Теорія ймовірностей
та математична статистика
Вип. 97, 2017, с. 48–57

Teoriya Ĭmovirnosteĭ
ta Matematychna Statystyka
No. 97, 2017, pp. 48–57

КОНСИСТЕНТНІСТЬ ОЦІНКИ ОРТОГОНАЛЬНОЇ РЕГРЕСІЇ В НЕЯВНІЙ ЛІНІЙНІЙ МОДЕЛІ З ПОХИБКАМИ У ЗМІННИХ

О. О. ДАШКОВ, О. Г. КУКУШ

Анотація. Розглядається неявна лінійна модель регресії з похибками у змінних, де істинні точки лежать на деякій гіперплощині в евклідовому просторі, а сукупна кореляційна матриця похибок є пропорційною до одиничної матриці. Вивчається оцінка ортогональної регресії цієї гіперплощини. Наведено достатні умови консистентності та строгої консистентності оцінки. Результати застосовуються до явної множинної моделі регресії з вільним членом і похибками у змінних.

Ключові слова і фрази. Консистентність оцінки, множинна модель регресії з похибками у змінних, неявна лінійна модель регресії, оцінка ортогональної регресії, оцінка повних найменших квадратів.

2010 *Mathematics Subject Classification.* Primary 62J05; Secondary 62H12, 65F20.

1. Вступ

Функціональна множинна лінійна модель регресії з вільним членом і похибками у змінних задається рівняннями:

$$\begin{cases} y_i = b_0 + x^\top \xi_i + \varepsilon_i, \\ x_i = \xi_i + \delta_i. \end{cases} \tag{1}$$

Тут $\xi_i \in \mathbb{R}^{m-1}$ — це невідомі невипадкові вектори, ε_i, δ_i — випадкові похибки, а $b_0 \in \mathbb{R}$, $x \in \mathbb{R}^{m-1}$ — це параметри регресії, які потрібно оцінити за спостереженнями (y_i, x_i), $i = \overline{1, n}$. Подібні моделі спостережень використовують в економетриці та при обробці сигналів. Тут і надалі всі вектори є векторами-стовпцями.

Зазвичай цю модель зводять до лінійної регресії без вільного члена, в якій уже розглядають оцінку повних найменших квадратів (Total Least Squares estimator, TLS-оцінка). Вона досліджувалась, зокрема, у роботах [1–3, 5], де наведено достатні умови слабкої і строгої консистентності TLS-оцінки при різних припущеннях про модель спостережень. Найбільш слабкі умови для випадку, який ми будемо розглядати, містяться у статті [7].

Для того, щоб звільнитися від вільного члена в моделі (1), вводять додатковий регресор, який тотожно рівний одиниці та спостерігається без похибок. Але такий підхід призводить до того, що регресори стають нерівноправними. Щоб запобігти цьому, ми розглядаємо одразу неявну лінійну модель спостережень в евклідовому просторі $\mathbb{R}^m$, $m \geq 2$, де приховані змінні лежать на деякій гіперплощині, яку треба оцінити.

У такій моделі найбільш поширеною є оцінка ортогональної регресії (ООР), яка задає гіперплощину з найменшою сумою квадратів відстаней від спостережуваних точок. Рівноправність усіх змінних дозволяє застосувати геометричні міркування при дослідженні властивостей оцінки.

Використовуючи переваги неявної моделі, ми знаходимо достатні умови консистентності та строгої консистентності ООР. Ці умови накладаються на власні значення вибіркової коваріаційної матриці прихованих змінних. Вони не є наслідком із відповідних умов консистнтності з роботи [7], якщо останні застосувати до зведеної моделі без вільного члена.

Стаття побудована таким чином. У розділі 2 вводиться неявна лінійна модель спостережень в евклідовому просторі, задається ООР і формулюються припущення про похибки вимірювання. Розділи 3 та 4 містять достатні умови консистентності та строгої консистентності цієї оцінки. У розділі 5 ці результати застосовуються до TLS-оцінки у множинній лінійній регресії з вільним членом та похибками у змінних. У розділі 6 множинна модель із вільним членом зводиться до моделі без вільного члена і проводиться порівняння отриманих результатів із відомими теоремами про консистентність TLS-оцінки. Останній розділ роботи окреслює напрямок подальших досліджень.

У цій статті риска згори позначає усереднення перших n членів послідовності чисел, векторів чи матриць. Число в дужках згори позначає номер члена усередненої послідовності. Якщо він відсутній, то за замовчуванням вважається, що це n-й член послідовності.

2. Оцінка і припущення про модель

Розглянемо неявну лінійну модель спостережень в евклідовому просторі $\mathbb{R}^m$, $m \geq 2$:

$$\begin{cases} z_i = \eta_i + \gamma_i, \\ (\eta_i, \tau) = d. \end{cases} \tag{2}$$

Тут z_i — спостережувані вектори в $\mathbb{R}^m$, η_i — приховані змінні, які лежать на гіперплощині

$$\Gamma_{\tau d} = \{u \in \mathbb{R}^m : (u, \tau) = d\}, \tag{3}$$

τ — одиничний вектор нормалі до гіперплощини, $d \in \mathbb{R}$; γ_i — випадкові похибки вимірювання. За спостереженнями z_i, $i = \overline{1, n}$, оцінюють гіперплощину (3).

Надалі глобальними припущеннями про модель (2) будуть такі умови.

(i) Вектори γ_i, $i \geq 1$, є незалежними в сукупності з нульовим середнім.

(ii) Вектори γ_i, $i \geq 1$, мають однакову коваріаційну матрицю $\mathbf{cov}(\gamma_i) = \sigma^2 I_m$, де $\sigma^2 > 0$ невідоме.

Для довільних наборів u_i, $i = \overline{1, n}$, та v_i, $i = \overline{1, n}$, векторів з $\mathbb{R}^m$ введемо вибіркову коваріаційну матрицю

$$S_{uv} = \frac{1}{n} \sum_{i=1}^{n} (u_i - \bar{u})(v_i - \bar{v})^\top, \quad \bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i, \quad \bar{v} = \frac{1}{n} \sum_{i=1}^{n} v_i.$$

У моделі (2) ООР параметрів τ, d задається парою $\hat{\tau}$, $\hat{d}$, що доставляє мінімум цільовій функції

$$Q(\tau, d) = \sum_{i=1}^{n} \rho^2(z_i, \Gamma_{\tau d}), \quad \|\tau\| = 1, \quad d \in \mathbb{R},$$

де $\rho(z, \Gamma)$ позначає відстань від точки z до гіперплощини Γ.

У [6] показано, що ООР $\hat{\tau}$ є нормованим власним вектором матриці S_{zz} , який відповідає найменшому власному значенню $\lambda_{\min}(S_{zz})$, а оцінка

$$\hat{d} = (\bar{z}, \hat{\tau}).$$

Для симетричної матриці $S_{\eta\eta}$ ми впорядковуємо всі власні значення з урахуванням їх кратності:

$$\lambda_{\min}(S_{\eta\eta}) \leq \lambda_2(S_{\eta\eta}) \leq \cdots \leq \lambda_{\max}(S_{\eta\eta}).$$

Надалі Σ позначає $\sigma^2 I_m = \mathbf{cov}(\gamma_i)$.

3. Достатні умови консистентності ООР

Теорема 1. *Нехай у моделі (2)–(3) виконані умови (i), (ii), а також наступні умови: для деякого* $1 \le r \le 2$

$$\sup_{i\ge 1} \mathsf{E}\,\|\gamma_i\|^{2r} < \infty,$$

$$\frac{\lambda_{\max}\Big(S^{(n)}_{\eta\eta}\Big)}{n\lambda_2^2\Big(S^{(n)}_{\eta\eta}\Big)} \to 0, \quad n\to\infty,$$

$$n^{1-1/r}\lambda_2\Big(S^{(n)}_{\eta\eta}\Big) \to \infty, \quad n\to\infty.$$

Тоді оцінка $\hat\tau$ *є консистентною, тобто при* $n\to\infty$

$$\min\big\{\|\hat\tau-\tau\|, \|\hat\tau+\tau\|\big\} \xrightarrow{\mathsf{P}} 0.$$

Доведення. Оскільки ООР еквіваріантна відносно ортогональних перетворень (цю властивість оцінки описано в [6]), то без обмеження загальності можна вважати, що $\tau = (1,0,\dots,0)^\top$. У такому випадку

$$S^{(n)}_{\eta\eta} = \begin{pmatrix} 0 & 0_{1\times(m-1)} \\ 0_{(m-1)\times 1} & A_n \end{pmatrix},$$

де A_n — квадратна матриця розміру $m-1$, бо вектори $\eta_i - \bar\eta^{(n)}$ лежать на гіперплощині, перпендикулярній до вектора τ, і при цьому

$$\lambda_{\min}(A_n) = \lambda_2\Big(S^{(n)}_{\eta\eta}\Big).$$

Тоді

$$S_{zz}\hat\tau = \lambda_{\min}(S_{zz})\hat\tau, \quad S_{zz} = S_{\eta\eta} + S_{\eta\gamma} + S_{\gamma\eta} + S_{\gamma\gamma}. \tag{4}$$

Оскільки S_{zz} симетрична, то

$$\lambda_{\min}(S_{zz}) \le (S_{zz}\tau,\tau) = (S_{\gamma\gamma}\tau,\tau) = \sigma^2 + \big((S_{\gamma\gamma}-\Sigma)\tau,\tau\big). \tag{5}$$

З іншого боку

$$\lambda_{\min}(S_{zz}) = (S_{zz}\hat\tau,\hat\tau) = (S_{\eta\eta}\hat\tau,\hat\tau) + (S_{\eta\gamma}\hat\tau,\hat\tau) + (S_{\gamma\eta}\hat\tau,\hat\tau) + (S_{\gamma\gamma}\hat\tau,\hat\tau).$$

Оскільки $S_{\eta\eta}$ невід'ємно визначена та $\|\hat\tau\| = 1$, то

$$\lambda_{\min}(S_{zz}) - \sigma^2 \ge -(\|S_{\eta\gamma}\| + \|S_{\gamma\eta}\| + \|S_{\gamma\gamma}-\Sigma\|). \tag{6}$$

Із (5) і (6) випливає, що

$$\big|\lambda_{\min}(S_{zz}) - \sigma^2\big| \le \|S_{\eta\gamma}\| + \|S_{\gamma\eta}\| + \|S_{\gamma\gamma}-\Sigma\|. \tag{7}$$

Перепишемо першу рівність (4) у вигляді

$$S_{\eta\eta}\hat\tau = \Big(\big(\lambda_{\min}(S_{zz})-\sigma^2\big)I_m - S_{\eta\gamma} - S_{\gamma\eta} - (S_{\gamma\gamma}-\Sigma)\Big)\hat\tau. \tag{8}$$

Позначимо $B_n = \begin{pmatrix} 0 & 0_{1\times(m-1)} \\ 0_{(m-1)\times 1} & A_n^{-1} \end{pmatrix}$ і нехай P_2 — матриця ортопроектора на гіперплощину $\{\tau\}^\perp$. Тоді

$$S^{(n)}_{\eta\eta}B_n = B_n S^{(n)}_{\eta\eta} = \begin{pmatrix} 0 & 0 \\ 0 & I_{m-1} \end{pmatrix} = P_2.$$

Помноживши обидві частини рівності (8) зліва на B_n, отримаємо

$$P_2\hat\tau = B_n\Big(\big(\lambda_{\min}(S_{zz})-\sigma^2\big)I_m - S_{\eta\gamma} - S_{\gamma\eta} - \big(S_{\gamma\gamma}-\sigma^2 I_m\big)\Big)\hat\tau,$$

а тому, враховуючи (7) і те, що операторна норма матриці не перевищує норми Фробеніуса, маємо

$$\|P_2\hat{\tau}\| \le 2\|B_n\| \cdot \big(\|S_{\eta\gamma}\|_F + \|S_{\gamma\eta}\|_F + \|S_{\gamma\gamma} - \sigma^2 I_m\|_F\big). \tag{9}$$

Тут і надалі $\|A\|_F$ позначає норму Фробеніуса матриці, тобто корінь із суми квадратів усіх її матричних елементів; операторна норма матриці — це норма відповідного оператора в евклідовому просторі.

Оскільки $\|\hat{\tau}\| = 1$, то $\|P_2\hat{\tau}\|^2 + (\hat{\tau}, \tau)^2 = 1$ і тому виконується

$$\min\big(\|\hat{\tau} - \tau\|, \|\hat{\tau} + \tau\|\big) = \sqrt{\left(1 - \sqrt{1 - \|P_2\hat{\tau}\|^2}\right)^2 + \|P_2\hat{\tau}\|^2}\,. \tag{10}$$

Тоді нам достатньо доводити збіжність $P_2\hat{\tau} \xrightarrow{\mathsf{P}} 0, n \to \infty$, а враховуючи нерівність (9), достатньо показати збіжності $\|B_n\| \cdot \left\|S_{\gamma\eta}^{(n)}\right\|_F \xrightarrow{\mathsf{P}} 0$ та

$$\|B_n\| \cdot \left\|S_{\gamma\gamma}^{(n)} - \Sigma\right\|_F \xrightarrow{\mathsf{P}} 0, \quad n \to \infty.$$

Оскільки B_n — блочно-діагональна матриця, а матриця A_n^{-1} симетрична, то

$$\|B_n\| = \max\big(0, \|A_n^{-1}\|\big) = \lambda_{\max}\big(A_n^{-1}\big) = \frac{1}{\lambda_2\left(S_{\eta\eta}^{(n)}\right)}. \tag{11}$$

Для оцінювання $\|S_{\gamma\eta}\|_F$ скористаємось нерівністю Розенталя для $2r \ge 2$ (у формулюванні з роботи [7]):

$$\begin{aligned}\mathsf{E}\left\|S_{\gamma\eta}^{(n)}\right\|_F^{2r} &= \frac{1}{n^{2r}}\,\mathsf{E}\left\|\sum_{i=1}^n \gamma_i(\eta_i - \bar{\eta})^\top\right\|_F^{2r} \le \\ &\le \frac{\alpha}{n^{2r}}\sum_{i=1}^n \mathsf{E}\|\gamma_i\|^{2r}\|\eta_i - \bar{\eta}\|^{2r} + \frac{\beta}{n^{2r}}\left(\sum_{i=1}^n \mathsf{E}\|\gamma_i\|^2\|\eta_i - \bar{\eta}\|^2\right)^r.\end{aligned} \tag{12}$$

Далі,

$$\mathsf{E}\|\gamma_i\|^{2r} \le \sup_{j\ge1}\mathsf{E}\|\gamma_j\|^{2r}, \quad \left(\mathsf{E}\|\gamma_i\|^2\right)^r \le \sup_{j\ge1}\mathsf{E}\|\gamma_j\|^{2r}, \quad i \ge 1,$$

$$\sum_{i=1}^n \|\eta_i - \bar{\eta}\|^2 = \operatorname{tr}\left(\sum_{i=1}^n (\eta_i - \bar{\eta})(\eta_i - \bar{\eta})^\top\right) = n \cdot \operatorname{tr}\left(S_{\eta\eta}^{(n)}\right) \le nm\lambda_{\max}\left(S_{\eta\eta}^{(n)}\right),$$

$$\sum_{i=1}^n \|\eta_i - \bar{\eta}\|^{2r} \le \left(\sum_{i=1}^n \|\eta_i - \bar{\eta}\|^2\right)^r \le n^r m^r \lambda_{\max}\left(S_{\eta\eta}^{(n)}\right)^r.$$

Тому

$$\mathsf{E}\|S_{\gamma\eta}\|_F^{2r} = \frac{O\left(\lambda_{\max}\left(S_{\eta\eta}^{(n)}\right)^r\right)}{n^r}, \quad n \to \infty. \tag{13}$$

З (11), (13) і умов теореми випливає, що

$$\|B_n\| \cdot \left\|S_{\gamma\eta}^{(n)}\right\|_F \xrightarrow{\mathsf{P}} 0, \quad n \to \infty. \tag{14}$$

Тепер оцінимо

$$\mathsf{E}\|S_{\gamma\gamma} - \Sigma\|_F^r \le \left(\left(\mathsf{E}\left\|\overline{\gamma\gamma^\top} - \Sigma\right\|_F^r\right)^{1/r} + \left(\mathsf{E}\left\|\bar{\gamma}\,\overline{\gamma^\top}\right\|_F^r\right)^{1/r}\right)^r. \tag{15}$$

За умов теореми 1 послідовність $\left\{\mathsf{E}\left\|\gamma_i\gamma_i^\top - \Sigma\right\|^r, i \ge 1\right\}$ обмежена і випадкові матриці $\left\{\gamma_i\gamma_i^\top - \Sigma, i \ge 1\right\}$ незалежні у сукупності, тому можемо скористатися нерівністю

Розенталя для $1 \le r \le 2$ (у формулюванні з роботи [7]):

$$\mathsf{E}\Big\|\overline{\gamma\gamma^\top} - \Sigma\Big\|_F^r = \frac{1}{n^r}\,\mathsf{E}\Big\|\sum_{i=1}^n \bigl(\gamma_i\gamma_i^\top - \Sigma\bigr)\Big\|_F^r \le$$
$$\le \frac{\alpha}{n^r}\sum_{i=1}^n \mathsf{E}\bigl\|\gamma_i\gamma_i^\top - \Sigma\bigr\|_F^r = O\bigl(n^{1-r}\bigr). \tag{16}$$

Також маємо

$$\mathsf{E}\Big\|\bar{\gamma}\,\overline{\gamma^\top}\Big\|_F^r = \mathsf{E}\Big\|\bar{\gamma}^{(n)}\Big\|^{2r} = \frac{1}{n^{2r}}\,\mathsf{E}\Big\|\sum_{i=1}^n \gamma_i\Big\|^{2r} \le$$
$$\le \frac{\alpha}{n^{2r}}\sum_{i=1}^n \mathsf{E}\|\gamma_i\|^{2r} + \frac{\beta}{n^{2r}}\Bigl(\sum \mathsf{E}\|\gamma_i\|^2\Bigr)^r = O\bigl(n^{-r}\bigr). \tag{17}$$

Отже, з (15) $\|B_n\| \cdot \Big\|S_{\gamma\gamma}^{(n)} - \Sigma\Big\| \xrightarrow{\mathsf{P}} 0$, $n \to \infty$, що і доводить теорему. □

Теорема 2. *Нехай в умовах теореми 1 послідовність $\bar{\eta}^{(n)}$, $n \ge 1$, обмежена. Тоді ООР $\bigl(\hat{\tau}; \hat{d}\bigr)$ є консистентною оцінкою, тобто при $n \to \infty$*

$$\min\Bigl\{\|\hat{\tau} - \tau\| + \bigl|\hat{d} - d\bigr|,\ \|\hat{\tau} + \tau\| + \bigl|\hat{d} + d\bigr|\Bigr\} \xrightarrow{\mathsf{P}} 0. \tag{18}$$

Зауваження. Саме така збіжність у (18) пов'язана з тим, що пара $\tau = \tau_0$, $d = d_0$ задає ту саму гіперплощину, що й пара $\tau = -\tau_0$, $d = -d_0$.

Доведення. За властивостями ООР $\hat{d} = (\hat{\tau}; \bar{z})$, при цьому $d = (\tau; \bar{\eta})$, тому

$$\|\hat{\tau} - \tau\| + |\hat{d} - d| = \|\hat{\tau} - \tau\| + |(\hat{\tau}, \bar{z}) - (\tau, \bar{\eta})| = \|\hat{\tau} - \tau\| + |(\hat{\tau}, \bar{\gamma}) - (\hat{\tau} - \tau, \bar{\eta})| \le$$
$$\le \|\hat{\tau} - \tau\|(\|\bar{\eta}\| + 1) + \|\bar{\gamma}\| \cdot \|\hat{\tau}\|.$$

Аналогічно отримуємо, що

$$\|\hat{\tau} + \tau\| + \bigl|\hat{d} + d\bigr| \le \|\hat{\tau} + \tau\|(\|\bar{\eta}\| + 1) + \|\bar{\gamma}\| \cdot \|\hat{\tau}\|.$$

А тому

$$\min\Bigl\{\|\hat{\tau} - \tau\| + \bigl|\hat{d} - d\bigr|,\ \|\hat{\tau} + \tau\| + \bigl|\hat{d} + d\bigr|\Bigr\} \le$$
$$\le \min\{\|\hat{\tau} + \tau\|, \|\hat{\tau} - \tau\|\}(\|\bar{\eta}\| + 1) + \|\bar{\gamma}\| \cdot \|\hat{\tau}\|. \tag{19}$$

Оскільки зараз виконані умови теореми 1, то

$$\min\{\|\hat{\tau} + \tau\|, \|\hat{\tau} - \tau\|\} \xrightarrow{\mathsf{P}} 0, \quad n \to \infty.$$

А враховуючи, що $\|\hat{\tau}\| = 1$, $\bigl\|\bar{\eta}^{(n)}\bigr\| = O(1)$, $n \to \infty$, і за законом великих чисел $\bar{\gamma}^{(n)} \xrightarrow{\mathsf{P}} 0$, $n \to \infty$, то з (19) випливає твердження теореми. □

Зауваження. Додаткова умова обмеженості послідовності $\bar{\eta}^{(n)}$, $n \ge 1$, виникає з того, що навіть коли ми достатньо точно оцінили вектор нормалі до шуканої гіперплощини, то не завжди можемо точно оцінити відстань до цієї гіперплощини.

4. Достатні умови строгої консистентності ООР

Теорема 3. *Нехай справджуються умови (i), (ii) та для деяких $r \geq 2$, $n_0 \geq 1$ виконується таке:*

$$\sup_{i \geq 1} \mathsf{E} \|\gamma_i\|^{2r} < \infty,$$

$$\sum_{n=n_0}^{\infty} \left(\frac{\lambda_{\max}\left(S_{\eta\eta}^{(n)}\right)}{n\lambda_2^2\left(S_{\eta\eta}^{(n)}\right)} \right)^r < \infty,$$

$$\sum_{n=n_0}^{\infty} \left(\frac{1}{\sqrt{n}\lambda_2\left(S_{\eta\eta}^{(n)}\right)} \right)^r < \infty.$$

Тоді оцінка $\hat{\tau}$ є строго консистентною, тобто при $n \to \infty$

$$\min\{\|\hat{\tau} - \tau\|, \|\hat{\tau} + \tau\|\} \to 0 \quad \textit{м. н.}$$

Доведення. Із (11) і (13) випливає, що

$$\mathsf{E} \sum_{n=n_0}^{\infty} \|B_n\|^{2r} \cdot \left\|S_{\gamma\eta}^{(n)}\right\|_F^{2r} = \sum_{n=n_0}^{\infty} \mathsf{E}\left\|S_{\gamma\eta}^{(n)}\right\|_F^{2r} \cdot \|B_n\|^{2r} < \infty,$$

і тому $\|B_n\| \cdot \left\|S_{\gamma\eta}^{(n)}\right\|_F \to 0$, $n \to \infty$ м. н. А отже, враховуючи (9) і (10), нам залишається довести збіжність $\|B_n\| \cdot \left\|S_{\gamma\gamma}^{(n)} - \Sigma\right\| \to 0$, $n \to \infty$, м. н. Для цього скористаємося нерівністю Розенталя для $r \geq 2$:

$$\mathsf{E}\left\|\overline{\gamma\gamma^\top}^{(n)} - \Sigma\right\|_F^r = \frac{1}{n^r} \mathsf{E}\left\|\sum_{i=1}^{n} \left(\gamma_i\gamma_i^\top - \Sigma\right)\right\|_F^r \leq$$

$$\leq \frac{\alpha}{n^r} \sum_{i=1}^{n} \mathsf{E}\left\|\gamma_i\gamma_i^\top - \Sigma\right\|_F^r + \frac{\beta}{n^r}\left(\sum \mathsf{E}\left\|\gamma_i\gamma_i^\top - \Sigma\right\|_F^2\right)^{r/2} = O\left(n^{-r/2}\right).$$

Далі, врахувавши (15) і (17), за умов теореми 3 маємо, що

$$\mathsf{E} \sum_{n=n_0}^{\infty} \|B_n\|^r \|S_{\gamma\gamma} - \Sigma\|_F^r = \sum_{n=n_0}^{\infty} \mathsf{E}\|S_{\gamma\gamma} - \Sigma\|_F^r \|B_n\|^r < \infty,$$

а тому м. н. $\|B_n\| \cdot \|S_{\gamma\gamma} - \Sigma\|_F \to 0$, $n \to \infty$, що і доводить теорему. □

Наступна теорема доводиться аналогічно до теореми 2.

Теорема 4. *Нехай в умовах теореми 3 послідовність $\bar{\eta}^{(n)}$, $n \geq 1$, обмежена. Тоді ООР $\left(\hat{\tau}; \hat{d}\right)$ є строго консистентною, тобто при $n \to \infty$*

$$\min\left\{\|\hat{\tau} - \tau\| + \left|\hat{d} - d\right|, \|\hat{\tau} + \tau\| + \left|\hat{d} + d\right|\right\} \to 0 \quad \textit{м. н.}$$

5. Застосування до множинної лінійної моделі

Результати попередніх розділів можна використати при дослідженні моделі (1). Щоб перейти від неї до неявної моделі (2), покладаємо

$$z_i = \begin{pmatrix} y_i \\ x_i \end{pmatrix}, \quad \eta_i = \begin{pmatrix} b_0 + x^\top \xi_i \\ \xi_i \end{pmatrix}, \quad \gamma_i = \begin{pmatrix} \varepsilon_i \\ \delta_i \end{pmatrix}, \quad i = \overline{1, n}, \tag{20}$$

$$\tau = \left(\frac{1}{\sqrt{1 + \|x\|^2}}; \frac{-x^\top}{\sqrt{1 + \|x\|^2}} \right)^\top, \quad d = \frac{b_0}{\sqrt{1 + \|x\|^2}}\,. \tag{21}$$

Застосуємо теореми 1–4 до переписаної моделі (2) для отримання наслідків у моделі (1). Для цього спочатку накладемо відповідні умови: у моделі (1) будуть діяти припущення *(i)*, *(ii)*, де γ_i з (20).

Тепер за оцінками $\hat{\tau}$, $\hat{d}$ побудуємо оцінки $\hat{x}$, $\hat{b}$ параметрів регресії x, b_0. Нехай $\hat{\tau}_1$ позначає першу координату вектора $\hat{\tau}$, а $P\hat{\tau}$ — частину вектора $\hat{\tau}$ без першої координати. Відповідні позначення введемо і для вектора τ. При $\hat{\tau}_1 \neq 0$ покладемо

$$\hat{x} = \frac{-1}{\hat{\tau}_1} P\hat{\tau}, \quad \hat{b} = \frac{\hat{d}}{\hat{\tau}_1}, \tag{22}$$

інакше обидві оцінки покладемо рівними нулю.

Тепер сформулюємо теореми-наслідки.

Теорема 5. *Нехай у моделі (1) виконані умови (i), (ii) та для деякого* $1 \le r \le 2$ *справджується таке:*

$$\sup_{i\ge 1} \mathsf{E}\,\|\gamma_i\|^{2r} < \infty, \tag{23}$$

$$\frac{\lambda_{\max}(S_{\xi\xi})}{n\lambda^2_{\min}(S_{\xi\xi})} \to 0, \quad n \to \infty, \tag{24}$$

$$n^{1-1/r}\lambda_{\min}(S_{\xi\xi}) \to \infty, \quad n \to \infty. \tag{25}$$

Тоді $\hat{x} \xrightarrow{\mathsf{P}} x$, $n \to \infty$.

Теорема 6. *Нехай в умовах теореми 5 послідовність* $\bar{\xi}^{(n)}$, $n \ge 1$, *обмежена. Тоді* $\hat{b} \xrightarrow{\mathsf{P}} b_0$, $n \to \infty$.

Теорема 7. *Нехай в моделі (1) виконані умови (i), (ii) та для деяких* $r \ge 2$, $n_0 \ge 1$ *справджується таке:*

$$\sup_{i\ge 1} \mathsf{E}\,\|\gamma_i\|^{2r} < \infty,$$

$$\sum_{n=n_0}^{\infty} \left(\frac{\lambda_{\max}(S_{\xi\xi})}{n\lambda^2_{\min}(S_{\xi\xi})}\right)^r < \infty, \tag{26}$$

$$\sum_{n=n_0}^{\infty} \left(\frac{1}{\sqrt{n}\lambda_{\min}(S_{\xi\xi})}\right)^r < \infty. \tag{27}$$

Тоді $\hat{x} \to x$ *м. н.*, $n \to \infty$.

Теорема 8. *Нехай в умовах теореми 7 послідовність* $\bar{\xi}^{(n)}$, $n \ge 1$, *обмежена. Тоді* $\hat{b} \to b_0$ *м. н.*, $n \to \infty$.

Для пояснення правильності теорем 5–8 спочатку дослідимо умови, які накладаються на власні числа матриці $S_{\xi\xi}$. Оскільки

$$S_{\eta\eta} = \begin{pmatrix} x^\top S_{\xi\xi}x & x^\top S_{\xi\xi} \\ S_{\xi\xi}x & S_{\xi\xi} \end{pmatrix} = \begin{pmatrix} 1 & x^\top \\ 0 & I_{m-1} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & S_{\xi\xi} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x & I_{m-1} \end{pmatrix},$$

то з теореми Фішера – Куранта про мінімаксну характеризацію власних чисел симетричної матриці [4] отримуємо, що $\lambda_{\min}(S_{\xi\xi}) \le \mathrm{const}\cdot\lambda_2(S_{\eta\eta})$. Також

$$\lambda_{\max}(S_{\eta\eta}) \le \mathrm{tr}(S_{\eta\eta}) \le \left(1 + \|x\|^2\right)\mathrm{tr}(S_{\xi\xi}) \le m\left(1 + \|x\|^2\right)\lambda_{\max}(S_{\xi\xi}).$$

А отже, коли виконані умови однієї з теорем 5–8, то виконуються й умови відповідної теореми 1–4 для переписаної неявної моделі (2).

З огляду на вигляд оцінок у (22) бачимо, що при зміні оцінок $\hat{\tau}$ та $\hat{d}$ на протилежні, вирази для $\hat{x}$ та $\hat{b}$ не змінюються. Тому, враховуючи, що

$$x = \frac{-P\tau}{\tau_1}, \quad b_0 = \frac{d}{\tau_1}, \quad \tau_1 \neq 0,$$

зі збіжності $\min\{\|\hat{\tau}-\tau\|, \|\hat{\tau}+\tau\|\} \to 0$, $n \to \infty$ (за ймовірністю чи майже напевно) випливає відповідна збіжність $\hat{x} \to x$, $n \to \infty$. Аналогічно отримуємо, що з результату кожної з теорем 1–4 для неявної моделі (2) випливає результат відповідної теореми 5–8 для моделі (1).

Таким чином, теореми 5–8 справді є наслідками з теорем 1–4.

6. Порівняння з відомими результатами про консистентність

Порівняємо результати теорем 5–8 із відомими умовами консистентності TLS-оцінки, коли останні застосовані до зведеної моделі регресії без вільного члена:

$$\begin{cases} b_i = x_0^\top a_i^0 + \tilde{b}_i, \\ a_i = a_i^0 + \tilde{a}_i, \end{cases} \tag{28}$$

яка отримана з (1) замінами

$$a_i = \begin{pmatrix} 1 \\ x_i \end{pmatrix}, \ a_i^0 = \begin{pmatrix} 1 \\ \xi_i \end{pmatrix}, \ \tilde{a}_i = \begin{pmatrix} 0 \\ \delta_i \end{pmatrix}, \ b_i = y_i, \ x_0 = \begin{pmatrix} b_0 \\ x \end{pmatrix}, \ \tilde{b}_i = \varepsilon_i, \ i = \overline{1, n}. \tag{29}$$

Позначимо $\tilde{c}_i = \left(\tilde{a}_i^\top; \tilde{b}_i\right)^\top$, тоді припущення *(i)*, *(ii)* про модель (1) перетворюються на такі умови.

(iii) Вектори $\tilde{c}_i$, $i \geq 1$, незалежні в сукупності з нульовим середнім.

(iv) Вектори $\tilde{c}_i$, $i \geq 1$, мають однакову коваріаційну матрицю

$$\mathbf{cov}(\tilde{c}_i) = \sigma^2 \cdot \mathrm{diag}(0, 1, 1, \ldots, 1)$$

із невідомим множником $\sigma^2 > 0$.

У моделі (28) оцінка повних найменших квадратів $\hat{x}_{TLS}$ параметра x_0 є розв'язком оптимізаційної задачі

$$\min_{(\Delta a_i \in \mathbb{R}^m, \, \Delta b_i \in \mathbb{R})} \sum_{i=1}^{n} \Big(\|\Delta a_i\|^2 + |\Delta b_i|^2 \Big)$$

за умови, що існує такий $x \in \mathbb{R}^m$, для якого виконано

$$b_i - \Delta b_i = x^\top (a_i - \Delta a_i), \quad i = \overline{1, n}.$$

У роботі [7] наведено найбільш слабкі умови консистентності і строгої консистентності TLS-оцінки у моделі (28) з обмеженнями *(iii)*, *(iv)*. Відповідні теореми мають такі формулювання, де $A_0 \coloneqq [a_1^0, a_2^0, \ldots, a_n^0]^\top$.

Теорема 9. *Нехай у моделі (28) виконані умови (iii), (iv) і для деякого* $1 \leq r \leq 2$

$$\sup_{i \geq 1} \mathsf{E} \, \|\tilde{c}_i\|^{2r} < \infty, \tag{30}$$

$$n^{-1/r} \lambda_{\min}\big(A_0^\top A_0\big) \to \infty, \quad n \to \infty. \tag{31}$$

Тоді $\hat{x}_{TLS} \xrightarrow{\mathsf{P}} x_0$, $n \to \infty$.

Теорема 10. *Нехай у моделі (28) виконані умови (iii), (iv) та для деяких* $r \geq 2$, $n_0 \geq 1$ *справджується таке:*

$$\sup_{i \geq 1} \mathsf{E} \, \|\tilde{c}_i\|^{2r} < \infty,$$

$$\sum_{n=n_0}^{\infty}\left(\frac{\sqrt{n}}{\lambda_{\min}(A_0^\top A_0)}\right)^r < \infty. \tag{32}$$

Тоді $\hat{x}_{TLS} \to x_0$ *м. н.*, $n \to \infty$.

Оскільки теореми 9, 10 дають умови консистентності одночасно оцінок параметрів b_0 та x, то ми будемо порівнювати їх із теоремами 6 і 8 відповідно. Бачимо, що умови (23) і (30) однакові, а умови (31) і (32) отримуються із (25) і (27) відповідно, якщо замінити $\lambda_{\min}(S_{\xi\xi})$ на $\frac{1}{n}\lambda_{\min}\big(A_0^\top A_0\big)$. Тому спочатку дослідимо матрицю $A_0^\top A_0$. Враховуючи позначення (29), маємо

$$A_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \xi_1 & \xi_2 & \dots & \xi_n \end{pmatrix}^\top.$$

Тоді

$$\frac{1}{n}A_0^\top A_0 = \frac{1}{n}\sum_{i=1}^{n} a_i^0 a_i^{0\top} = \begin{pmatrix} 1 & \bar{\xi}^\top \\ \bar{\xi} & \overline{\xi\xi^\top} \end{pmatrix}.$$

Для порівняння чисел $\lambda_{\min}(S_{\xi\xi})$ та $\frac{1}{n}\lambda_{\min}\big(A_0^\top A_0\big)$ використаємо властивості спектра симетричних невід'ємно визначених матриць. Для такої матриці $A \in \mathbb{R}^{m\times m}$ позначатимемо

$$\det(A - \lambda I_m) = (-1)^m\lambda^m + \dots + b(A)\lambda^2 - c(A)\lambda + \det(A). \tag{33}$$

Тоді $c(A)$ дорівнює сумі алгебраїчних доповнень до діагональних елементів матриці A. Далі впорядковуємо всі власні значення A:

$$0 \le \lambda_1 \le \lambda_2 \le \dots \le \lambda_m. \tag{34}$$

За теоремою Вієта

$$\det(A) = \lambda_1\lambda_2\dots\lambda_m, \quad c(A) = \sum_{j=1}^{m}\frac{\lambda_1\lambda_2\dots\lambda_m}{\lambda_j}. \tag{35}$$

Тоді, враховуючи впорядкованість власних чисел, маємо

$$\frac{1}{m}\lambda_1 = \frac{\lambda_1\lambda_2\dots\lambda_m}{m\lambda_2\lambda_3\dots\lambda_m} \le \frac{\det(A)}{c(A)} \le \frac{\lambda_1\lambda_2\dots\lambda_m}{\lambda_2\lambda_3\dots\lambda_m} = \lambda_1. \tag{36}$$

Якщо для кожного $i = \overline{2, m}$ від i-го рядка матриці $\frac{1}{n}A_0^\top A_0$ відняти перший, помножений на $(i-1)$-й елемент вектора $\bar{\xi}$, то отримаємо, що

$$\det\left(\frac{1}{n}A_0^\top A_0\right) = \det(S_{\xi\xi}). \tag{37}$$

Використовуючи цей результат для розмірності на одиницю менше, одержуємо

$$c\left(\frac{1}{n}A_0^\top A_0\right) = c(S_{\xi\xi}) + \det\left(\overline{\xi\xi^\top}\right) \ge c(S_{\xi\xi}). \tag{38}$$

Тому з (36), із врахуванням (37) і (38), випливає, що

$$\lambda_{\min}\left(\frac{1}{n}A_0^\top A_0\right) \le \frac{m\det\big(\frac{1}{n}A_0^\top A_0\big)}{c\big(\frac{1}{n}A_0^\top A_0\big)} \le \frac{m\det(S_{\xi\xi})}{c(S_{\xi\xi})} \le m\lambda_{\min}(S_{\xi\xi}). \tag{39}$$

Тобто умови (25) і (27) є слабшими за умови (31) і (32) відповідно. А для типового випадку, коли $\lambda_{\max}(S_{\xi\xi}) = O(\lambda_{\min}(S_{\xi\xi}))$, $n \to \infty$, умови (24) і (26) будуть випливати з умов (25) і (27) відповідно.

7. Висновки

Ми розглянули неявну лінійну модель регресії з похибками у змінних і отримали досить м'які умови консистентності ООР. Як наслідок ми одержали теореми про консистентність TLS-оцінки у множинній лінійній регресії зі скалярним відгуком, вільним членом і похибками у змінних, при цьому відповідні твердження не випливають із результатів [7]. Цікаво було б подібним чином дослідити TLS-оцінку у відповідній множинній регресії з векторним відгуком і вільним членом.

Список літератури

1. W. Fuller, *Measurement error models*, Wiley, New York, 1987.
2. P. Gallo, *Consistency of regression estimates when some variables are subject to errors*, Comm. Statist. Theory Methods, **11** (1982), no. 9, 973–983.
3. L. J. Gleser, *Estimation in multivariate "errors-in-variables" regression model: large sample results*, Ann. Statist., **9** (1981), no. 1, 24–44.
4. A. I. Kostrikin, Yu. I. Manin, *Linear Algebra and Geometry*, Series in Algebra, Logic and Applications, vol. 1, CRC Press, 1989.
5. A. Kukush, S. Van Huffel, *Consistency of elementwise-weighted total lest squares estimator in a mulivariate errors-in-variables model $AX = B$*, Metrika, **59** (2004), no. 1, 75–97.
6. S. V. Masiuk, A. G. Kukush, S. V. Shklyar, M. I. Chepurny, I. A. Likhtarov (ed.), *Radiation risk estimation: Based on measurement error models*, 2nd ed., De Gruyter Ser. Math. Life Sci., vol. 5, De Gruyter, 2017.
7. S. V. Shklyar, *Conditions for the consistency of the total least squares estimator in an errors-in-variables linear regression model*, Theory Probab. Math. Statist., **83** (2011), 175–190.

Кафедра математичного аналізу, механіко-математичний факультет, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64/13, м. Київ, Україна, 01601

Адреса електронної пошти: oodashkov@gmail.com

Кафедра математичного аналізу, механіко-математичний факультет, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64/13, м. Київ, Україна, 01601

Адреса електронної пошти: alexander_kukush@univ.kiev.ua

Стаття надійшла до редколегії 4.10.2017

CONSISTENCY OF ORTHOGONAL REGRESSION ESTIMATOR IN IMPLICIT LINEAR ERRORS-IN-VARIABLES MODEL

O. O. DASHKOV, A. G. KUKUSH

Abstract. An implicit linear errors-in-variables model is considered, where the true points belong to a hyperplane in a Euclidean space and the total error variance-covariance matrix is proportional to the identity matrix. The orthogonal regression estimator of the hyperplane is studied. Sufficient conditions for the consistency and strong consistency of the estimator are presented. The results are applied to an explicit multiple errors-in-variables model with intercept.

СОСТОЯТЕЛЬНОСТЬ ОЦЕНКИ ОРТОГОНАЛЬНОЙ РЕГРЕССИИ В НЕЯВНОЙ ЛИНЕЙНОЙ МОДЕЛИ С ОШИБКАМИ В ПЕРЕМЕННЫХ

А. А. ДАШКОВ, А. Г. КУКУШ

Аннотация. Рассматривается неявная линейная модель регрессии с ошибками в переменных, где истинные точки лежат на некоторой гиперплоскости в евклидовом пространстве, а совместная корреляционная матрица ошибок пропорциональна единичной матрице. Изучается оценка ортогональной регрессии этой гиперплоскости. Приведены достаточные условия состоятельности и строгой состоятельности оценки. Результаты применяются к явной множественной модели регрессии со свободным членом и ошибками в переменных.