УДК 519.21

І. О. Сенько (Київ. нац. ун-т ім. Т. Шевченка)

КОНСИСТЕНТНІСТЬ ПОКРАЩЕНОЇ ОЦІНКИ НАЙМЕНШИХ КВАДРАТІВ У ВЕКТОРНІЙ ЛІНІЙНІЙ МОДЕЛІ З ПОХИБКАМИ ВИМІРЮВАННЯ

We consider the vector linear errors-in-variables model. For this model, we construct an adjusted least-squares estimator and prove its weak and strong consistency under various assumptions about measurement errors.

Рассматривается векторная линейная модель с погрешностями в переменных, для которой построена исправленная оценка наименьших квадратов. Доказаны слабая и строгая состоятельность этой оценки при различных предположениях о погрешностях измерений.

1. Вступ. Моделі з похибками в змінних є важливим узагальненням регресійних моделей, оскільки існує велика кількість практичних застосувань, де всі дані, що є 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 в задачах геологічних та океанографічних досліджень. Загальний огляд моделей з похибками в змінних можна знайти в [1]. Для коректного розв'язування подібних задач необхідно враховувати наявність похибок як у залежних, так і в незалежних змінних.

Розв'язання матричного рівняння $AX = B$ при великій кількості рядків у матрицях A, B по суті є розв'язанням переозначеної системи лінійних рівнянь. Подібна модель для одновимірного випадку, коли X є вектором, розглядалася, зокрема, у [2]. Якщо сукупна коваріаційна матриця адитивних похибок у спостережених матрицях A, B відома з точністю до сталого множника, то отримують оцінку повних найменших квадратів [3] або її модифікації. Наприклад, зважену оцінку найменших квадратів розглянуто в [4]. Оптимальність покращеної оцінки найменших квадратів для одновимірного випадку було доведено в [5]. Також оцінки, побудовані подібними способами, розглядалися в [6] для поліноміальної моделі та в [7] для білінійної векторної моделі. У випадку, коли жодної інформації про кореляційну матрицю похибок у наявності немає, можна використати ідею кластеризації, як це було зроблено у [8].

У даній роботі розглянуто переозначену систему лінійних рівнянь $AX = B$. Сукупна коваріаційна матриця похибок у спостереженнях матриці A є відомою, а матриці B — невідомою. Вивчається покращена оцінка найменших квадратів та її консистентність, коли кількість рядків у матриці A нескінченно зростає.

Надалі будемо використовувати 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 через $\|Z\|$ операторну норму 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 [9] до матриці Z. Математичне сподівання та дисперсію будемо позначати 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 $X_m = \mathsf{o_P}(1)$, $m \to \infty$, означає, що $X_m \xrightarrow{\mathsf{P}} 0$, $m \to \infty$, а $X_m = \mathsf{O_P}(1)$,

$m \to \infty$, означає, що послідовність випадкових величин $\{\|X_m\|\}$ стохастично обмежена. Для квадратної симетричної матриці V через $\lambda_{\min}(V)$ та $\lambda_{\max}(V)$ будемо позначати відповідно її найменше та найбільше власні числа. Через const будемо позначати довільну сталу, яка не залежить від кількості вимірювань m.

Опишемо коротко будову статті. У пункті 2 наведено опис моделі спостережень та показано зв'язок множинної векторної моделі регресії з матричним рівнянням. У пункті 3 побудовано покращену оцінку найменших квадратів за умови, що є відомою сукупна кореляційна матриця похибок спостережень прихованих змінних. У пункті 4 наведено умови, за яких ця оцінка є слабко та строго консистентною. Пункт 5 присвячено дослідженню швидкості збіжності оцінки. Доведення теорем винесено у додаток.

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; $b_i^0 \in \mathbb{R}^d$, $i \geqslant 1$, — вектори, що задають істинні значення вихідних даних; елементи цих векторів позначатимемо відповідно a_{ij}^0 та b_{ij}^0. Для істинних значень виконується рівність

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

Будемо спостерігати за векторами $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}^d$, $1 \leqslant i \leqslant m$, 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 похибки вимірювань. Елементи цих векторів будемо позначати аналогічно через a_{ij}, $\tilde{a}_{ij}$, b_{ij}, $\tilde{b}_{ij}$. Тоді векторна лінійна модель з похибками у змінних задається рівностями

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

Використовуючи позначення $A = \begin{bmatrix} a_1 & a_2 & \ldots & a_m \end{bmatrix}^T = (a_{ij})_{i=1,\ j=1}^{m\quad\ n}$ 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

$$AX \approx B$$

за умови, що

$$A_0 X_0 = B_0.$$

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

Основним припущенням щодо 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.*

3. Побудова оцінки. Скористаємось методом виправленої оціночної функції, описаним у [10] (розд. 7). Розпочнемо з цільової функції, яка відповідає методу найменших квадратів:

$$q_{\mathrm{LS}}\left(X,A,B\right)\stackrel{\mathrm{df}}{=}\|AX-B\|_{\mathrm{F}}^{2}.$$

У просторі матриць $\mathbb{R}^{n\times d}$ розглянемо скалярний добуток

$$\left(U,V\right)\stackrel{\mathrm{df}}{=}\mathrm{tr}\left(UV^{T}\right),\qquad U,V\in\mathbb{R}^{m\times d}.$$

Тоді норма Фробеніуса $\|\cdot\|_{\mathrm{F}}$ буде породжуватися цим скалярним добутком. Похідна $\dfrac{\partial q_{\mathrm{LS}}}{\partial X}$ є лінійним функціоналом на $\mathbb{R}^{n\times d}$,

$$\frac{1}{2}\frac{\partial q_{\mathrm{LS}}}{\partial X}\left(H\right)=\mathrm{tr}\left(A^{T}\left(AX-B\right)H^{T}\right)=\left(A^{T}\left(AX-B\right),H\right),\quad H\in\mathbb{R}^{n\times d}.$$

За відсутності похибок у змінних, тобто коли $\tilde{A}=0$ та $\tilde{B}=0$, оцінку найменших квадратів отримуємо за допомогою мінімізації цільової функції q_{LS}, а отже, оцінку можна обчислити з рівняння $\dfrac{\partial q_{\mathrm{LS}}}{\partial X}=0$. Тому оціночною функцією для методу найменших квадратів буде функція

$$\psi_{\mathrm{LS}}\left(X,A,B\right)\stackrel{\mathrm{df}}{=}A^{T}\left(AX-B\right),$$

і за необтяжливих умов оцінка найменших квадратів буде консистентною при $m\to\infty$, якщо похибки у вимірюваннях відсутні.

Використовуючи вищезгаданий метод виправленої оціночної функції, будемо шукати нову оціночну функцію $\psi\left(X,A,B\right)$ так, щоб виконувалося співвідношення

$$\mathsf{E}\left(\psi\left(X,A_0+\tilde{A},B\right)\Big|\,B\right)=\psi_{\mathrm{LS}}\left(X,A_0,B\right)\quad\forall X,A_0. \tag{1}$$

Розв'язком (1) у класі поліноміальних матричнозначних функцій є функція

$$\psi=A^{T}\left(AX-B\right)-\mathsf{E}\,\tilde{A}^{T}\tilde{A}X=\left(A^{T}A-V_{\tilde{A}}\right)X-A^{T}B.$$

Тоді оцінка $\hat{X}=\hat{X}_m$ визначається з рівняння

$$\psi\big(\hat{X},A,B\big)=0,\qquad\hat{X}\in\mathbb{R}^{n\times d}. \tag{2}$$

Точніше, покладемо

$$\hat{X}=\left(A^{T}A-V_{\tilde{A}}\right)^{\dagger}A^{T}B. \tag{3}$$

Цю оцінку називатимемо *покращеною оцінкою найменших квадратів*.

Якщо матриця $A^TA-V_{\tilde{A}}$ невироджена, то $\hat{X}$ в точності задовольняє рівняння (2) і є його єдиним розв'язком. При доведенні теорем про консистентність буде показано, що при необмеженому зростанні m ця матриця буде невиродженою з імовірністю, що прямує до одиниці.

4. Слабка та строга консистентність оцінки. Для доведення консистентності будемо використовувати наступні припущення:

(ii) *рядки матриці $\tilde{A}$ є незалежними випадковими векторами, тобто незалежними є вектори $\{\tilde{a}_i,i\geqslant1\}$, та рядки матриці $\tilde{B}$ незалежні, тобто незалежними є вектори $\{\tilde{b}_i,i\geqslant1\}$;*

(iii) $\mathsf{E}\,\tilde{a}_{ij}^4 \leqslant \mathrm{const}$; $\mathsf{E}\,\tilde{b}_{il}^2 \leqslant \mathrm{const}$, $i \geqslant 1$, $1 \leqslant j \leqslant n$, $1 \leqslant j \leqslant n$;

(iii′) $\exists i_0 > 1\ \forall i \geqslant i_0$: *компоненти векторів* $\tilde{a}_i$ *попарно незалежні, а також* $\mathsf{E}\,\tilde{a}_{ij}^2 \leqslant \mathrm{const}$; $\mathsf{E}\,\tilde{b}_{il}^2 \leqslant \mathrm{const}$, $i \geqslant 1$, $1 \leqslant j \leqslant n$, $1 \leqslant j \leqslant n$;

(iv) $\dfrac{\lambda_{\max}(V_{A_0}) + m}{\lambda_{\min}^2(V_{A_0})} \to 0$, $m \to \infty$, *де* $V_{A_0} \stackrel{\mathrm{df}}{=} A_0^T A_0$.

***Зауваження* 1.** Умова (iv) — це умова П. Галло, яку вперше для подібних задач розглянуто в [11].

Теорема 1. *Нехай виконуються припущення* (ii), (iv) *та одне з припущень* (iii) *або* (iii)′. *Тоді* $\hat{X} \xrightarrow{\mathsf{P}} X_0$, $m \to \infty$.

Доведення. Перетворимо рівняння (3) таким чином:

$$\hat{X} = \left(A^T A - V_{\tilde{A}}\right)^{\dagger} A^T B = \left(A^T A - V_{\tilde{A}}\right)^{\dagger} A^T \big(B_0 + \tilde{B}\big).$$

Якщо існує $\left(A^T A - V_{\tilde{A}}\right)^{-1}$, то це рівняння можна записати у вигляді

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

Із припущення (iv) випливає, що існує $m_0 \in \mathbb{N}$ таке, що для будь-якого $m \geqslant m_0$ виконується $\det V_{A_0} \neq 0$. Тоді для $m \geqslant m_0$

$$V_{A_0}^{-1}\left(A^T A - V_{\tilde{A}}\right) \hat{X} = V_{A_0}^{-1}\left(A^T A_0 X_0 + A^T \tilde{B}\right). \tag{4}$$

Для консистентності оцінки (3) достатньо показати, що при $m \to \infty$

$$V_{A_0}^{-1}\left(A^T A - V_{\tilde{A}}\right) \xrightarrow{\mathsf{P}} I_n, \tag{5}$$

$$V_{A_0}^{-1}\left(A^T A_0\right) \xrightarrow{\mathsf{P}} I_n, \tag{6}$$

$$V_{A_0}^{-1} A^T \tilde{B} \xrightarrow{\mathsf{P}} 0. \tag{7}$$

Доведення цих збіжностей при виконанні умов теореми 1 винесено в додаток.

З урахуванням виконання (5)–(7) рівність (4) набирає вигляду

$$\big(I_n + o_{\mathsf{P}}(1)\big)\hat{X} = \big(I_n + o_{\mathsf{P}}(1)\big)X_0 + o_{\mathsf{P}}(1), \qquad m \to \infty,$$

а отже,

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

Теорему доведено.

Для доведення строгої консистентності введемо посилені умови, порівняно з умовами (iii) та (iv):

(v) $\exists r \geqslant 2$: $\mathsf{E}\,|\tilde{a}_{ij}|^{2r} \leqslant \mathrm{const}$; $\mathsf{E}\left|\tilde{b}_{il}\right|^{2r} \leqslant \mathrm{const}$, $i \geqslant 1$, $1 \leqslant j \leqslant n$, $1 \leqslant l \leqslant d$;

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

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

Теорема 2. *Якщо виконуються припущення* (ii), (v), (vi), *то*

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

Доведення вміщено в додатку.

5. Порядок збіжності. Нехай виконуються умови теореми 1. Тоді з імовірністю, що прямує до 1 при $m \to \infty$, матимемо

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

Позначимо $\hat{\Delta} = \hat{X} - X_0$ і розглянемо $m \geqslant m_0$, при яких матриця V_{A_0} буде невиродженою. Тоді

$$\left(A^T A - V_{\tilde{A}}\right) \hat{\Delta} = A^T\left(A_0 X_0 + \tilde{B}\right) - \left(A^T A - V_{\tilde{A}}\right) X_0,$$
$$V_{A_0}^{-1} \left(A^T A - V_{\tilde{A}}\right) \hat{\Delta} = V_{A_0}^{-1} A^T \tilde{B} + V_{A_0}^{-1} \left(A^T A_0 X_0 - \left(A^T A - V_{\tilde{A}}\right) X_0\right) \stackrel{\text{df}}{=} R_1 + R_2. \tag{8}$$

З (5) маємо

$$\left(A^T A - V_{\tilde{A}}\right) \hat{\Delta} = \left(I_n + \mathsf{o_P}(1)\right)\hat{\Delta}, \qquad m \to \infty. \tag{9}$$

У процесі доведення збіжності (7) (див. додаток) отримуємо

$$\mathsf{E}\, \|R_1\|_{\mathsf{F}}^2 \leqslant \text{const} \frac{\lambda_{\max}\left(V_{A_0}\right) + m}{\lambda_{\min}^2\left(V_{A_0}\right)}.$$

Тоді

$$R_1 = \sqrt{\frac{\lambda_{\max}\left(V_{A_0}\right) + m}{\lambda_{\min}^2\left(V_{A_0}\right)}} \mathsf{O_P}(1), \qquad m \to \infty. \tag{10}$$

Запишемо розклад $R_2 = R_{21} + R_{22}$, де

$$R_{21} \stackrel{\text{df}}{=} V_{A_0}^{-1} \left(A^T A_0 X_0 - V_{A_0} X_0\right) = V_{A_0}^{-1} \tilde{A}^T A_0 X_0,$$

$$R_{22} \stackrel{\text{df}}{=} V_{A_0}^{-1} \left((A^T A - V_{\tilde{A}}) X_0 - V_{A_0} X_0\right) =$$

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

Розглянемо R_{21}. Використовуючи доведення теореми 1 (див. додаток), маємо

$$\mathsf{E}\, \|V_{A_0}^{-1} \tilde{A}^T A_0\|_{\mathsf{F}}^2 \leqslant \text{const} \frac{\lambda_{\max}\left(V_{A_0}\right)}{\lambda_{\min}^2\left(V_{A_0}\right)}.$$

Отже,

$$\mathsf{E}\, \|R_{21}\| \leqslant \text{const} \frac{\lambda_{\max}\left(V_{A_0}\right)}{\lambda_{\min}^2\left(V_{A_0}\right)},$$
$$R_{21} = \frac{\sqrt{\lambda_{\max}\left(V_{A_0}\right)}}{\lambda_{\min}\left(V_{A_0}\right)} \mathsf{O_P}(1), \qquad m \to \infty. \tag{11}$$

Далі, так само оцінимо R_{22} :

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right)\right\|_{\mathsf{F}}^2\leqslant \mathrm{const}\frac{m}{\lambda_{\min}^2\left(V_{A_0}\right)},$$

$$\mathsf{E}\left\|V_{A_0}^{-1}\tilde{A}^TA_0\right\|_{\mathsf{F}}^2\leqslant \mathrm{const}\frac{\lambda_{\max}\left(V_{A_0}\right)}{\lambda_{\min}^2\left(V_{A_0}\right)}.$$

Отже,

$$R_{22}=\frac{\sqrt{m}+\sqrt{\lambda_{\max}\left(V_{A_0}\right)}}{\lambda_{\min}\left(V_{A_0}\right)}\mathrm{O}_{\mathsf{P}}(1),\qquad m\to\infty. \tag{12}$$

Остаточно з (8) $-$(12) при виконанні умов теореми 1 отримуємо

$$\hat{X}-X_0=\frac{\sqrt{m}+\sqrt{\lambda_{\max}\left(V_{A_0}\right)}}{\lambda_{\min}\left(V_{A_0}\right)}\mathrm{O}_{\mathsf{P}}(1),\qquad m\to\infty.$$

6. Висновки. У роботі розглянуто множинну векторну модель $AX=B$, для якої побудовано покращену оцінку найменших квадратів (3), що є консистентною у випадку, коли залежні змінні моделі спостерігаються з похибками. Досліджено швидкість збіжності цієї оцінки до істинного параметра X_0.

Подальший інтерес викликає встановлення умов асимптотичної нормальності оцінки (3) та побудова модифікації оцінки для малого та середнього обсягу вибірки. Модифікована оцінка повинна мати таку ж асимптотичну ефективність, але бути більш стійкою з обчислювальної точки зору для реального обсягу вибірки.

7. Додаток. ***7.1. Доведення збіжностей* (5) – (7).** 1. Нехай виконано припущення (ii) – (iv). Доведемо (5). Маємо

$$V_{A_0}^{-1}\left(A^TA-V_{\tilde{A}}\right)=I_n+V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right)+V_{A_0}^{-1}\left(\tilde{A}^TA_0+A_0^T\tilde{A}\right), \tag{13}$$

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right)\right\|_{\mathrm{F}}^2\leqslant\left\|V_{A_0}^{-1}\right\|^2\mathsf{E}\left\|\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right\|_{\mathrm{F}}^2,$$

$$\left\|V_{A_0}^{-1}\right\|=\frac{1}{\lambda_{\mathsf{min}}\left(V_{A_0}\right)}.$$

З умов (ii), (iii) випливає, що

$$\mathsf{E}\left\|\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right\|_{\mathrm{F}}^2=\mathsf{E}\sum_{i,k=1}^{n}\left(\sum_{j=1}^{m}\left(\tilde{a}_{ji}\tilde{a}_{jk}-\mathsf{E}\,\tilde{a}_{ji}\tilde{a}_{jk}\right)\right)^2=$$

$$=\sum_{i,k=1}^{n}\mathsf{var}\left(\sum_{j=1}^{m}\tilde{a}_{ji}\tilde{a}_{jk}\right)=\sum_{i,k=1}^{n}\sum_{j=1}^{m}\mathsf{var}\left(\tilde{a}_{ji}\tilde{a}_{jk}\right).$$

З умови (iii) отримуємо

$$\mathsf{var}\left(\tilde{a}_{ji}\tilde{a}_{jk}\right)=\mathsf{E}\left(\tilde{a}_{ji}\tilde{a}_{jk}\right)^2-\left(\mathsf{E}\,\tilde{a}_{ji}\tilde{a}_{jk}\right)^2\leqslant\mathsf{E}\left(\tilde{a}_{ji}\tilde{a}_{jk}\right)^2\leqslant\sqrt{\mathsf{E}\,\tilde{a}_{ji}^4}\sqrt{\mathsf{E}\,\tilde{a}_{jk}^4}\leqslant\mathrm{const}. \tag{14}$$

Тому

$$\mathsf{E}\left\|\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right\|_{\mathrm{F}}^2 \leqslant \mathrm{const}\cdot m.$$

З припущення (iv) випливає, що

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right)\right\|_{\mathrm{F}}^2 \leqslant \frac{\mathrm{const}\cdot m}{\lambda_{\mathrm{min}}^2\left(V_{A_0}\right)}\to 0, \qquad m\to\infty.$$

Отже,

$$V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A}-V_{\tilde{A}}\right)\xrightarrow{\mathsf{P}} 0, \qquad m\to\infty. \tag{15}$$

Далі, з використанням (ii) та (iii) отримуємо

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^T A_0+A_0^T\tilde{A}\right)\right\|_{\mathrm{F}}^2 \leqslant \left\|V_{A_0}^{-1}\right\|^2\mathsf{E}\left\|\left(\tilde{A}^T A_0+A_0^T\tilde{A}\right)\right\|_{\mathrm{F}}^2 \leqslant$$

$$\leqslant \frac{1}{\lambda_{\mathsf{min}}^2\left(V_{A_0}\right)}\mathsf{E}\left(\left\|\tilde{A}^T A_0\right\|_{\mathrm{F}}+\left\|A_0^T\tilde{A}\right\|_{\mathrm{F}}\right)^2 = \frac{4}{\lambda_{\mathsf{min}}^2\left(V_{A_0}\right)}\mathsf{E}\left\|\tilde{A}^T A_0\right\|_{\mathrm{F}}^2, \tag{16}$$

$$\mathsf{E}\left\|\tilde{A}^T A_0\right\|_{\mathrm{F}}^2 = \sum_{i,k=1}^{n}\mathsf{E}\left(\sum_{j=1}^{m}a_{ij}^0\tilde{a}_{ik}\right)^2 =$$

$$= \sum_{i,k=1}^{n}\mathsf{E}\left(\sum_{j=1}^{m}\left(a_{ji}^0\right)^2\left(\tilde{a}_{jk}\right)^2+\sum_{\substack{j,j'=1\\ j\neq j'}}^{m}a_{ji}^0a_{j'i}^0\tilde{a}_{jk}\tilde{a}_{j'k}\right) =$$

$$= \sum_{i,k=1}^{n}\left(\sum_{j=1}^{m}\left(a_{ji}^0\right)^2\mathsf{E}\left(\tilde{a}_{jk}\right)^2+2\sum_{\substack{j,j'=1\\ j\leqslant j'}}^{m}a_{ji}^0a_{j'i}^0\,\mathsf{E}\left(\tilde{a}_{jk}\tilde{a}_{j'k}\right)\right) =$$

$$= \sum_{i,k=1}^{n}\sum_{j=1}^{m}\left(a_{ji}^0\right)^2\mathsf{E}\left(\tilde{a}_{jk}\right)^2 \leqslant$$

$$\leqslant \mathrm{const}\cdot\|A_0\|_{\mathrm{F}}^2 \leqslant \mathrm{const}\cdot\lambda_{\max}\left(V_{A_0}\right).$$

Тому

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^T A_0+A_0^T\tilde{A}\right)\right\|_{\mathrm{F}}^2 \leqslant \frac{\mathrm{const}\cdot\lambda_{\max}\left(V_{A_0}\right)}{\lambda_{\mathsf{min}}^2\left(V_{A_0}\right)}\to 0, \qquad m\to\infty,$$

що випливає з (iv). Отже,

$$V_{A_0}^{-1}\left(\tilde{A}^T A_0+A_0^T\tilde{A}\right)\xrightarrow{\mathsf{P}} 0, \qquad m\to\infty. \tag{17}$$

З (13), (15) і (17) маємо (5).

Для доведення (6) запишемо

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

З (16) маємо

$$\mathsf{E}\left\|V_{A_0}^{-1}\tilde{A}^T A_0\right\|_{\mathrm{F}}^2 \leqslant \left\|V_{A_0}^{-1}\right\|^2 \mathsf{E}\left\|\tilde{A}^T A_0\right\|_{\mathrm{F}}^2 \leqslant \frac{\mathrm{const}}{\lambda_{\min}^2\left(V_{A_0}\right)}\lambda_{\max}\left(V_{A_0}\right) \xrightarrow[m\to\infty]{} 0.$$

Отже,

$$V_{A_0}^{-1}\tilde{A}^T A_0 \xrightarrow{\mathsf{P}} 0, \qquad m \to \infty.$$

Звідси і отримуємо (6).

Доведемо співвідношення (7):

$$\left\|V_{A_0}^{-1}A^T\tilde{B}\right\|_{\mathrm{F}}^2 \leqslant \frac{1}{\lambda_{\min}^2\left(V_{A_0}\right)}\left\|A^T\tilde{B}\right\|_{\mathrm{F}}^2.$$

З умов (ii) та (iii) маємо

$$\mathsf{E}\left\|A^T\tilde{B}\right\|_{\mathrm{F}}^2 = \sum_{j=1}^{n}\sum_{l=1}^{d}\mathsf{E}\left(\sum_{i=1}^{m}a_{ij}\tilde{b}_{il}\right)^2 = \sum_{j=1}^{n}\sum_{l=1}^{d}\sum_{i=1}^{m}\mathsf{E}\,a_{ij}^2\,\mathsf{E}\,\tilde{b}_{il}^2 \leqslant \sum_{j=1}^{n}\sum_{i=1}^{m}\mathsf{E}\,a_{ij}^2 =$$

$$= \sum_{j,i}\mathsf{E}\left((a_{ij}^0)^2 + 2a_{ij}^0\tilde{a}_{ij} + (\tilde{a}_{ij})^2\right) = \sum_{j,i}(a_{ij}^0)^2 + \sum_{j,i}\mathsf{E}(\tilde{a}_{ij})^2 \leqslant$$

$$\leqslant \|A_0\|_{\mathrm{F}}^2 + \sum_{j=1}^{n}\sum_{i=1}^{m}\left(\mathsf{E}\,\tilde{a}_{ij}^4\right)^{1/2} \leqslant \mathrm{const}\left(\lambda_{\max}\left(V_{A_0}\right) + m\right).$$

Тоді

$$\mathsf{E}\left\|A^T\tilde{B}\right\|_{\mathrm{F}}^2 \leqslant \mathrm{const}\frac{\lambda_{\max}\left(V_{A_0}\right) + m}{\lambda_{\min}^2\left(V_{A_0}\right)} \to 0, \qquad m \to \infty,$$

за припущенням (iv). Це і доводить справедливість (7).

2. Розглянемо випадок, коли виконано припущення (ii), (iii′), (iv). Тоді доведення не змінюється, за винятком того, що у (14) оцінювання відбувається з використанням припущення (iii′):

$$\mathsf{var}\left(\tilde{a}_{ji}\tilde{a}_{jk}\right) \leqslant \mathsf{E}\left(\tilde{a}_{ji}\tilde{a}_{jk}\right)^2 = \mathsf{E}\,\tilde{a}_{ji}^2\mathsf{E}\,\tilde{a}_{jk}^2 \leqslant \mathrm{const}.$$

7.2. Нерівність Розенталя. При доведенні строгої консистентності буде суттєво використовуватися наступне твердження [12, с. 244].

Лема 1 (нерівність Розенталя). *Нехай* $\{\eta_i, i \geqslant 1\}$ *— послідовність незалежних випадкових величин,* $\mathsf{E}\,\eta_i = 0$, $i \geqslant 1$. *Тоді для будь-яких* $t \in \mathbb{R}$, $t \geqslant 2$ *і* $m \geqslant 1$

$$\mathsf{E}\left|\sum_{i=1}^{m}\eta_i\right|^t \leqslant c(t)\max\left(\sum_{i=1}^{m}\mathsf{E}\left|\eta_i\right|^t, \left(\sum_{i=1}^{m}\mathsf{E}\,\eta_i^2\right)^{t/2}\right),$$

де величина $c(t)$ *залежить лише від* t *і не залежить від* m.

7.3. Доведення теореми 2. Можна показати, що збіжність у (5)–(7) за даних умов є збіжністю з імовірністю 1. Тоді з (4) випливатиме доведення теореми.

Оцінимо момент різниці $\tilde{A}^T\tilde{A} - V_{\tilde{A}}$, де r — число з умови (v):

$$\mathsf{E}\left\|\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right\|_{\mathrm{F}}^r = \mathsf{E}\left\|\sum_{i=1}^m\left(\tilde{a}_i\tilde{a}_i^T - \mathsf{E}\,\tilde{a}_i\tilde{a}_i^T\right)\right\|_{\mathrm{F}}^r \leqslant$$

$$\leqslant \mathrm{const}\cdot\max\left(\sum_{i=1}^m \mathsf{E}\left\|\tilde{a}_i\tilde{a}_i^T - \mathsf{E}\,\tilde{a}_i\tilde{a}_i^T\right\|_{\mathrm{F}}^r, \left(\sum_{i=1}^m \mathsf{E}\left\|\tilde{a}_i\tilde{a}_i^T - \mathsf{E}\,\tilde{a}_i\tilde{a}_i^T\right\|_{\mathrm{F}}^2\right)^{r/2}\right),$$

$$\mathsf{E}\left\|\tilde{a}_i\tilde{a}_i^T - \mathsf{E}\,\tilde{a}_i\tilde{a}_i^T\right\|_{\mathrm{F}}^r = \mathsf{E}\left(\sum_{j,k=1}^n\left(\tilde{a}_{ij}\tilde{a}_{ik} - \mathsf{E}\,\tilde{a}_{ij}\tilde{a}_{ik}\right)^2\right)^{r/2} \leqslant$$

$$\leqslant \mathrm{const}\cdot\mathsf{E}\sum_{j,k=1}^n\left(|\tilde{a}_{ij}\tilde{a}_{ik}|^r + r|\tilde{a}_{ij}\tilde{a}_{ik}|^{r-1}\,\mathsf{E}\,|\tilde{a}_{ij}\tilde{a}_{ik}|\right) =$$

$$= \mathrm{const}\cdot\sum_{j,k=1}^n\left(\mathsf{E}\,|\tilde{a}_{ij}\tilde{a}_{ik}|^r + r\,\mathsf{E}\,|\tilde{a}_{ij}\tilde{a}_{ik}|^{r-1}\,\mathsf{E}\,|\tilde{a}_{ij}\tilde{a}_{ik}|\right) \leqslant$$

$$\leqslant \mathrm{const}\cdot\sum_{j,k=1}^n\left(\sqrt{\mathsf{E}\,|\tilde{a}_{ij}|^{2r}}\sqrt{\mathsf{E}\,|\tilde{a}_{ik}|^{2r}} + \right.$$

$$\left. + r\sqrt{\mathsf{E}\,|\tilde{a}_{ij}|^{2r-2}}\sqrt{\mathsf{E}\,|\tilde{a}_{ik}|^{2r-2}}\,\mathsf{E}\sqrt{|\tilde{a}_{ij}|^2|\tilde{a}_{ik}|^2}\right) \leqslant \mathrm{const},$$

$$\mathsf{E}\left\|\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right\|_{\mathrm{F}}^2 \leqslant \left(\mathsf{E}\left\|\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right\|_{\mathrm{F}}^r\right)^{\frac{r}{2}} \leqslant \mathrm{const}.$$

Оскільки $r \geqslant 2$, то з леми 1 маємо

$$\mathsf{E}\left\|\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right\|_{\mathrm{F}}^r \leqslant \mathrm{const}\cdot m^{r/2}.$$

Далі,

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right)\right\|_{\mathrm{F}}^r \leqslant \left\|V_{A_0}^{-1}\right\|^r\mathsf{E}\left\|\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right\|_{\mathrm{F}}^r \leqslant \mathrm{const}\cdot\frac{m^{r/2}}{\lambda_{\min}^r(V_{A_0})}.$$

Тоді з умови (vi) за лемою Бореля–Кантеллі отримуємо

$$V_{A_0}^{-1}\left(\tilde{A}^T\tilde{A} - V_{\tilde{A}}\right) \xrightarrow{\mathsf{P1}} 0, \qquad m\to\infty.$$

Так само можна довести, що

$$V_{A_0}^{-1}\left(\tilde{A}^T A_0 + A_0^T\tilde{A}\right) \xrightarrow{\mathsf{P1}} 0, \qquad m\to\infty.$$

Дійсно,

$$\mathsf{E}\left\|A_0^T\tilde{A}\right\|_{\mathrm{F}}^{2r}=\mathsf{E}\left\|\sum_{i=1}^{m}a_i^0\tilde{a}_i^T\right\|_{\mathrm{F}}^{2r}\leqslant \mathrm{const}\cdot\max\left(\sum_{i=1}^{m}\mathsf{E}\left\|a_i^0\tilde{a}_i^T\right\|_{\mathrm{F}}^{2r},\left(\sum_{i=1}^{m}\left\|a_i^0\tilde{a}_i^T\right\|_{\mathrm{F}}^{2}\right)^r\right),$$

$$\mathsf{E}\left\|a_i^0\tilde{a}_i^T\right\|_{\mathrm{F}}^{2r}\leqslant \mathrm{const}\cdot\|a_i^0\|^{2r}\,\mathsf{E}\,\|\tilde{a}_i\|^{2r}\overset{\text{(v)}}{\leqslant}\mathrm{const}\cdot\|a_i^0\|^{2r},$$

$$\mathsf{E}\left\|A_0^T\tilde{A}\right\|_{\mathrm{F}}^{2r}\leqslant \mathrm{const}\cdot\max\left(\sum_{i=1}^{m}\|a_i^0\|^{2r},\left(\sum_{i=1}^{m}\|a_i^0\|^2\right)^r\right)\leqslant$$

$$\leqslant \mathrm{const}\left(\|A_0\|_{\mathrm{F}}^2\right)^r\leqslant \mathrm{const}\cdot\lambda_{\max}^r\left(V_{A_0}\right).$$

Тому

$$\mathsf{E}\left\|V_{A_0}^{-1}\left(\tilde{A}^TA_0+A_0^T\tilde{A}\right)\right\|_{\mathrm{F}}^{2r}\leqslant\left\|V_{A_0}^{-1}\right\|^{2r}\cdot 2\,\mathsf{E}\left\|A_0^T\tilde{A}\right\|_{\mathrm{F}}^{2r}\leqslant \mathrm{const}\cdot\frac{\lambda_{\max}^r\left(V_{A_0}\right)}{\lambda_{\min}^{2r}\left(V_{A_0}\right)}.$$

Оскільки

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

то за лемою Бореля – Кантеллі

$$V_{A_0}^{-1}\left(\tilde{A}^TA_0+A_0^T\tilde{A}\right)\xrightarrow{\mathsf{P1}}0,\qquad m\to\infty.$$

Отже, збіжність у (5) буде в сенсі збіжності м. н.

Аналогічно матриці в (6) збігаються з імовірністю 1.

Для доведення збіжності з імовірністю 1 у (7) розглянемо

$$\mathsf{E}\left[\left\|A^T\tilde{B}\right\|_{\mathrm{F}}^{2r}\middle|\,\tilde{A}\right]=\mathsf{E}\left[\left\|\sum_{i=1}^{m}a_i\tilde{b}_i^T\right\|_{\mathrm{F}}^{2r}\middle|\,\tilde{A}\right]\leqslant$$

$$\leqslant \mathrm{const}\left(\sum_{i=1}^{m}\mathsf{E}\left[\left\|a_i\tilde{b}_i^T\right\|_{\mathrm{F}}^{2r}\middle|\,\tilde{A}\right]+\left(\sum_{i=1}^{m}\mathsf{E}\left[\left\|a_i\tilde{b}_i^T\right\|_{\mathrm{F}}^{2}\middle|\,\tilde{A}\right]\right)^r\right).$$

З використанням припущення (v) маємо

$$\mathsf{E}\left[\left\|a_i\tilde{b}_i^T\right\|_{\mathrm{F}}^{2r}\middle|\,\tilde{A}\right]\leqslant\mathsf{E}\left[\|a_i\|^{2r}\tilde{b}_i\|^{2r}\middle|\,\tilde{A}\right]\leqslant \mathrm{const}\|a_i\|^{2r}\,\mathsf{E}\left\|\tilde{b}_i^T\right\|^{2r}\leqslant\mathrm{const}\|a_i\|^{2r}.$$

Далі,

$$\mathsf{E}\left(\sum_{i=1}^{m}\mathsf{E}\left[\left\|a_i\tilde{b}_i^T\right\|_{\mathrm{F}}^{2r}\middle|\,\tilde{A}\right]\right)\leqslant \mathrm{const}\sum_{i=1}^{m}\mathsf{E}\,\|a_i\|^{2r}\leqslant \mathrm{const}\sum_{i=1}^{m}\left(\|a_i^0\|^{2r}+\mathsf{E}\,\|\tilde{a}_i\|^{2r}\right)\leqslant$$

$$\leqslant \mathrm{const}\left(\|A_0\|_{\mathrm{F}}^{2r}+m\right)=\mathrm{const}\left(\lambda_{\max}^r\left(V_{A_0}\right)+m\right),$$

$$\mathsf{E}\left(\sum_{i=1}^{m}\mathsf{E}\left[\left\|a_i\tilde{b}_i^T\right\|_{\mathrm{F}}^2\middle|\tilde{A}\right]\right)^r \leqslant \mathrm{const}\left(\sum_{i=1}^{m}\mathsf{E}\,\|a_i\|^2\right)^r \leqslant \mathrm{const}\sum_{i=1}^{m}\left(\|a_i^0\|^2+\mathsf{E}\,\|\tilde{a}_i\|^2\right)^r \leqslant$$

$$\leqslant \mathrm{const}\left(\lambda_{\max}\left(V_{A_0}\right)+m\right)^r \leqslant \mathrm{const}\Big(\lambda_{\max}^r\left(V_{A_0}\right)+m^r\Big).$$

Тому

$$\mathsf{E}\left\|A^T\tilde{B}\right\|_{\mathrm{F}}^{2r} \leqslant \mathrm{const}\cdot\frac{\lambda_{\max}^r\left(V_{A_0}\right)+m^r}{\lambda_{\min}^{2r}\left(V_{A_0}\right)}.$$

Тоді, використовуючи припущення (vi) та лему Бореля – Кантеллі, маємо

$$V_{A_0}^{-1}A^T\tilde{B}\xrightarrow{\mathsf{P1}}0,\qquad m\to\infty.$$

Зрештою, матриці у (5) – (7) збігаються з імовірністю 1, і теорему доведено.

1. *Fuller W. A.* Measurement error models. – New York: Wiley, 1987. – 440 p.
2. *Cheng C.-L., Van Ness J.* Structural and functional models revisited // Recent Advances in Total Least Squares Techniques and Error-in-Variables Modeling: SIAM Proceedings Series (Leuven, Belgium, August 21 – 24, 1996). – Philadelphia, 1997. – P. 37 – 50.
3. *Van Huffel S., Vandewalle J.* The total least squares problem: Computation aspects and analysis. – Philadelphia: SIAM, 1991. – 300 p.
4. *Kukush A., Van Huffel S.* Consistency of elementwise-weighted total least squares estimator in a multivariate errors-in-variables model $AX = B$ // Metrika. – 2004. – **59**. – P. 75 – 97.
5. *Kukush A., Otto Mashke E.* The efficiency of adjusted least squares in the linear functional relationship // J. Multivar. Anal. – 2003. – **87**. – P. 261 – 274.
6. *Cheng C.-L., Schneeweiss H.* Polynomial regression with errors in the variables // J. R. Statist. Soc. B. – 2000. – **60**. – P. 189 – 199.
7. *Kukush A., Markovsky I., Van Huffel S.* Consistent estimator in the bilinear multivariate errors-in-variables model // Metrika. – 2003. – **57**. – P. 253 – 285.
8. *Кукуш О. Г., Полеха М. Я.* Конзистентна оцінка у векторній моделі з похибками у змінних при невідомій коваріаційній структурі похибок // Укр. мат. журн. – 2007. – **59**, № 8. – P. 1026 – 1033.
9. *Гантмахер Ф. Р.* Теория матриц. – М.: Наука, 1988. – 576 с.
10. *Carroll R. J., Ruppert D., Sefanski L. A.* Measurement error in nonlinear models: number 63 in monographs on statistics and applied probability. – London: Chapman & Hall, 1995. – 305 p.
11. *Gallo P. P.* Consistency of regression estimates when some variables are subject to error // Communs Statist. B-Theory Methods. – 1982. – **11**. – P. 973 – 893.
12. *Härdle W., Kerkyacharian G., Picard D., Tsybakov A.* Wavelets approximation and statistical applications. – New York: Springer-Verlag, 1998. – 265 p.

Одержано 04.04.12