty.$$

Також з умови (vii) випливає умова (iv).

(viii) : Рядки матриці $\tilde{A}$ є однаково розподіленими та рядки матриці $\tilde{B}$ є однаково розподіленими, тобто однаково розподілені вектори $\tilde{a}_i \sim \tilde{a}$, $i \geqslant 1$, і однаково розподіленими є вектори $\tilde{b}_i \sim \tilde{b}$, $i \geqslant 1$.

Будемо позначати $V_{\tilde{a}} = \operatorname{Cov}(\tilde{a}_i) = \frac{1}{m} V_{\tilde{A}}$ та $V_{\tilde{b}} = \operatorname{Cov}(\tilde{b}_i) = \frac{1}{m} \mathsf{E}\, \tilde{B}^T \tilde{B}$. Для того, щоб відрізняти вектори $\tilde{a}_i$ від елементів вектора $\tilde{a}$, останні позначатимуться як $\tilde{a}(i)$, $i = 1, \ldots, n$. Так само елементи вектора $\tilde{b}$ будуть позначатися $\tilde{b}(i)$, $i = 1, \ldots, d$.

Подальші припущення є такими:

(ix) : $\exists\, \delta > 0\;\; \mathsf{E}\, |\tilde{a}(j)|^{4+\delta} < \infty$, $j = 1, \ldots, n$.

(x) : $\exists\, \tau > 0\;\; \sup_{m \geqslant 1} \frac{1}{m} \sum_{i=1}^{m} \left\| a_i^0 \right\|^{2+\tau} < \infty$.

(xi) : $\exists \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} a_i^0 \overset{\text{def}}{=} a^0$.

Елементи вектора a^0 будемо позначати $a^0(i)$, $i = 1, \ldots, n$.

Останнє у цьому розділі припущення вводиться для того, щоб отримати додатно визначену асимптотичну коваріаційну матрицю.

(xii) : $V_{\tilde{a}} > 0$.

Теорема 3.1 (Асимптотична нормальність)**.** *1) Нехай виконані припущення* (i)–(iii) *та* (vii)–(xi)*. Тоді*

$$\sqrt{m} \cdot \operatorname{vec}\bigl(\hat{X} - X_0\bigr) \xrightarrow{d} N(0, \Sigma), \qquad m \to \infty,$$

де матриця Σ визначена рівняннями (6) *і* (11).

2) Якщо додатково виконане припущення (xii)*, то* $\Sigma > 0$.

Доведення. 1) З припущення (vii) випливає, що існує $m_0 \in \mathbb{N}$ таке, що для будь-якого $m \geqslant m_0$ буде виконуватися $\det V_{A_0} \neq 0$. Тоді для $m \geqslant m_0$ рівняння (1) можна переписати у вигляді:

$$\left(A^T A - V_{\tilde{A}}\right) \hat{X} = A^T \bigl(B_0 + \tilde{B}\bigr),$$

або

$$V_{A_0}^{-1} \left(A^T A - V_{\tilde{A}}\right) \hat{X} = V_{A_0}^{-1} \bigl(A^T A_0 X_0 + A^T \tilde{B}\bigr).$$

Далі відокремимо у лівій частині різницю $\hat{X} - X_0$:

$$V_{A_0}^{-1} \bigl(A^T A - V_{\tilde{A}}\bigr) \bigl(\hat{X} - X_0\bigr) = V_{A_0}^{-1} A^T \tilde{B} + V_{A_0}^{-1} \left(A^T A_0 X_0 - \left(A^T A - V_{\tilde{A}}\right) X_0\right). \tag{2}$$

Якщо виконані припущення (i)–(iii) та (vii) то має місце теорема 2.1 про слабку консистентність, для доведення якої в [7] було встановлено співвідношення:

$$V_{A_0}^{-1} \left(A^T A - V_{\tilde{A}}\right) = I_n + o_{\mathsf{P}}(1). \tag{3}$$

Тоді (2) перетворюється на

$$\left(I_n + o_{\mathsf{P}}(1)\right) \left(\hat{X} - X_0\right) = V_{A_0}^{-1} \bigl(A^T \tilde{B} + V_{\tilde{A}} X_0 - A^T \tilde{A} X_0\bigr) \tag{4}$$

Позначимо

$$l = \operatorname{vec}\bigl(A^T \tilde{B} + V_{\tilde{A}} X_0 - A^T \tilde{A} X_0\bigr) = \operatorname{vec}\left(\sum_{i=1}^{m} a_i \tilde{b}_i^T + \sum_{i=1}^{m} \mathsf{E}\, \tilde{a}_i \tilde{a}_i^T X_0 - \sum_{i=1}^{m} a_i \tilde{a}_i^T X_0 \right).$$

Як показано у додатку,

$$\frac{l}{\sqrt{m}} \xrightarrow{d} N(0, \Sigma_1), \qquad m \to \infty, \tag{5}$$

де Σ_1 визначено у (11).

Тому

$$\sqrt{m}\, \left(I_n + o_{\mathsf{P}}(1)\right) \operatorname{vec}\bigl(\hat{X} - X_0\bigr) = \left(I_d \otimes \left(\frac{V_{A_0}}{m} \right)^{-1} \right) \frac{l}{\sqrt{m}},$$

звідси

$$\sqrt{m}\,\mathrm{vec}\big(\hat{X} - X_0\big) \xrightarrow{d} \big(I_d \otimes V_{A_\infty}^{-1}\big)\, N\,(0, \Sigma_1) = N\,(0, \Sigma)\,,$$

де асимптотична коваріаційна матриця

$$\Sigma = \big(I_d \otimes V_{A_\infty}^{-1}\big)\, \Sigma_1 \,\big(I_d \otimes V_{A_\infty}^{-1}\big)^T. \tag{6}$$

2) Доведення винесено у Додаток. □

Наслідок 3.1.

$$\left\|\sqrt{m}\,\Sigma^{-1/2}\,\mathrm{vec}\big(\hat{X} - X_0\big)\right\|^2 = m\cdot \mathrm{vec}^T\big(\hat{X} - X_0\big)\Sigma^{-1}\,\mathrm{vec}\big(\hat{X} - X_0\big) \xrightarrow{d} \chi^2_{nd}, \quad m \to \infty, \tag{7}$$

3.2. **Побудова довірчого еліпсоїда для** X_0**.** Для того, щоб побудувати довірчий еліпсоїд, потрібно мати консистентну оцінку матриці Σ. Для цього знадобиться наступне припущення.

(xiii) : Відомі треті та четверті моменті вектора похибок у прихованих змінних, тобто відомі набори $\{\mathsf{E}\,\tilde{a}(i)\tilde{a}(j)\tilde{a}(k) \mid i, j, k = 1, \ldots, d\}$ та

$$\{\mathsf{E}\,\tilde{a}(i)\tilde{a}(j)\tilde{a}(k)\tilde{a}(l) \mid i, j, k, l = 1, \ldots, d\}.$$

Ці моменти можна оцінити, наприклад, при наявності повторних вимірювань.

Із припущення (xiii) випливає, що у виразі (11) відомим є доданок $\mathsf{E}\,(\tilde{a}\otimes\tilde{a})\,(\tilde{a}\otimes\tilde{a})$, а множник $a^0(j)$ в елементі $a^0(j)\,\mathsf{E}\,\tilde{a}(i)\tilde{a}(k)\tilde{a}(l)$ матриці $\mathsf{E}\left[\left(\tilde{a}\otimes a^0\right)(\tilde{a}\otimes\tilde{a})\right]$ консистентно оцінюється як

$$\widehat{a^0(j)} = \frac{1}{m}\sum_{p=1}^{m} a_{pj} \xrightarrow{\mathsf{P}} a^0(j), \qquad m \to \infty.$$

Позначимо консистентну оцінку матриці $\mathsf{E}\left[\left(\tilde{a}\otimes a^0\right)(\tilde{a}\otimes\tilde{a})\right]$ через $U_{A,\mathrm{app}}$, тобто

$$U_{A,\mathrm{app}} \xrightarrow[m\to\infty]{\mathsf{P}} \mathsf{E}\left[\left(\tilde{a}\otimes a^0\right)(\tilde{a}\otimes\tilde{a})\right].$$

Побудуємо консистентну оцінку для V_{A_∞}. Використовуючи (3), маємо:

$$\left(\frac{1}{m}A_0^T A_0\right)^{-1}\left(\frac{1}{m}A^T A - V_{\tilde{a}}\right) \xrightarrow{\mathsf{P}} I_n, \qquad m \to \infty,$$

$$\hat{V}_{A_\infty} \stackrel{\mathrm{def}}{=} \frac{1}{m}A^T A - V_{\tilde{a}} \xrightarrow{\mathsf{P}} V_{A_\infty}.$$

Для оцінювання коваріаційної матриці $V_{\tilde{b}}$ застосуємо рівність $B = A_0X_0 + \tilde{B}$. Тоді:

$$\mathsf{E}\,B^T B = X_0^T\left(A_0^T A_0\right)X_0 + \mathsf{E}\,\tilde{B}^T\tilde{B},$$

$$\frac{1}{m}\,\mathsf{E}\,B^T B = X_0^T\left(\frac{1}{m}A_0^T A_0\right)X_0 + V_{\tilde{b}}.$$

Оскільки, $\frac{1}{m}\,\mathsf{E}\,B^T B = \frac{1}{m}\sum_{i=1}^m \mathsf{E}\,b_i b_i^T$, то доведемо, що для послідовності матриць $\{b_i b_i^T, i \geqslant 1\}$ виконується закон великих чисел, тобто

$$\frac{1}{m}\sum_{i=1}^m b_i b_i^T - \frac{1}{m}\sum_{i=1}^m \mathsf{E}\,b_i b_i^T \xrightarrow{\mathsf{P}} 0, \qquad m \to \infty.$$

Маємо

$$\frac{1}{m}\sum_{i=1}^m b_i b_i^T = \frac{1}{m}\sum_{i=1}^m b_i^0 b_i^{0T} + \frac{1}{m}\sum_{i=1}^m b_i^0 \tilde{b}_i^T + \frac{1}{m}\sum_{i=1}^m \tilde{b}_i b_i^{0T} + \frac{1}{m}\sum_{i=1}^m \tilde{b}_i\tilde{b}_i^T.$$

Доведемо, що
$$\frac{1}{m}\sum_{i=1}^{m} b_i^0 \tilde{b}_i^T \to \mathsf{E}\left(\frac{1}{m}\sum_{i=1}^{m} b_i^0 \tilde{b}_i^T\right) = 0, \qquad m\to\infty.$$
Для цього застосуємо закон великих чисел Чебишова до послідовностей незалежних випадкових величин $\{b_{ij}^0 \tilde{b}_{ik}, i \geqslant 1\}$, $j,k = 1,\dots,n$. Потрібно встановити, що
$$\operatorname{Var}\left(b_{ij}^0 \tilde{b}_{ik}\right) = o(i), \qquad i\to\infty,$$
$$\operatorname{Var}\left(b_{ij}^0 \tilde{b}_{ik}\right) = \left(b_{ij}^0\right)^2 \mathsf{E}\left(\tilde{b}_{ik}\right)^2 \leqslant \|b_i^0\|^2 \cdot \mathrm{const} = \mathrm{const}\cdot \|X_0 a_i^0\|^2 \leqslant \mathrm{const}\cdot \|X_0\|^2 \left\|a_i^0\right\|^2.$$
Оскільки, $\frac{1}{m}\sum_{i=1}^{m} a_i^0 a_i^{0T} \to V_\infty$, $m\to\infty$, то
$$\frac{1}{m} a_m^0 a_m^{0T} = \frac{1}{m}\sum_{i=1}^{m} a_i^0 a_i^{0T} - \frac{m-1}{m}\cdot\frac{1}{m-1}\sum_{i=1}^{m-1} a_i^0 a_i^{0T} \xrightarrow[m\to\infty]{} V_{A_\infty} - 1\cdot V_{A_\infty} = 0,$$
$$\|a_i^0 a_i^{0T}\|_{\mathsf{F}} = \left(\sum_{j=1,k=1}^{m}\left(a_{ij}^0 a_{ik}^0\right)^2\right)^{1/2} = \left(\sum_{j=1}^{m}\left(a_{ij}^0\right)^2\sum_{k=1}^{m}\left(a_{ik}^0\right)^2\right)^{1/2} = \|a_i\|^2.$$
Отже, $\|a_i\|^2 = o(i)$, $i\to\infty$, а тому
$$\operatorname{Var}\left(b_{ij}^0 \tilde{b}_{ik}\right) \leqslant \mathrm{const}\cdot\|a_i\|^2 = o(i), \qquad i\to\infty.$$
Потім
$$\frac{1}{m}\sum_{i=1}^{m} b_i^{0T}\tilde{b}_i = \left(\frac{1}{m}\sum_{i=1}^{m} b_i^0\tilde{b}_i^T\right)^T \to 0, \qquad m\to\infty.$$
Далі, враховуючи припущення (ii) та (viii), поелементно застосовуємо посилений закон великих чисел, теорема 3 з [9, гл. IV.3], до послідовності матриць $\left\{\tilde{b}_i\tilde{b}_i^T, i\geqslant 1\right\}$, і отримуємо
$$\frac{1}{m}\sum_{i=1}^{m}\tilde{b}_i\tilde{b}_i^T \xrightarrow{\mathsf{P1}} \operatorname{Cov}\left(\tilde{b}\right) = V_{\tilde{b}}, \qquad m\to\infty.$$
Отже,
$$\frac{1}{m}\sum_{i=1}^{m} b_i b_i^T - \frac{1}{m}\sum_{i=1}^{m}\mathsf{E}\, b_i b_i^T = \frac{1}{m}\sum_{i=1}^{m} b_i^0\tilde{b}_i^T + \frac{1}{m}\sum_{i=1}^{m}\tilde{b}_i b_i^{0T} + \frac{1}{m}\sum_{i=1}^{m}\tilde{b}_i\tilde{b}_i^T - V_{\tilde{b}}$$
$$\xrightarrow[m\to\infty]{\mathsf{P}} 0 + 0 + V_{\tilde{b}} - V_{\tilde{b}} = 0.$$

Тому консистентною оцінкою для $V_{\tilde{b}}$ буде
$$\hat{V}_{\tilde{b}} \overset{\mathrm{def}}{=} \frac{1}{m}\sum_{i=1}^{m} b_i^T b_i - \hat{X}^T\left(\frac{1}{m}A^T A - V_{\tilde{a}}\right)\hat{X} \xrightarrow{\mathsf{P}} V_{\tilde{b}}, \qquad m\to\infty.$$

Зрештою,
$$\begin{aligned}\hat{\Sigma}_1 \overset{\mathrm{def}}{=} {} & \hat{V}_{\tilde{b}}\otimes\hat{V}_{A_\infty} + \hat{V}_{\tilde{b}}\otimes V_{\tilde{a}} \\ & + \left(\hat{X}\otimes I_n\right) \\ & \quad\times\left[V_{\tilde{a}}\otimes\hat{V}_{A_\infty} + \mathsf{E}\left(\tilde{a}\otimes\tilde{a}\right)\left(\tilde{a}\otimes\tilde{a}\right) - \operatorname{vec}\left(V_{\tilde{a}}\right)\operatorname{vec}^T\left(V_{\tilde{a}}\right) + U_{A,\mathrm{app}} + U_{A,\mathrm{app}}^T\right] \\ & \quad\times\left(\hat{X}\otimes I_n\right)^T \xrightarrow{\mathsf{P}} \Sigma_1, \qquad m\to\infty,\end{aligned}$$
та
$$\hat{\Sigma} \overset{\mathrm{def}}{=} \left(I_d\otimes\hat{V}_{A_\infty}^{-1}\right)\hat{\Sigma}_1\left(I_d\otimes\hat{V}_{A_\infty}^{-1}\right)^T \xrightarrow{\mathsf{P}} \Sigma, \qquad m\to\infty.$$

Тоді, враховуючи (7), довірчий еліпсоїд з довірчою ймовірністю γ для $\mathrm{vec}\,(X_0)$ у просторі $\mathbb{R}^{nd}$ координат $(x_{11}, x_{21}, \ldots, x_{n1}, x_{12}, x_{22}, \ldots, x_{n2}, \ldots, x_{1d}, x_{2d}, \ldots, x_{nd})$ задається нерівністю

$$\left(\mathrm{vec}(X) - \mathrm{vec}\big(\hat{X}\big)\right)^T \left(\frac{1}{m}\hat{\Sigma}\right)^{-1} \left(\mathrm{vec}(X) - \mathrm{vec}\big(\hat{X}\big)\right) \leqslant \chi^2_{nd,\gamma},$$

де $X = (x_{ij})_{i=1,\, j=1}^{\;n \qquad d}$, а $\chi^2_{nd,\gamma}$ — квантиль рівня γ розподілу хі-квадрат з nd ступенями вільності, тобто $\mathsf{P}\left(\chi^2_{nd} \leqslant \chi^2_{nd,\gamma}\right) = \gamma$.

4. Модифікація для малої вибірки

При практичному застосуванні розглянутої оцінки виникає проблема пов'язана з тоим, що матриця $\left(A^T A - V_{\tilde{A}}\right)^{-1}$ існує з імовірністю, що прямує до 1 за умов консистентності оцінки. Якщо розглядати умови строгої консистентності, то вона буде існувати після деякої випадкової кількості спостережень і цих спостережень може бути потрібно накопичити як завгодно випадково багато. Тому бажано модифікувати цю оцінку так, щоб асимптотичні властивості зберігалися і обернена матриця існувала при невеликій кількості вимірювань з імовірністю 1. Це дозволить отримати більшу точність оцінювання, внаслідок зведення до нуля ймовірності випадкових відхилень при обчисленні $\left(A^T A - V_{\tilde{A}}\right)^{\dagger}$.

Модифікація будується наступним чином. Позначимо

$$T = \begin{bmatrix} B^T B & B^T A \\ A^T B & A^T A \end{bmatrix} = \begin{bmatrix} B & A \end{bmatrix}^T \cdot \begin{bmatrix} B & A \end{bmatrix}, \qquad W = \begin{bmatrix} 0 & 0 \\ 0 & V_{\tilde{A}} \end{bmatrix}$$

— матриці розмірності $(n+d) \times (n+d)$.

Нехай λ_A — найменший додатний корінь рівняння $\det(T - \lambda W) = 0$, якщо він існує, тоді означимо

$$\mu = \begin{cases} \frac{m-\alpha}{m}, & \lambda_A > 1 + \frac{1}{m}, \\ \lambda_A \cdot \frac{m-\alpha}{m+1}, & \lambda_A \leqslant 1 + \frac{1}{m}, \end{cases}$$

де α можна вибрати з проміжку $(0, m)$. Відповідно до рекомендацій з [2], можна взяти $\alpha = nd + 1$, для кількості вимірювань $m > nd + 1$. Модифікована оцінка визначається наступним чином:

$$\hat{X}_{\mathrm{MALS}} \stackrel{\mathrm{def}}{=} \left(A^T A - \mu V_{\tilde{A}}\right)^{\dagger} A^T B.$$

Введемо наступне припущення.

(xiv) : Розподіли $L(\tilde{a}(j))$, $1 \leqslant j \leqslant n$, $L(\tilde{b}(l))$, $1 \leqslant l \leqslant d$, не мають атомів, тобто $\forall j = 1, \ldots, n\ \forall z \in \mathbb{R}\colon\ \mathsf{P}(\tilde{a}(j) = z) = 0$ і аналогічно з $\tilde{b}(l), 1 \leqslant l \leqslant d$.

Тоді має місце наступне твердження.

Лема 4.1. *Нехай виконані припущення* (i), (ii) *та* (xiv). *Тоді для кожного $m \geqslant n+d$ буде виконуватися: $T > 0$ м.н.*

Доведення. Доведемо, що

$$\forall m \geqslant n + d\colon\ \mathrm{rank}[B \quad A] = n + d \quad \text{м.н.} \tag{8}$$

З (8) випливає, що $\forall x \in \mathbb{R}^{n+d}$

$$x^T T x = x^T [B \quad A]^T [B \quad A] x = \big\|[B \quad A]x\big\|^2 \geqslant 0,$$

отже $x^T T x = 0 \Leftrightarrow [B \quad A]x = 0$. Маємо $[B \quad A] \in \mathbb{R}^{m\times(n+d)}$, тому із того, що $\mathrm{rank}[B \quad A] = n+d$ м.н., випливає $\mathrm{Ker}[B \quad A] = \{0\}$ м.н., тобто $[B \quad A]x = 0 \Leftrightarrow x = 0$. Отже, $\forall x \in \mathbb{R}^{n+d} \setminus \{0\}\colon x^T T x > 0$ м.н., тобто $T > 0$ м.н.

Для доведення (8) помітимо, що з припущень (i) та (ii) випливає незалежність векторів $\left\{\left(b_i^T \quad a_i^T\right)^T, 1 \leqslant i \leqslant n+d\right\}$ у сукупності як випадкових векторів.Тому, використовуючи умову (xiv) маємо, що для будь-якого набору чисел $\alpha_1, \ldots, \alpha_{n+d}$ з $\mathbb{R}$:

$$\mathsf{P}\left(\alpha_1 \left(b_1^T \quad a_1^T\right)^T + \cdots + \alpha_{n+d}\left(b_{n+d}^T \quad a_{n+d}^T\right)^T = 0\right) = 0,$$

отже $\left\{\left(b_i^T \quad a_i^T\right)^T, 1 \leqslant i \leqslant n+d\right\}$ — лінійно незалежна система м.н. і тоді для всіх $m \geqslant n+d$

$$\operatorname{rank}\begin{bmatrix} b_1^T & a_1^T \\ b_2^T & a_2^T \\ \ldots & \ldots \\ b_{n+d}^T & a_{n+d}^T \end{bmatrix} = n+d = \operatorname{rank}\begin{bmatrix} b_1^T & a_1^T \\ b_2^T & a_2^T \\ \ldots & \ldots \\ b_m^T & a_m^T \end{bmatrix}. \quad \square$$

Лема 4.2. *Нехай виконані припущення* (i), (ii) *та* (xiv). *Тоді для будь-якого $m \geqslant n+d$:*

1) *μ та, відповідно, X_{MALS} коректно визначені майже напевне;*
2) *$A^TA - \mu V_{\tilde{A}} \geqslant \frac{\alpha+1}{m+1} A^TA > 0$ м.н.*

Доведення. Усі рівності та нерівності у доведенні розглядаються майже напевне.

1) Маємо $T - \lambda W = T^{-1/2}\left(I_{n+d} - \lambda T^{-1/2} W T^{-1/2}\right) T^{-1/2}$. Тому

$$\det(T - \lambda W) = 0 \Leftrightarrow \det\left(I_{n+d} - \lambda T^{-1/2} W T^{-1/2}\right) = 0$$
$$\Leftrightarrow \det\left(T^{-1/2} W T^{-1/2} - \frac{1}{\lambda} I_{n+d}\right) = 0.$$

Оскільки, $T^{-1/2} W T^{-1/2} \geqslant 0$, то $\lambda_A = \left(\lambda_{\max}\left(T^{-1/2} W T^{-1/2}\right)\right)^{-1}$ існує м.н. Тому коректно визначені μ та X_{MALS}.

2) Далі, $\mu \leqslant \lambda_A \frac{m-\alpha}{m+1}$, отже,

$$A^TA - \mu V_{\tilde{A}} \geqslant A^TA - \lambda_A \frac{m-\alpha}{m+1} V_{\tilde{A}} = \frac{m-\alpha}{m+1}\left(A^TA - \lambda_A V_{\tilde{A}}\right) + \frac{\alpha+1}{m+1} A^TA. \qquad (9)$$

Розглянемо

$$T - \lambda_A W = T^{-1/2}\left(I_{n+d} - \lambda_A T^{-1/2} W T^{-1/2}\right) T^{-1/2}.$$

Оскільки $\lambda_A = \lambda_{\max}^{-1}\left(T^{-1/2} W T^{-1/2}\right)$, то $\lambda_A T^{-1/2} W T^{-1/2} \leqslant I_{n+d}$. Тому $T - \lambda_A W \geqslant 0$.

Далі, $T - \lambda_A W = \begin{bmatrix} B^TB & B^TA \\ A^TB & A^TA - \lambda_A V_{\tilde{A}} \end{bmatrix} \geqslant 0 \Rightarrow A^TA - \lambda_A V_{\tilde{A}} \geqslant 0$.

Оскільки $T = \begin{bmatrix} B^TB & B^TA \\ A^TB & A^TA \end{bmatrix} > 0$, то $A^TA > 0$.

Використовуючи розклад (9), маємо $A^TA - \lambda_A V_{\tilde{A}} \geqslant \frac{\alpha+1}{m+1} A^TA > 0$. $\square$

Отже, для будь-якого $m \geqslant n+d$ з ймовірністю 1 існує $\hat{X}_{\mathrm{MALS}} = \left(A^TA - \mu V_{\tilde{A}}\right)^{-1} \times A^TB$, тобто оцінка існує вже при кількості вимірювань, яка співставна з розмірностями матриць, до того ж обчислюється за допомогою обертання, а не псевдообертання деякої додатновизначеної матриці, що дозволяє очікувати суттєво кращих обчислювальних властивостей модифікованої оцінки.

Для доведення того, що асимптотичні властивості модифікованої оцінки зберігаються, буде необхідне ще одне припущення.

(xv) : $V_{\tilde{b}} = \operatorname{Cov}(\tilde{b}) > 0$.

Лема 4.3. *Нехай виконані умови* (i)–(iii), (vii), (viii), (xiv) *та* (xv). *Тоді*

$$\sqrt{m}(1-\mu) \xrightarrow{\mathsf{P}} 0, \qquad m \to \infty.$$

Доведення леми уміщено в додатку.

За допомогою цього факту доводиться теорема про збереження асимптотичних властивостей модифікованої оцінки.

Теорема 4.1. *Нехай виконуються припущення* (i)–(iii), (vii), (viii), (xiv), *та* (xv). *Тоді*

$$\sqrt{m}\left(\hat{X}_{\mathrm{MALS}} - \hat{X}_{\mathrm{ALS}}\right) \xrightarrow{\mathsf{P}} 0, m \to \infty.$$

Доведення. З

$$\left(A^T A - \mu V_{\tilde{A}}\right) \hat{X}_{\mathrm{MALS}} = A^T B$$

і

$$\left(A^T A - V_{\tilde{A}}\right) \hat{X}_{\mathrm{ALS}} = A^T B$$

маємо

$$\left(A^T A - \mu V_{\tilde{A}}\right) \hat{X}_{\mathrm{MALS}} - \left(A^T A - V_{\tilde{A}}\right) \hat{X}_{\mathrm{ALS}} = 0.$$

Таким чином,

$$\left(A^T A - \mu V_{\tilde{A}}\right) \left(\hat{X}_{\mathrm{MALS}} - \hat{X}_{\mathrm{ALS}}\right) = -(1-\mu) V_{\tilde{A}} \hat{X}_{\mathrm{ALS}}.$$

Далі,

$$V_{A_0}^{-1}\left(A^T A - \mu V_{\tilde{A}}\right) \sqrt{m}\left(\hat{X}_{\mathrm{MALS}} - \hat{X}_{\mathrm{ALS}}\right) = -\sqrt{m}(1-\mu) V_{A_0}^{-1} V_{\tilde{A}} \hat{X}_{\mathrm{ALS}}$$

$$\begin{aligned}
V_{A_0}^{-1}\left(A^T A - \mu V_{\tilde{A}}\right) &= V_{A_0}^{-1}\left(A^T A - V_{\tilde{A}}\right) + (1-\mu) V_{A_0}^{-1} V_{\tilde{A}} \\
&= V_{A_0}^{-1}\left(A^T A - V_{\tilde{A}}\right) + (1-\mu)\left(\frac{1}{m} V_{A_0}\right)^{-1} \operatorname{Cov}(\tilde{a}) \\
&\xrightarrow[m\to\infty]{\mathsf{P}} I_n + 0 \cdot V_{A_\infty}^{-1} \operatorname{Cov}(\tilde{a}) = I_n - \sqrt{m}(1-\mu) V_{A_0}^{-1} V_{\tilde{A}} \hat{X}_{\mathrm{ALS}} \\
&= -\sqrt{m}(1-\mu)\left(\frac{1}{m} V_{A_0}\right)^{-1} \operatorname{Cov}(\tilde{a}) \hat{X}_{\mathrm{ALS}} \\
&\xrightarrow[m\to\infty]{\mathsf{P}} 0 \cdot V_{A_\infty}^{-1} \operatorname{Cov}(\tilde{a}) X_0 = 0.
\end{aligned}$$

Отже, враховуючи (3), маємо $\sqrt{m}\left(\hat{X}_{\mathrm{MALS}} - \hat{X}_{\mathrm{ALS}}\right) \xrightarrow{\mathsf{P}} 0$, $m \to \infty$. □

5. Висновки

У даній роботі отримано умови асимптотичної нормальності покращеної оцінки найменших квадратів та наведено консистентну оцінку відповідної асимптотичної коваріаційної матриці. Також було побудовано модифікацію даної оцінки, яка обчислюється за допомогою оберненої матриці, яка існує з ймовірністю 1 вже при невеликій кількості вимірювань. Ця модифікація при цьому зберігає асимптотичні властивості немодифікованої оцінки.

6. Додаток

6.1. **Доведення збіжності** (5)**.** Позначимо

$$\begin{aligned}
\eta_i &= \frac{1}{\sqrt{m}} \operatorname{vec}\left(a_i \tilde{b}_i^T + \mathsf{E}\, \tilde{a}_i \tilde{a}_i^T X_0 - a_i \tilde{a}_i^T X_0\right) \\
&= \frac{1}{\sqrt{m}} \operatorname{vec}\left(a_i \tilde{b}_i^T - a_i^0 \tilde{a}_i^T X_0 - \left(\tilde{a}_i \tilde{a}_i^T - \mathsf{E}\, \tilde{a}_i \tilde{a}_i^T\right) X_0\right).
\end{aligned}$$

Тоді

$$\frac{l}{\sqrt{m}} = \sum_{i=1}^{m} \eta_i, \qquad \mathsf{E}\, \eta_i = 0.$$

Перевіримо умови центральної граничної теореми у формі Ляпунова для випадкових векторів, теорема 19 у [8, гл. IV.4]. Доведемо, що існує $\lim_{m\to\infty}\sum_{i=1}^{m}\mathrm{Cov}\,(\eta_i)$.

Будемо позначати $\xi_1=\xi_1(i)=\mathrm{vec}\left(a_i\tilde b_i^T\right)$, $\xi_2=\xi_2(i)=\mathrm{vec}\left(a_i^0\tilde a_i^TX_0\right)$, $\xi_3=\xi_3(i)=\mathrm{vec}\left(\left(\tilde a_i\tilde a_i^T-\mathsf{E}\,\tilde a_i\tilde a_i^T\right)X_0\right)$. Тоді $\eta_i=\frac{1}{\sqrt m}(\xi_1-\xi_2-\xi_3)$ та

$$
\begin{aligned}
\mathrm{Cov}(\eta_i)&=\frac{1}{m}\,\mathsf{E}(\xi_1-\xi_2-\xi_3)(\xi_1-\xi_2-\xi_3)^T\\
&=\frac{1}{m}\left(\mathsf{E}\,\xi_1\xi_1^T+\mathsf{E}\,\xi_2\xi_2^T+\mathsf{E}\,\xi_3\xi_3^T+\mathsf{E}\,\xi_2\xi_3^T+\mathsf{E}\,\xi_3\xi_2^T\right)\\
&=\mathsf{E}\,\xi_1\xi_1^T+\mathsf{E}\left(\xi_2+\xi_3\right)\left(\xi_2+\xi_3\right)^T.
\end{aligned}\tag{10}
$$

З (i) випливає, що $\mathsf{E}\,\xi_1\xi_2^T=\mathsf{E}\,\xi_2\xi_1^T=\mathsf{E}\,\xi_1\xi_3^T=\mathsf{E}\,\xi_3\xi_1^T=0$.

Розглянемо доданки у (10) окремо, використовуючи властивості кронекерівського добутку [5, розд. 12]. Перший доданок:

$$
\begin{aligned}
\mathsf{E}\,\xi_1\xi_1^T&=\mathsf{E}\,\mathrm{vec}\left(a_i\tilde b_i^T\right)\mathrm{vec}^T\left(a_i\tilde b_i^T\right)=\mathsf{E}\left(\tilde b_i\otimes a_i\right)\left(\tilde b_i^T\otimes a_i^T\right)=\mathsf{E}\left(\tilde b_i\tilde b_i^T\otimes a_ia_i^T\right)\\
&=\mathsf{E}\left(\tilde b_i\tilde b_i^T\otimes a_i^0a_i^{0T}+\tilde b_i\tilde b_i^T\otimes\tilde a_i\tilde a_i^T\right)=V_{\tilde b}\otimes a_i^0a_i^{0T}+V_{\tilde b}\otimes V_{\tilde a}.
\end{aligned}
$$

Другий доданок:

$$
\begin{aligned}
\mathsf{E}\,\xi_2\xi_2^T&=\mathsf{E}\,\mathrm{vec}\left(a_i^0\tilde a_i^TX_0\right)\mathrm{vec}^T\left(a_i^0\tilde a_i^TX_0\right)=\mathsf{E}\left(\left(X_0^T\tilde a_i\right)\otimes a_i^0\right)\left(\left(X_0^T\tilde a_i\right)\otimes a_i^0\right)^T\\
&=\mathsf{E}\left(\left(X_0^T\tilde a_i\right)\otimes\left(I_na_i^0\right)\right)\left(\left(X_0^T\tilde a_i\right)\otimes\left(I_na_i^0\right)\right)^T\\
&=\mathsf{E}\Big(\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes a_i^0\right)\Big)\Big(\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes a_i^0\right)\Big)^T\\
&=\left(X_0^T\otimes I_n\right)\mathsf{E}\Big[\left(\tilde a_i\tilde a_i^T\right)\otimes\left(a_i^0a_i^{0T}\right)\Big]\left(X_0^T\otimes I_n\right)^T.
\end{aligned}
$$

Третій доданок:

$$
\begin{aligned}
\mathsf{E}\,\xi_3\xi_3^T&=\mathsf{E}\,\mathrm{vec}\left(\left(\tilde a_i\tilde a_i^T-\mathsf{E}\,\tilde a_i\tilde a_i^T\right)X_0\right)\mathrm{vec}^T\left(\left(\tilde a_i\tilde a_i^T-\mathsf{E}\,\tilde a_i\tilde a_i^T\right)X_0\right)\\
&=\mathsf{E}\Big(\left(X_0^T\tilde a_i\right)\otimes\tilde a_i-\mathsf{E}\left(X_0^T\tilde a_i\right)\otimes\tilde a_i\Big)\Big(\left(X_0^T\tilde a_i\right)^T\otimes\tilde a_i^T-\mathsf{E}\left(X_0^T\tilde a_i\right)^T\otimes\tilde a_i^T\Big)\\
&=\mathsf{E}\left[\Big(\left(X_0^T\tilde a_i\right)\otimes\left(I_n\tilde a_i\right)\Big)\Big(\left(X_0^T\tilde a_i\right)^T\otimes\left(I_n\tilde a_i\right)^T\Big)\right]\\
&\quad-\left[\mathsf{E}\left(X_0^T\tilde a_i\right)\otimes\left(I_n\tilde a_i\right)\right]\left[\mathsf{E}\left(X_0^T\tilde a_i\right)^T\otimes\left(I_n\tilde a_i\right)^T\right]\\
&=\mathsf{E}\left[\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes\tilde a_i\right)\right]\left[\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes\tilde a_i\right)\right]^T\\
&\quad-\mathsf{E}\left[\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes\tilde a_i\right)\right]\mathsf{E}\left[\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes\tilde a_i\right)\right]^T\\
&=\left(X_0^T\otimes I_n\right)\Big[\mathsf{E}\left(\tilde a_i\otimes\tilde a_i\right)\left(\tilde a_i\otimes\tilde a_i\right)-\mathrm{vec}\left(V_{\tilde a}\right)\mathrm{vec}^T(V_{\tilde a})\Big]\left(X_0^T\otimes I_n\right)^T.
\end{aligned}
$$

Четвертий доданок:

$$
\begin{aligned}
\mathsf{E}\,\xi_2\xi_3^T&=\mathsf{E}\,\mathrm{vec}\left(a_i^0\tilde a_i^TX_0\right)\mathrm{vec}^T\left(\left(\tilde a_i\tilde a_i^T-\mathsf{E}\,\tilde a_i\tilde a_i^T\right)X_0\right)\\
&=\mathsf{E}\left(\left(X_0^T\tilde a_i\right)\otimes a_i^0\right)\Big(\left(X_0^T\tilde a_i\right)^T\otimes\tilde a_i^T-\mathsf{E}\left(X_0^T\tilde a_i\right)^T\otimes\tilde a_i^T\Big)\\
&=\mathsf{E}\left[\Big(\left(X_0^T\tilde a_i\right)\otimes\left(I_na_i^0\right)\Big)\Big(\left(X_0^T\tilde a_i\right)^T\otimes\left(I_n\tilde a_i\right)^T\Big)\right]\\
&\quad-\mathsf{E}\Big[\left(X_0^T\tilde a_i\right)\otimes\left(I_na_i^0\right)\Big]\mathsf{E}\Big[\left(X_0^T\tilde a_i\right)^T\otimes\left(I_n\tilde a_i\right)^T\Big]\\
&=\mathsf{E}\left[\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes a_i^0\right)\right]\left[\left(X_0^T\otimes I_n\right)\left(\tilde a_i\otimes\tilde a_i\right)\right]^T\\
&=\left(X_0^T\otimes I_n\right)\mathsf{E}\left[\left(\tilde a_i\otimes a_i^0\right)\left(\tilde a_i\otimes\tilde a_i\right)\right]\left(X_0^T\otimes I_n\right)^T.
\end{aligned}
$$

Аналогічно, п'ятий доданок:

$$\mathsf{E}\,\xi_3\xi_2^T = \left(X_0^T \otimes I_n\right) \mathsf{E}\left[\left(\tilde{a}_i \otimes a_i^0\right)\left(\tilde{a}_i \otimes \tilde{a}_i\right)\right]^T \left(X_0^T \otimes I_n\right)^T.$$

Оскільки, з урахуванням припущення (viii), перший – третій доданки не залежать від i, то знайдемо границю лише для четвертого, а разом з тим і п'ятого доданків:

$$\begin{aligned}
\lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} \mathsf{E}\,\xi_2\xi_3^T &= \left(X_0^T \otimes I_n\right)\left[\lim_{m\to\infty}\sum_{i=1}^{m}\frac{1}{m}\,\mathsf{E}\left[\left(\tilde{a}_i \otimes a_i^0\right)\left(\tilde{a}_i \otimes \tilde{a}_i\right)\right]\right]\left(X_0^T \otimes I_n\right)^T \\
&= \left(X_0^T \otimes I_n\right)\left[\lim_{m\to\infty}\sum_{i=1}^{m}\frac{1}{m}\,\mathsf{E}\left[\left(\tilde{a} \otimes a_i^0\right)\left(\tilde{a} \otimes \tilde{a}\right)\right]\right]\left(X_0^T \otimes I_n\right)^T \\
&= \left(X_0^T \otimes I_n\right)\mathsf{E}\left[\left(\tilde{a} \otimes \left[\lim_{m\to\infty}\frac{1}{m}\sum_{i=1}^{m} a_i^0\right]\right)\left(\tilde{a} \otimes \tilde{a}\right)\right]\left(X_0^T \otimes I_n\right)^T \\
&= \left(X_0^T \otimes I_n\right)\mathsf{E}\left[\left(\tilde{a} \otimes a^0\right)\left(\tilde{a} \otimes \tilde{a}\right)\right]\left(X_0^T \otimes I_n\right)^T.
\end{aligned}$$

Тепер можна розглянути суму

$$\begin{aligned}
\sum_{i=1}^{m} \operatorname{Cov}(\eta_i) = \frac{1}{m}\sum_{i=1}^{m}\Big(& V_{\tilde{b}} \otimes \left(a_i^0 a_i^{0T} + V_{\tilde{a}}\right) \\
&+ \left(X_0^T \otimes I_n\right)\mathsf{E}\left(\tilde{a}_i\tilde{a}_i^T\right)\left(a_i^0 a_i^{0T}\right)\left(X_0^T \otimes I_n\right)^T \\
&+ \left(X_0^T \otimes I_n\right)\Big[\mathsf{E}\left(\tilde{a}_i \otimes \tilde{a}_i\right)\left(\tilde{a}_i \otimes \tilde{a}_i\right) - \operatorname{vec}(V_{\tilde{a}})\operatorname{vec}^T(V_{\tilde{a}})\Big] \\
&\quad \times \left(X_0^T \otimes I_n\right)^T \\
&+ \left(X_0^T \otimes I_n\right)\mathsf{E}\left[\left(\tilde{a}_i \otimes a_i^0\right)\left(\tilde{a}_i \otimes \tilde{a}_i\right)\right]\left(X_0^T \otimes I_n\right)^T \\
&+ \left(X_0^T \otimes I_n\right)\mathsf{E}\left[\left(\tilde{a}_i \otimes a_i^0\right)\left(\tilde{a}_i \otimes \tilde{a}_i\right)\right]^T \\
&\qquad\qquad\qquad \times \left(X_0^T \otimes I_n\right)^T\Big) \qquad (11)
\end{aligned}$$

$$\begin{aligned}
\xrightarrow[m\to\infty]{} \; & V_{\tilde{b}} \otimes V_{A_\infty} + V_{\tilde{b}} \otimes V_{\tilde{a}} \\
& + \left(X_0^T \otimes I_n\right) \\
& \quad \times \Big[V_{\tilde{a}} \otimes V_{A_\infty} + \mathsf{E}\left(\tilde{a} \otimes \tilde{a}\right)\left(\tilde{a} \otimes \tilde{a}\right) - \operatorname{vec}(V_{\tilde{a}})\operatorname{vec}^T(V_{\tilde{a}}) \\
& \qquad\quad + \mathsf{E}\left[\left(\tilde{a} \otimes a^0\right)\left(\tilde{a} \otimes \tilde{a}\right)\right] + \mathsf{E}\left[\left(\tilde{a} \otimes a^0\right)\left(\tilde{a} \otimes \tilde{a}\right)\right]^T\Big] \\
& \quad \times \left(X_0^T \otimes I_n\right)^T \\
& \stackrel{\text{def}}{=} \Sigma_1.
\end{aligned}$$

Зокрема тоді, $\|\operatorname{Cov}(\eta_m)\| \to 0$, $m \to \infty$, а отже для будь-якого $x \in \mathbb{R}^{nd}$

$$x^T \operatorname{Cov}(\eta_m)\, x \to 0, \qquad m \to \infty.$$

Доведемо, що для деякого $\nu > 0$ виконується умова Ляпунова

$$\sum_{i=1}^{m} \mathsf{E}\,\|\eta_i\|^{2+\nu} \to 0, \qquad m \to \infty. \qquad (12)$$

Оцінимо суму

$$\sum_{i=1}^{m} \| \mathsf{E}\, \eta_i \|^{2+\nu} = \frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\, \|\xi_1 - \xi_2 - \xi_3\|^{2+\nu}$$
$$\leqslant \frac{3^{2+\nu}}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\left(\|\xi_1\|^{2+\nu} + \|\xi_2\|^{2+\nu} + \|\xi_3\|^{2+\nu}\right).$$

Із використанням умов (viii), (ix) та (x) оцінимо доданки окремо для $0 < \nu \leqslant \max(\delta, \tau)$.

Перший доданок:

$$\frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\, \|\xi_1\|^{2+\nu} = \frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E} \left\|\tilde{b}_i \otimes a_i\right\|^{2+\nu}$$
$$\leqslant \frac{2^{2+\nu}}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\left(\left\|\tilde{b}_i \otimes a_i^0\right\|^{2+\nu} + \left\|\tilde{b}_i \otimes \tilde{a}_i\right\|^{2+\nu}\right),$$

$$\left\|\tilde{b}_i \otimes a_i^0\right\|^{2+\nu} = \left(\sum_{j=1}^{n} \sum_{k=1}^{d} (a_{ij}^0 \tilde{b}_{ik})^2\right)^{1+\nu/2} = \left(\left\|a_i^0\right\|^2 \cdot \|\tilde{b}_i\|^2\right)^{1+\nu/2} = \left\|a_i^0\right\|^{2+\nu} \cdot \|\tilde{b}_i\|^{2+\nu},$$

$$\left\|\tilde{b}_i \otimes \tilde{a}_i\right\|^{2+\nu} = \|\tilde{a}_i\|^{2+\nu} \cdot \|\tilde{b}_i\|^{2+\nu}.$$

Тоді

$$\frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\, \|\xi_1\|^{2+\nu} = \frac{2^{2+\nu}}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\left(\left\|a_i^0\right\|^{2+\nu} \cdot \mathsf{E}\, \|\tilde{b}\|^{2+\nu} + \mathsf{E}\, \|\tilde{a}\|^{2+\nu} \cdot \mathsf{E}\, \|\tilde{b}\|^{2+\nu}\right)$$
$$\leqslant \frac{\text{const}}{m^{\nu/2}} \to 0, \qquad m \to \infty.$$

Другий доданок:

$$\frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\, \|\xi_2\|^{2+\nu} = \frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \left\|\left(X_0^T a_i^0\right) \otimes \tilde{a}_i\right\|^{2+\nu}$$
$$= \frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \left\|X_0^T a_i^0\right\|^{2+\nu} \mathsf{E}\, \|\tilde{a}_i\|^{2+\nu} \leqslant \frac{\text{const}}{m^{\nu/2}} \to 0, \qquad m \to \infty.$$

Третій доданок:

$$\frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E}\, \|\xi_3\|^{2+\nu} = \frac{1}{m^{1+\nu/2}} \sum_{i=1}^{m} \mathsf{E} \left\|\left(X_0^T \tilde{a}_i\right) \otimes \tilde{a}_i - \mathsf{E}\left(X_0^T \tilde{a}_i\right) \otimes \tilde{a}_i\right\|^{2+\nu}$$
$$= \frac{1}{m^{\nu/2}} \mathsf{E} \left\|\left(X_0^T \tilde{a}\right) \otimes \tilde{a} - \mathsf{E}\left(X_0^T \tilde{a}\right) \otimes \tilde{a}\right\|^{2+\nu} \to 0, \qquad m \to \infty.$$

Отже, збіжність у (12) має місце, і виконуються умови центральної граничної теореми, а тому

$$\frac{l}{\sqrt{m}} \xrightarrow{d} N\left(0, \Sigma_1\right).$$

6.2. **Доведення пункту 2 теореми 3.1.** З (vii) та (xii) випливає, що

$$V_{\tilde{b}} \otimes V_{A_\infty} + V_{\tilde{b}} \otimes V_{\tilde{a}} > 0,$$

а отже, враховуючи (10), $\Sigma_1 > 0$. Тоді з (vii) та (6) отримуємо $\Sigma > 0$.

6.3. **Доведення леми 4.3.** Розглянемо $T - W = \begin{bmatrix} B^T B & B^T A \\ A^T B & A^T A - V_{\tilde{A}} \end{bmatrix}$. Із (i)–(iii), (viii) випливає, що

$$m^{-1}\left(A^T A - V_{\tilde{A}}\right) \xrightarrow[m\to\infty]{\mathsf{P}} V_{A_\infty}, \qquad m^{-1} A^T B \xrightarrow[m\to\infty]{\mathsf{P}} V_{A_\infty} X_0,$$
$$m^{-1} B^T A \xrightarrow[m\to\infty]{\mathsf{P}} X_0 V_{A_\infty}.$$

Розглянемо $m^{-1} B^T B$:

$$\frac{1}{m} B^T B = \frac{1}{m}\sum_{i=1}^m \left(b_i^0 b_i^{0T} + b_i^0 \tilde{b}_i^T + \tilde{b}_i b_i^{0T} + \tilde{b}_i \tilde{b}_i^T\right),$$

$$\frac{1}{m}\sum_{i=1}^m b_i^0 b_i^{0T} = \frac{1}{m} B_0^T B_0 = \frac{1}{m} A_0 X_0 X_0^T A_0^T \to X_0 V_{A_\infty} X_0^T, \qquad m \to \infty.$$

Останнє випливає з припущення (vii).

Із (ii), (viii) та посиленого закону великих чисел, теорема 3 з [9, гл. IV.3], маємо $m^{-1}\sum_{i=1}^m \tilde{b}_i \tilde{b}_i^T \xrightarrow{\mathsf{P1}} \operatorname{Cov}(\tilde{b}) = V_{\tilde{b}}$, $m \to \infty$.

Доведемо, що $m^{-1}\sum_{i=1}^m b_i^0 \tilde{b}_i \xrightarrow{\mathsf{P}} 0$, $m \to \infty$ та $m^{-1}\sum_{i=1}^m \tilde{b}_i b_i^0 \xrightarrow{\mathsf{P}} 0$, $m \to \infty$. Застосуємо закон великих чисел Чебишова до $m^{-1}\sum_{i=1}^m b_i^0 \tilde{b}_i$. Маємо $\forall k \geqslant 1$

$$b_k^0 \tilde{b}_k^T = \left(b_{kl}^0 \tilde{b}_{kl}\right)_{i,j=1}^d, \qquad \mathsf{E}\, b_{kl}^0 \tilde{b}_{kj} = 0, \qquad \operatorname{Var} b_{kl}^0 \tilde{b}_{kj} = 0, \qquad \operatorname{Var} \tilde{b}_{kj} \leqslant \left(b_{kl}^0\right)^2 \sup_{k,j} \mathsf{E}\, \tilde{b}_{kl}^2.$$

Отже, достатньо показати, що $\forall i = 1, \dots, d$

$$\frac{\left(b_{kl}^0\right)^2}{k} \to 0, \qquad k \to \infty,$$

або те саме, що

$$\frac{\left\|b_k^0\right\|^2}{k} \to \infty, \qquad k \to \infty.$$

Дійсно,

$$\left\|b_k^0\right\|^2 = \left\|X_0^T a_k^0\right\|^2 \leqslant \|X_0\|_{\mathsf{F}}^2 \left\|a_k^0\right\|^2.$$

Позначимо $A_{0,m} = \begin{bmatrix} a_1^{0T} & \dots & a_m^{0T} \end{bmatrix}$. Тоді

$$\frac{1}{m}\left\|a_m^0\right\|^2 = \frac{1}{m}\left(\operatorname{tr} A_{0,m}^T A_{0,m} - \operatorname{tr} A_{0,m-1}^T A_{0,m-1}\right) \xrightarrow[m\to\infty]{} \operatorname{tr} V_{A_\infty} - \operatorname{tr} V_{A_\infty} = 0.$$

Отже, за законом великих чисел Чебишова $m^{-1}\sum_{i=1}^m b_i^0 \tilde{b}_i \xrightarrow{\mathsf{P}} 0$, $m \to \infty$, а тому і $m^{-1}\sum_{i=1}^m \tilde{b}_i b_i^0 \xrightarrow{\mathsf{P}} 0$, $m \to \infty$. Остаточно маємо:

$$\begin{aligned} \frac{1}{m}(T - W) &= \frac{1}{m}\begin{bmatrix} B^T B & B^T A \\ A^T B & A^T A - V_{\tilde{A}} \end{bmatrix} \\ &\xrightarrow[m\to\infty]{\mathsf{P}} \begin{bmatrix} X_0^T V_{A_\infty} X_0 + V_{\tilde{b}} & X_0^T V_{A_\infty} \\ V_{A_\infty} X_0 & V_{A_\infty} \end{bmatrix} = \begin{bmatrix} X_0 & I_n \end{bmatrix}^T V_{A_\infty} \begin{bmatrix} X_0 & I_n \end{bmatrix} + \begin{bmatrix} V_{\tilde{b}} & 0 \\ 0 & 0 \end{bmatrix} \\ &\stackrel{\text{def}}{=} S_\infty. \end{aligned}$$

Доведемо, що $S_\infty > 0$. Справді, $\forall x \in \mathbb{R}^{n+d}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $x_1 \in \mathbb{R}^n$, $x_2 \in \mathbb{R}^d$

$$\begin{aligned} x^T V_{A_\infty} x &= x^T \begin{bmatrix} X_0 & I_n \end{bmatrix}^T V_{A_\infty} \begin{bmatrix} X_0 & I_n \end{bmatrix} x + x^T \begin{bmatrix} V_{\tilde{b}} & 0 \\ 0 & 0 \end{bmatrix} x \\ &= \left\|V_{A_\infty}(X_0 x_1 + x_2)\right\|^2 + x_1^T V_{\tilde{b}} x_1 \geqslant 0. \end{aligned}$$

Тоді

$$x^T S_\infty x = 0 \Leftrightarrow \begin{cases} X_0 x_1 + x_2 = 0, \\ x_1 = 0. \end{cases} \Leftrightarrow \begin{cases} x_1 = 0, \\ x_2 = 0. \end{cases}$$

Отже, $\forall x \in \mathbb{R}^{n+d} \setminus \{0\}$

$$x^T S_\infty x > 0.$$

Тоді

$$\begin{aligned} &\mathsf{P}\left(\frac{1}{m}(T-W) > 0\right) \to 1,\ m \to \infty \\ &\Leftrightarrow \mathsf{P}\left((T-W) > 0\right) \to 1,\ m \to \infty \\ &\Leftrightarrow \mathsf{P}\left(\lambda_A > 1\right) \to 1, m \to \infty. \end{aligned}$$

Далі,

$$\begin{aligned} \{\lambda_A > 1\} &\subset \left\{\frac{m-\alpha}{m+1} < \mu < \frac{m-\alpha}{m}\right\} \\ &= \left\{\frac{\alpha}{m} \leqslant 1 - \mu < \frac{1+\alpha}{m+1}\right\} \\ &= \left\{\frac{\alpha}{m} \leqslant \sqrt{m}(1-\mu) < (1+\alpha)\frac{\sqrt{m}}{m+1}\right\}. \end{aligned}$$

Звідси отримуємо, що $\sqrt{m}(1-\mu) \xrightarrow{\mathsf{P}} 0$, $m \to \infty$.

Література

1. R. J. Carrol, D. Ruppert, and L. A. Sefanski, *Measurement Error in Nonlinear Models*, Monographs on Statistics and Applied Probability, vol. 63, Chapman & Hall/CRC, London, 1995.
2. C.-L. Cheng, H. Schneeweiss, and M. Thamerus, *A small sample estimator for a polynomial regression with errors in the variables*, J. R. Statist. Soc. B **62** (2000), no. 4, 699–709.
3. W. A. Fuller, *Measurement Error Models*, Wiley, New York, 1987.
4. P. P. Gallo, *Consistency of regression estimates when some variables are subject to error*, Comm. Statist. B-Theory Methods **11** (1982), 973–983.
5. Р. Беллман, *Введение в теорию матриц*, “Мир”, Москва, 1969.
6. Ф. Р. Гантмахер, *Теория матриц*, “Наука”, Москва, 1988.
7. І. О. Сенько, *Консистентність покращеної оцінки найменших квадратів у векторній лінійній моделі з похибками вимірювання.* (подано до друку)
8. В. В. Петров, *Предельные теоремы для сумм независимых случайных величин*, “Наука”, Москва, 1987.
9. А. Н. Ширяев, *Вероятность*, “Наука”, Москва, 1980.

Кафедра математичного аналізу, механіко-математичний факультет, Національний університет ім. Тараса Шевченка, проспект Глушкова, 6, Київ 03127, Україна

Адреса електронної пошти: ivan_senko@ukr.net

Надійшла 25/10/2012