Теорія ймовір. та матем. статист.
Вип. 92, 2015, стор. 79–88

Teor. Ĭmovir. ta Matem. Statyst.
No. 92, 2015, pp. 79–88

ЗБІЖНІСТЬ ОЦІНОК У ПОЛІНОМІАЛЬНІЙ ФУНКЦІОНАЛЬНІЙ МОДЕЛІ З ПОХИБКАМИ ВИМІРЮВАННЯ

УДК 519.21

О. Г. КУКУШ І Я. В. ЦАРЕГОРОДЦЕВ

Анотація. Розглядається поліноміальна функціональна модель з похибками вимірювання. Відомі дисперсія похибки у регресії та її коваріація з похибкою у відгуку. Виправлена оцінка найменших квадратів параметра регресії адаптує звичайну оцінку найменших квадратів до наявної похибки в регресорі. Наведено умови строгої консистентності оцінки, слабші, ніж у статті Cheng and Schnee-weiss (1998) [Journal of the Royal Statistical Society B, №1, 189–199], а також умови асимптотичної нормальності оцінки.

Abstract. A polynomial measurement error model is considered. The variance of the errors in the regressor variable and its covariance with the errors of the response variable are assumed known. The adjusted least squares estimator of regression parameters adapts an ordinary least squares estimator for the errors in the regressor. Conditions for the strong consistency of the estimator are presented that are weaker compared with Cheng and Schneeweiss (1998) [Journal of the Royal Statistical Society B, №1, 189–199], as well as conditions for the asymptotic normality of the estimator.

Аннотация. Рассматривается полиномиальная модель с ошибками измерений. Известны дисперсия ошибки в регрессоре и ее ковариация с ошибкой в отклике. Исправленная оценка наименьших квадратов параметра регрессора адаптирует обычную оценку наименьших квадратов к наличию ошибки в регрессоре. Приведены условия строгой состоятельности оценки, более слабые, нежели в статье Cheng and Schneeweiss (1998) [Journal of the Royal Statistical Society B, №1, 189–199], а также условия асимптотической нормальности оценки.

1. Вступ

Поліноміальний функціональний зв'язок задається рівнянням

$$\begin{aligned} y_i &= \beta_0 + \beta_1\xi_i + \beta_2\xi_i^2 + \cdots + \beta_k\xi_i^k + \varepsilon_i, \\ x_i &= \xi_i + \delta_i, \qquad i = 1, \ldots, n, \end{aligned}$$

де похибки $(\delta_i, \varepsilon_i)$, $i \geq 1$, утворюють незалежні та однаково розподілені (н.о.р.) пари випадкових величин з нульовими середніми і коваріційною матрицею

$$\Omega := \begin{pmatrix} \sigma_\delta^2 & \sigma_{\delta\varepsilon} \\ \sigma_{\delta\varepsilon} & \sigma_\varepsilon^2 \end{pmatrix}. \tag{1}$$

Тут ξ_i, $i \geq 1$, є неспостережуваними невипадковими величинами, які називають прихованими змінними. Це відповідає так званому *функціональному* випадку моделі з похибками вимірювання. У *структурному* випадку ξ_i, $i \geq 1$, утворювали би послідовність н.о.р. випадкових величин.

2000 *Mathematics Subject Classification.* Primary 62F12, 62J02.

Ключові слова і фрази. Асимптотична нормальність, виправлена оцінка найменших квадратів, консистентність оцінки, модель з похибками вимірювання, модифікація оцінок для малої вибірки, поліноміальна регресія.

Автори щиро вдячні проф. Г. Шнеєвайссу (H. Schneeweiss, Мюнхен, Німеччина), який звернув нашу увагу на поліноміальну модель з похибками вимірювання, а також канд. фіз.-мат. наук С. В. Шкляру (Київ) за плідні обговорення.

Параметри σ_δ^2 та $\sigma_{\delta\varepsilon}$ вважаємо відомими. За спостереженями (x_i, y_i), $i = 1, \ldots, n$, треба оцінити параметри регресії $\beta_0, \ldots, \beta_k$ та, можливо, дисперсію σ_ε^2.

C.-L. Cheng та H. Schneeweiss [4] побудували виправлену оцінку найменших квадратів (adjusted least squares estimator; ALS-оцінка) параметрів регресії; там же див. огляд літератури з розглядуваної поліноміальної моделі регресії. C.-L. Cheng та ін. [5] відзначили, що ALS-оцінка є нестійкою для малих та середніх виборок, і побудували модифіковану ALS-оцінку (MALS-оцінку), яка краще себе поводить для малих і середніх виборок та є асимптотично еквівалентною до ALS-оцінки при $n \to \infty$. Тому на практиці краще використовувати MALS-оцінку, особливо коли обсяг вибирки невиликий.

C.-L. Cheng та A. Kukush [6] на основі ALS-оцінки побудували критерій згоди для функціональної поліноміальної моделі з похибками вимірювання; оскільки ALS-оцінка гарно працює і в структурному випадку, то цей критерій можна застосувати і в структурній моделі. P. Hall та Y. Ma [7] для структурної поліноміальної моделі побудували більш потужний критерій згоди.

Метою даного дослідження є послабити умови консистентності ALS-оцінки параметра регресії в порівнянні з [4], а також дати умови асимптотичної нормальності цієї оцінки (чого не було зроблено в [4]). Для консистентності ми не вимагаємо збіжності певних вибіркових моментів прихованої змінної ξ, а натомість вимагаємо обмеженість вибіркового моменту порядку $2k$ та відділеність від нуля визначника певної матриці, яка є аналогом інформаційної матриці для скінченної вибірки. Проте для асимптотичної нормальності оцінки ми вимагатимемо збіжність вибіркових моментів до порядку $4k-2$ включно. Звичайно, MALS-оцінка успадковує асимптотичні властивості ALS-оцінки. Зауважимо, що аналогічні результати для ALS-оцінки у векторній лінійній моделі отримано в [11, 12].

Стаття побудована наступним чином. У розділі 2 ми описуємо моделі спостережень та вводимо ALS-оцінки $\hat\beta_A$ та $\hat\sigma^2_{\varepsilon,A}$. Розділ 3 містить результати про строгу консистентність ALS-оцінок. У розділі 4 доводиться асимптотична нормальність $\hat\beta_A$ і здійснюється консистенте оцінювання її асимптотичної коваріційної матриці. Розділ 5 окреслює напрямки подальших досліджень.

Вживаються наступні позначення. Символ E означає математичне сподівання, **cov** — коваріаційну матрицю випадкового вектора; верхня риска означає усереднення за номером спостереження i, наприклад, $\overline{\xi} = \frac{1}{n}\sum_{i=1}^n \xi_i$, $\overline{xy} = \frac{1}{n}\sum_{i=1}^n x_i y_i$ тощо. Верхній індекс T означає транспонування; в роботі всі вектори є векторами-стовпчиками. Збіжність з імовірністю 1 та за розподілом позначається відповідно $\xrightarrow{\mathsf{P1}}$, $\xrightarrow{\mathrm{d}}$. Послідовність випадкових векторів чи випадкових матриць, що збігається до нуля з імовірністю 1, позначається $o(1)$. Запис $\varepsilon \stackrel{\mathrm{d}}{=} \varepsilon_1$ означає, що величини ε, ε_1 мають однаковий імовірнісний розподіл. Додатні сталі, які не залежать від обсягу вибірки n, позначаються const; можливі записи типу $2 \cdot \mathrm{const} = \mathrm{const}$.

2. Модель спостережень і виправлена оцінка найменших квадратів

При фіксованому $k \ge 1$ ми розглядаємо модель спостережень

$$y_i = \rho_i^{\mathsf{T}} \beta + \varepsilon_i, \qquad x_i = \xi_i + \delta_i, \qquad i = 1, \ldots, n, \tag{2}$$

де $\rho_i := (1, \xi_i, \xi_i^2, \ldots, \xi_i^k)^{\mathsf{T}}$ та $\beta := (\beta_0, \beta_1, \ldots, \beta_k)^{\mathsf{T}}$; ξ_i, $i = 1, \ldots, n$, — це детерміновані значення прихованої змінної ξ; похибки $(\delta_i, \varepsilon_i)$, $i = 1, \ldots, n$, утворюють н.о.р. пари випадкових величин з нульовими середніми і коваріаційною матрицею Ω, заданою в (1), з додатними дисперсіями σ_δ^2, σ_ε^2.

Вважаємо відомими всі моменти $\mathsf{E}\,\delta^j$, $j = 2, \ldots, 2k$, та $\mathsf{E}\,\delta^j \varepsilon$, $j = 1, \ldots, k$. Слідуючи [4], побудуємо виправлену оцінку найменших квадратів (ALS-оцінку) параметра β.

Якби спостерігались (ξ_i, y_i), $i = 1, \dots, n$, то звичайна оцінка найменших квадратів будувалась би за допомогою оціночної функції

$$S_{LS}^{(\beta)}(\xi, y; \beta) := \rho y - \rho\rho^{\mathsf{T}}\beta, \qquad \rho = \rho(\xi) := \left(1, \xi, \dots, \xi^k\right)^{\mathsf{T}}.$$

Це незсунена оціночна функція, тобто при всіх $\beta \in \mathbb{R}^{k+1}$ виконується

$$\mathsf{E}_\beta\, S_{LS}^{(\beta)}(\xi, y; \beta) = 0. \tag{3}$$

Тут і нижче E_β означає математичне сподівання за умови, що β — це істинне значення параметра регресії, тобто що $y = \rho^{\mathsf{T}}(\xi)\beta + \varepsilon$, $\varepsilon \overset{\mathrm{d}}{=} \varepsilon_1$; ξ — детермінована величина.

Замість ξ ми спостерігаємо $x = \xi + \delta$, $(\delta, \varepsilon) \overset{\mathrm{d}}{=} (\delta_1, \varepsilon_1)$. Далі будується виправлена оціночна функція $S_C^{(\beta)}(x, y; \beta)$, така що

$$\mathsf{E}_\beta\, S_C^{(\beta)}(x, y; \beta) = \mathsf{E}_\beta\, S_{LS}^{(\beta)}(\xi, y; \beta), \qquad \beta \in \mathbb{R}^{k+1}.$$

Завдяки рівності (3) нова оціночна функція також буде незсуненою, що є необхідною умовою консистентності ALS-оцінки, яка буде задаватися оціночною функцією $S_C^{(\beta)}$.

Щоб побудувати $S_C^{(\beta)}$, треба сконструювати матрицю $H(x)$ та вектор $h(x, y)$ такі, що

$$\mathsf{E}\, H(x) = \rho\rho^{\mathsf{T}}, \tag{4}$$

$$\mathsf{E}_\beta\, h(x, y) = \mathsf{E}_\beta\, \rho y; \qquad \xi \in \mathbb{R},\ \beta \in \mathbb{R}^{k+1}, \tag{5}$$

і тоді виправлена оціночна функція дорівнюватиме

$$S_C^{(\beta)} = h(x, y) - H(x)\beta. \tag{6}$$

Для побудови $H(x)$ спочатку при кожному $0 \le r \le 2k$ будуються многочлени

$$t_r = \sum_{j=0}^{r} a_{rj} x^j, \tag{7}$$

де a_{rj} є функціями від $\mathsf{E}\,\delta^p$, $p = 0, \dots, r$, причому для цих многочленів (7) виконується

$$\mathsf{E}\, t_r(x) = \xi^r, \qquad \xi \in \mathbb{R}.$$

Наприклад, $t_0 = 1$, $t_1 = x$, $t_2 = x_2 - \sigma_\delta^2$, $t_3 = x^3 - 3x\sigma_\delta^2 - \mathsf{E}\,\delta^3$. Деталі побудови t_r дивись у [4]. Тепер розв'язок рівняння (4) задається матрицею H з елементами

$$H_{ij}(x) = t_{i+j}(x), \qquad i, j = 0, \dots, k. \tag{8}$$

При цьому задіяні відомі моменти $\mathsf{E}\,\delta^p$, $p = 0, \dots, 2k$.

Далі, розв'язок рівняння (5) має вигляд

$$h = (h_r)_{r=0}^{k}, \qquad h_r = t_r y - \sum_{j=0}^{r} b_{rj} t_j, \tag{9}$$

причому b_{rj} є функціями від $\mathsf{E}\,\delta^p$, $p = 0, \dots, r$, та $\mathsf{E}\,\delta^p\varepsilon$, $p = 0, \dots, r$. Зокрема $h_0 = y$, $h_1 = xy - \sigma_{\delta\varepsilon}$, $h_2 = (x^2 - \sigma_\delta^2)y - \mathsf{E}\,\delta^2\varepsilon - 2\sigma_{\delta\varepsilon}x$. Деталі побудови h_r дивись у [4].

Після того, як побудовано оціночну функцію (6), ALS-оцінка $\hat\beta_A$ задається рівнянням

$$\frac{1}{n}\sum_{i=1}^{n} S_C^{(\beta)}(x_i, y_i; \beta) = 0,$$

або

$$\overline{H}\beta = \overline{h}, \qquad \beta \in \mathbb{R}^{k+1}. \tag{10}$$

Тут $\overline{H}$ та $\overline{h}$ — це усереднені значення матриць $H_{(i)} := H(x_i)$ та векторів $h_{(i)} := h(x_i, y_i)$.

Означення 1. Оцінка $\hat{\beta}_A$ задається виразом

$$\hat{\beta}_A = \overline{H}^{-1}\overline{h}, \tag{11}$$

якщо матриця $\overline{H}$ є невиродженою; якщо ж вона вироджується, то покладаємо $\hat{\beta}_A = 0$.

Також побудуємо оцінку дисперсії σ_ε^2. Якби спостерігались (ξ_i, y_i), $i = 1, \ldots, n$, то наступна оціночна функція була б незсуненою:

$$S_{LS}^{(\sigma_\varepsilon^2)}\left(\xi, y; \beta, \sigma_\varepsilon^2\right) := y^2 - y\rho^{\mathsf{T}}(\xi)\beta - \sigma_\varepsilon^2,$$

тобто

$$\mathsf{E}_{\beta,\sigma_\varepsilon^2} S_{LS}^{(\sigma_\varepsilon^2)}\left(\xi, y; \beta, \sigma_\varepsilon^2\right) = 0, \qquad \beta \in \mathbb{R}^{k+1},\ \sigma_\varepsilon^2 > 0.$$

Виправлена оціночна функція задається рівністю

$$S_C^{(\sigma_\varepsilon^2)}\left(x, y; \beta, \sigma_\varepsilon^2\right) = y^2 - h^{\mathsf{T}}(x, y)\beta - \sigma_\varepsilon^2.$$

З огляду на рівність (5), вона теж буде незсуненою, оскільки

$$\mathsf{E}_{\beta,\sigma_\varepsilon^2} S_C^{(\sigma_\varepsilon^2)}\left(x, y; \beta, \sigma_\varepsilon^2\right) = \mathsf{E}_{\beta,\sigma_\varepsilon^2} S_{LS}^{(\sigma_\varepsilon^2)}\left(\xi, y; \beta, \sigma_\varepsilon^2\right).$$

Для оцінювання σ_ε^2 маємо додаткове до (10) рівняння

$$\frac{1}{n}\sum_{i=1}^{n} S_C^{(\sigma_\varepsilon^2)}\left(x_i, y_i; \beta, \sigma_\varepsilon^2\right) = 0;$$

$$\sigma_\varepsilon^2 = \overline{y^2} - \overline{h}^{\mathsf{T}}\beta.$$

Означення 2. Оцінка $\hat{\sigma}_{\varepsilon,A}^2$ задається виразом

$$\hat{\sigma}_{\varepsilon,A}^2 = \overline{y^2} - \overline{h}^{\mathsf{T}}\hat{\beta}_A. \tag{12}$$

3. Консистентність ALS-оцінок

Лема 3. *Нехай $r > 1$ — фіксоване дійсне число, $\{\eta_k, k \ge 1\}$ — послідовність н.о.р. випадкових величин з нульовим середнім, причому $\mathsf{E}\,|\eta_1|^r < \infty$. Нехай також задано числову послідовність $\{a_k, k \ge 1\}$, для якої*

$$\overline{|a|^r} = \frac{1}{n}\sum_{k=1}^{n} |a_k|^r \le \mathrm{const}\,.$$

Тоді

$$\overline{a\eta} = \frac{1}{n}\sum_{k=1}^{n} a_k\eta_k \xrightarrow{\mathsf{P1}} 0.$$

Доведення. Без обмеження загальності можна вважати, що $1 < r \le 2$. Для випадкових величин $X_n = a_n\eta_n$, $n \ge 1$, розглянемо ряд

$$\sum_{n=1}^{\infty} \frac{\mathsf{E}\,|X_n|^r}{n^r} = \mathsf{E}\,|\eta_1|^r \sum_{n=1}^{\infty} \frac{|a_n|^r}{n^r}. \tag{13}$$

Позначимо $A_0 = 0$, $A_n = \frac{1}{n}\sum_{k=1}^{n} |a_k|^r$, $n \ge 1$. Тоді $|a_n|^r = nA_n - (n-1)A_{n-1}$, $n \ge 1$. Ряд у правій частині (13) збігається, бо його сума дорівнює

$$\sum_{n=1}^{\infty} \frac{nA_n - (n-1)A_{n-1}}{n^r} = \sum_{m=1}^{\infty} A_m\,\frac{(m+1)^r - m^r}{m^{r-1}(m+1)^r},$$

а останній ряд збігається внаслідок обмеженості $\{A_m\}$. Тоді завдяки тому, що $1 < r \le 2$ та ряд (13) збігається, за теоремою 12 із [3, с. 222], маємо:

$$\overline{a\eta} = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{\mathsf{P1}} 0. \qquad \square$$

Введемо поняття, яке буде корисним при дослідженні строгої консистентності оцінок.

Означення 4. Нехай задано послідовність тверджень $\{A_n(\omega), n \ge 1\}$, які залежать від елементарної події $\omega \in \Omega$. Кажемо, що твердження $A_n = A_n(\omega)$ виконується *зрештою*, якщо існує така подія Ω_0, $\mathsf{P}(\Omega_0) = 1$, що для всіх $\omega \in \Omega_0$ знайдеться номер $n_0(\omega)$, для якого при всіх $n \ge n_0(\omega)$ твердження $A_n(\omega)$ виконується.

Наступні умови будуть потрібні для консистентності оцінки $\hat\beta_A$.

(i) $\overline{\xi^{2k}} = \frac{1}{n}\sum_{i=1}^{n} \xi_i^{2k} \le \mathrm{const}$.
(i′) При деякому $r > 1$ виконується $\mathsf{E}\,|\delta^{k-1}\varepsilon|^r < \infty$.
(ii) $\underline{\lim}_{n\to\infty} \det(\overline{\rho\rho^\mathsf{T}}) > 0$.

Теорема 5. *За умов (i), (i′), (ii) оцінка $\hat\beta_A$ строго консистентна, тобто $\hat\beta_A \xrightarrow{\mathsf{P1}} \beta$, $n \to \infty$.*

Доведення. а) Спочатку покажемо, що матриця $\overline{H}$ невироджена зрештою (в сенсі означення 4). Згідно з (8) матричні елементи $\overline{H}_{ij} = \overline{t}_{i+j}$. Різниця $\overline{t}_{i+j} - \mathsf{E}\,\overline{t}_{i+j}$ є алгебраїчною сумою доданків виду $z_p := \overline{x^p} - \mathsf{E}\,\overline{x^p}$, $0 \le p \le 2k$. Розписуючи $x^p = (\xi + \delta)^p$ за формулою бінома Ньютона, отримуємо, що z_p є алгебраїчною сумою доданків виду $g_{uv} := \overline{\xi^u\delta^v} - \mathsf{E}\,\overline{\xi^u\delta^v}$, $u, v \ge 0$, $u + v \le p$. При $u = 0$ маємо за ПЗВЧ, що $\overline{\delta^v} - \mathsf{E}\,\overline{\delta^v} \xrightarrow{\mathsf{P1}} 0$: адже $v \le 2k$ та $\mathsf{E}\,\delta^{2k} < \infty$. Якщо $v = 0$, то $g_{u0} = 0$. При $u, v \ge 1$, $u + v \le 2k$, маємо

$$g_{uv} = \overline{\xi^u(\delta^v - \mathsf{E}\,\delta^v)}. \tag{14}$$

Застосуємо лему 3:

$$\overline{\xi^{ru}} \le \mathrm{const}$$

при $r = \frac{2k}{2k-1}$; для цього ж r маємо

$$\mathsf{E}\,|\delta^v - \mathsf{E}\,\delta^v|^r \le \mathrm{const}\,\mathsf{E}\,|\delta|^{vr} \le \mathrm{const}(1 + \mathsf{E}\,|\delta|^{(2k-1)r}) < \infty;$$

тож лема забезпечує збіжність $g_{uv} \xrightarrow{\mathsf{P1}} 0$, $n \to \infty$. Таким чином,

$$\overline{H} - \mathsf{E}\,\overline{H} = \overline{H} - \overline{\rho\rho^\mathsf{T}} \xrightarrow{\mathsf{P1}} 0, \qquad n \to \infty. \tag{15}$$

Розглянемо матрицю $\Phi_n := \overline{\rho\rho^\mathsf{T}}$. За умовою (ii) вона невироджена при $n \ge n_0$. За умовою (i) її матричні елементи $\overline{\xi^{i+j}}$, $0 \le i, j \le k$, обмежені. Обернена матриця Φ_n^{-1} виписується через визначник та приєднану матрицю. Тому для евклідової матричної норми маємо при $n \ge n_0$:

$$\left\|\Phi_n^{-1}\right\| \le \frac{\mathrm{const}}{\det(\Phi_n)}, \qquad \overline{\lim_{n\to\infty}} \left\|\Phi_n^{-1}\right\| \le \frac{\mathrm{const}}{\underline{\lim}_{n\to\infty} \det(\Phi_n)} < \infty,$$
$$\left\|\Phi_n^{-1}\right\| \le \mathrm{const}\,. \tag{16}$$

Збіжність (15) та нерівність (16) показують, що зрештою $\left\|\overline{H} - \Phi_n\right\| < \left\|\Phi_n^{-1}\right\|^{-1}$. Як тільки остання нерівність виконується, то матриця $\overline{H}$ стає невиродженою за теоремою про збурення оборотного оператора; тож справді $\overline{H}$ невироджена зрештою. Тоді зрештою $\hat\beta_A$ задається формулою (11).

б) Поведінка $\overline{h}$.

З формули (9) маємо при $0 \le r \le k$:

$$h_r = t_r \rho^{\mathsf{T}} \beta + t_r \varepsilon - \sum_{j=0}^{r} b_{rj} t_j. \tag{17}$$

Ми хочемо перевірити, що

$$\overline{h}_r - \mathsf{E}_\beta \, \overline{h}_r \xrightarrow{\mathsf{P1}} 0.$$

У порівнянні з п. а) доведення, новим для нас є лише доданок

$$\overline{t_r \varepsilon} - \mathsf{E} \, \overline{t_r \varepsilon}, \qquad 0 \le r \le k.$$

Він являє собою алгебраїчну суму доданків виду

$$f_{ij} := \overline{\xi^i \delta^j \varepsilon} - \mathsf{E} \, \overline{\xi^i \delta^j \varepsilon}, \qquad i + j \le k. \tag{18}$$

При $i = 0$ за ПЗВЧ виконується $f_{0j} \xrightarrow{\mathsf{P1}} 0$. Нехай тепер $i \ge 1, \ j \le k - i$. Маємо $\overline{\xi^{2i}} \le \mathrm{const}$, та при $r > 1$ з умови (i') отримуємо

$$\mathsf{E} \, |\delta^j \varepsilon - \mathsf{E} \, \delta^j \varepsilon|^r \le \mathrm{const} \cdot \mathsf{E} \, |\delta^j \varepsilon|^r < \infty.$$

Тоді за лемою 3, $f_{ij} \xrightarrow{\mathsf{P1}} 0$. Таким чином,

$$\overline{h} - \mathsf{E}_\beta \, \overline{h} \xrightarrow{\mathsf{P1}} 0. \tag{19}$$

З огляду на (5), маємо тоді

$$h = \overline{\rho \rho^{\mathsf{T}}} \beta + o(1) = \Phi_n \beta + o(1).$$

в) З формули (11) маємо зрештою:

$$\hat{\beta}_A = (\Phi_n + o(1))^{-1} (\Phi_n \beta + o(1)) = \left(I + \Phi_n^{-1} o(1) \right)^{-1} \left(\beta + \Phi_n^{-1} o(1) \right). \tag{20}$$

Тут I позначає одиничну матрицю. За нерівністю (16) маємо

$$\left\| \Phi_n^{-1} o(1) \right\| \le \left\| \Phi_n^{-1} \right\| \cdot \| o(1) \| \le \mathrm{const} \cdot o(1) = o(1). \tag{21}$$

Нарешті, із співвідношень (20), (21) випливає шукана збіжність $\hat{\beta}_A \xrightarrow{\mathsf{P1}} \beta$. □

Розглянемо додаткову умову.

(iii) При деякому $c > 0$ виконується $\mathsf{E} \, |\varepsilon|^{2+c} < \infty$.

Теорема 6. *Нехай виконані умови (i), (i′), (ii), (iii). Тоді обидві оцінки $\hat{\beta}_A$, $\hat{\sigma}^2_{\varepsilon,A}$ строго консистентні, зокрема*

$$\hat{\sigma}^2_{\varepsilon,A} \xrightarrow{\mathsf{P1}} \sigma^2_\varepsilon, \qquad n \to \infty.$$

Доведення. Залишилося довести строгу консистентність оцінки дисперсії. Маємо

$$\overline{y^2} = \beta^{\mathsf{T}} \overline{\rho \rho^{\mathsf{T}}} \beta + 2 \beta^{\mathsf{T}} \overline{\rho \varepsilon} + \overline{\varepsilon^2}.$$

За лемою 3 $\overline{\rho \varepsilon} \xrightarrow{\mathsf{P1}} 0$, бо виконана умова (iii); також $\overline{\varepsilon^2} \xrightarrow{\mathsf{P1}} \sigma^2_\varepsilon$. Тоді $\overline{y^2} = \mathsf{E} \, \overline{y^2} + o(1)$, і з рівності (12) маємо розклад

$$\hat{\sigma}^2_{\varepsilon,A} = \mathsf{E} \, \overline{y^2} + o(1) - \left(\beta^{\mathsf{T}} \overline{\rho \rho^{\mathsf{T}}} + o(1) \right) (\beta + o(1)) = \sigma^2_\varepsilon + o(1). \qquad \square$$

Зауваження 7. Теорема 6 істотно послаблює умови консистентності оцінок у порівнянні з [4]: використані моменти похибок менших порядків, та замість існування невиродженої границі матриць Φ_n вимагається лише асимптотична відділеність $\det \Phi_n$ від нуля і обмеженість вибіркового моменту $\overline{\xi^{2k}}$; збіжність вибіркових моментів ми не вимагаємо.

4. Асимптотична нормальність $\hat{\beta}_A$

Наступна центральна гранична теорема є варіантом теореми Ляпунова [2, с. 73].

Теорема 8. *Нехай $\{z_i, i \ge 1\}$ — послідовність незалежних випадкових векторів, розподілених в $\mathbb{R}^d$ і з нульовими середніми, $\overline{\mathbf{cov}(z)} = \frac{1}{n}\sum_{i=1}^n \mathbf{cov}(z_i) \to S$, $n \to \infty$, причому існує таке $c > 0$, що*

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\,\|z_i\|^{2+c} \le \text{const}\,. \tag{22}$$

Тоді

$$\sqrt{n}\overline{z} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i \xrightarrow{\mathrm{d}} N(0, S).$$

Наступні умови забезпечують асимптотичну нормальність оцінки.

(iv) При кожному $r = 0, \ldots, 4k-2$ існує скінченна границя

$$\mu_r := \lim_{n\to\infty} \overline{\xi^r}, \tag{23}$$

причому матриця $H_\infty := (\mu_{i+j})_{i,j=0}^{k}$ є невиродженою.

(v) Нехай $\mathsf{E}\,\delta^{2k}\varepsilon^2 < \infty$ та при деякому $c > 0$ виконується:

$$\overline{|\xi|^{4k-2+c}} \le \text{const}, \qquad \mathsf{E}\,|\delta|^{4k-2+c} < \infty, \qquad \mathsf{E}\,|\varepsilon|^{2+c} < \infty, \qquad \mathsf{E}\,|\delta^{k-1}\varepsilon|^{2+c} < \infty.$$

Теорема 9. *Нехай виконані умови (iv) – (v). Тоді*

$$\sqrt{n}(\hat{\beta}_A - \beta) \xrightarrow{\mathrm{d}} N(0, \Sigma), \tag{24}$$

$$\Sigma = H_\infty^{-1} U H_\infty^{-1}, \qquad U = \lim_{n\to\infty} \overline{\mathbf{cov}(h - H\beta)} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{cov}\left(h_{(i)} - H_{(i)}\beta\right),$$

де $h_{(i)} = h(x_i, y_i)$, $H_{(i)} = H(x_i)$.

Доведення. Умови теореми 5 виконані, і тому

$$\overline{H} = \mathsf{E}\,\overline{H} + o(1) = \overline{\rho\rho^{\mathsf{T}}} + o(1).$$

Тоді внаслідок збіжності (23) маємо $\overline{H} \xrightarrow{\mathsf{P1}} H_\infty$, причому остання матриця додатно визначена. Тому $\overline{H}$ також додатно визначена — зрештою (див. означення 4). Маємо зрештою:

$$\sqrt{n}(\hat{\beta}_A - \beta) = \overline{H}^{-1} \cdot \sqrt{n} \cdot \overline{h - H\beta}. \tag{25}$$

До незалежних випадкових векторів $z_i := h_{(i)} - H_{(i)}\beta$ застосуємо теорему 8. За формулою (17)

$$z := h - H\beta = (t\rho^{\mathsf{T}} - H)\beta - Bt + t\varepsilon, \tag{26}$$

де елементи матриці B дорівнюють

$$B_{rj} = \begin{cases} b_{rj}, & 0 \le j \le r \le k; \\ 0, & j > r,\ 0 \le r \le k-1. \end{cases}$$

За побудовою функцій h, H маємо $\mathsf{E}\,z = 0$, і тоді за умовою (iv) коректно визначена матриця $U := \lim_{n\to\infty} \overline{\mathbf{cov}(z)} = \lim_{n\to\infty} \mathsf{E}\,\overline{zz^{\mathsf{T}}}$. Наприклад, μ_{4k-2} входить у вираз $\lim_{n\to\infty} \mathsf{E}\,\overline{(t\rho^{\mathsf{T}} - H)\beta\beta^{\mathsf{T}}(t\rho^{\mathsf{T}} - H)^{\mathsf{T}}}$. Тут важливо, що вираз $t_k\rho_k - H_{kk} = t_k\rho_k - t_{2k}$ як многочлен від ξ містить ξ^{2k-1}, але не містить ξ^{2k}. Крім того, вираз $t\varepsilon$ містить доданок $\delta^k\varepsilon$, який має скінченну дисперсію завдяки першій нерівності в умові (v).

Умова Ляпунова (22) виконується завдяки вимогам (v). Наприклад, вираз $t_k\rho_k - H_{kk}$ містить доданки виду $\mathrm{const}\cdot\xi^{2k-1}\delta$. При $r>2$ розглянемо

$$\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\,|\xi_i^{2k-1}\delta_i|^r = \mathsf{E}\,|\delta|^r\cdot\overline{|\xi|^{(2k-1)r}}<\infty,$$

якщо $r=2+\frac{c}{2k-1}$ зі сталою c з умови (v). Аналогічно з використанням умови (v) перевіряється умова Ляпунова для інших складових вектора z.

Тож за теоремою 8

$$\sqrt{n}\cdot\overline{h-H\beta}\xrightarrow{\mathrm{d}}N(0,U),$$

і за лемою Слуцького [1] з рівності (25) отримуємо:

$$\sqrt{n}(\hat\beta_A-\beta)\xrightarrow{\mathrm{d}}N\left(0,H_\infty^{-1}UH_\infty^{-1}\right). \qquad \square$$

Теорема 10. *Нехай виконана умова (iv), а також наступні умови.*

(vi) $\mathsf{E}\,\delta^{4k}<\infty$, $\mathsf{E}\,\delta^{2k}\varepsilon^2<\infty$ *та при деякому $c>0$ виконується:*

$$\mathsf{E}\,|\varepsilon|^{2+c}<\infty,\qquad \mathsf{E}\,|\delta^{2k-1}\varepsilon^2|^{1+c}<\infty,\qquad \mathsf{E}\,|\delta^{3k-1}\varepsilon|^{1+c}<\infty.$$

(vii) $\overline{|\xi|^{4k}}\le\mathrm{const}$.

Тоді наступна матриця $\hat\Sigma$ є строго консистентною оцінкою матриці Σ з (24):

$$\hat\Sigma=H^{-1}\hat UH^{-1},\qquad \hat U=\frac{1}{n}\sum_{i=1}^{n}\left(h_{(i)}-H_{(i)}\hat\beta_A\right)\left(h_{(i)}-H_{(i)}\hat\beta_A\right)^{\mathsf T},$$

тобто $\hat\Sigma\xrightarrow{\mathsf{P1}}\Sigma$ при $n\to\infty$.

Доведення. Зараз виконані умови теореми 9, тому має місце збіжність (24), та оцінка $\hat\beta_A$ строго консистентна.

З доведення теореми 9 бачимо, що $\overline{H}\xrightarrow{\mathsf{P1}}H_\infty$, тому залишається перевірити, що

$$\hat U\xrightarrow{\mathsf{P1}}U.$$

Позначимо $\Delta\hat\beta=\hat\beta_A-\beta$; $\Delta\hat\beta\xrightarrow{\mathsf{P1}}0$. Тоді

$$\hat U=\overline{(h-H\beta)(h-H\beta)^{\mathsf T}}+\mathrm{rest},$$

$$\mathrm{rest}=\overline{H\Delta\hat\beta\cdot(\Delta\hat\beta)^{\mathsf T}H}-\overline{(h-H\beta)(H\Delta\hat\beta)^{\mathsf T}}-\overline{H\Delta\hat\beta(h-H\beta)^{\mathsf T}}.$$

Внаслідок строгої консистентності $\hat\beta_A$ виконується: $\mathrm{rest}\xrightarrow{\mathsf{P1}}0$. Справді, завдяки умовам (vi), (vii) послідовність усереднень $\{\overline{x^{4k}},n\ge1\}$ обмежена м.н., тому середнє $\overline{\|H\|^2}$ м.н. залишається обмеженим при зростанні обсягу вибірки, і перший доданок залишку rest прямує до нуля м.н. Така ж збіжність 2-го та 3-го доданків випливає з нерівності Коші–Шварца, оскільки далі буде показано, що м.н. є обмеженою послідовність усереднень

$$\overline{\|h-H\beta\|^2}. \tag{27}$$

Тому залишилось перевірити, що

$$\overline{(h-H\beta)(h-H\beta)^{\mathsf T}}\xrightarrow{\mathsf{P1}}U=\lim_{n\to\infty}\overline{\mathbf{cov}(h-H\beta)}, \tag{28}$$

або

$$\overline{(h-H\beta)(h-H\beta)^{\mathsf T}}-\mathsf{E}\,\overline{(h-H\beta)(h-H\beta)^{\mathsf T}}\xrightarrow{\mathsf{P1}}0. \tag{29}$$

(Зокрема, із збіжності (28) випливатиме, що усереднення (27) обмежені м.н..) Для виразу $z=h-H\beta$ маємо розклад (26). Доданки z містять вирази $\xi^i\delta^j$ при $i+j\le2k$, $i\le2k-1$, а також вирази $\xi^u\delta^v\varepsilon$ при $u+v\le k$. При розгляді $zz^{\mathsf T}$ з'являються наступні доданки.

а) Вирази $\xi^i\delta^j$ при $i+j \le 4k$, $i \le 4k-2$. Ми повинні забезпечити збіжність

$$\overline{\xi^i\delta^j} - \mathsf{E}\,\overline{\xi^i\delta^j} \xrightarrow{\mathsf{P1}} 0. \tag{30}$$

При $i=0$, $j=4k$ збіжність (30) справедлива, бо $\mathsf{E}\,\delta^{4k} < \infty$ за умовою (vi). Нехай тепер $j \le 4k-1$. Маємо $\overline{|\xi|^{ri}} \le \text{const}$, при деякому $r>1$ завдяки умові (vii); також $\mathsf{E}\,|\delta^j|^q < \infty$, $q = \frac{4k}{j} > 1$. Збіжність (30) справджується за лемою 3.

б) Вирази $\xi^i\delta^j\varepsilon^2$ при $i+j \le 2k$. Ми повинні забезпечити збіжність

$$\overline{\xi^i\delta^j\varepsilon^2} - \mathsf{E}\,\overline{\xi^i\delta^j\varepsilon^2} \xrightarrow{\mathsf{P1}} 0. \tag{31}$$

При $i=0$, $j \le 2k$ маємо $\overline{\delta^j\varepsilon^2} - \mathsf{E}\,\delta^j\varepsilon^2 \xrightarrow{\mathsf{P1}} 0$, бо $\mathsf{E}\,|\delta^j\varepsilon^2| < \infty$ за умовою (vi). При $1 \le i \le 2k$, $j \le 2k-1$ маємо $\overline{|\xi|^{ri}} \le \text{const}$ при деякому $r>1$ завдяки умові (vii); $\mathsf{E}\,|\delta^j\varepsilon^2|^{1+c} < \infty$, де число c взято з умови (vi). Збіжність (31) справджується за лемою 3.

в) Вирази $\xi^i\delta^j\varepsilon$ при $i \le 3k-1$, $i+j \le 3k$. Треба забезпечити збіжність

$$\overline{\xi^i\delta^j\varepsilon} - \mathsf{E}\,\overline{\xi^i\delta^j\varepsilon} \xrightarrow{\mathsf{P1}} 0. \tag{32}$$

При $i=0$, $j \le 3k$ маємо $\overline{\delta^j\varepsilon} - \mathsf{E}\,\delta^j\varepsilon \xrightarrow{\mathsf{P1}} 0$, бо $\mathsf{E}\,|\delta^j\varepsilon| < \infty$ за умовою (vi). При $1 \le i \le 3k-1$, $j \le 3k-1$ маємо $\overline{|\xi|^{ri}} \le \text{const}$ при деякому $r>1$ завдяки умові (vii); $\mathsf{E}\,|\delta^j\varepsilon|^{1+c} < \infty$, де число c взято з умови (vi). Збіжність (32) має місце за лемою 3.

Розглянуті випадки встановлюють збіжність (29), що і доводить твердження теореми. $\square$

Зазначимо, що в статті [5] побудовано модифіковані оцінки $\hat\beta_M$, $\hat\sigma^2_{\varepsilon,M}$ для нашої поліноміальної регресії. Вони стійкіші за ALS-оцінки для випадку малої вибірки, бо $\hat\beta_M$ обчислюється подібно до (11), проте матриця $\overline{H}$ (яка може бути близькою до сингулярної) заміняється на певну додатно визначену матрицю; крім того, оцінка $\hat\sigma^2_{\varepsilon,M}$ завжди додатна, на відміну від $\hat\sigma^2_{\varepsilon,A}$. За умови невиродженості коваріаційної матриці Ω з (1) та граничної матриці H_∞ з (iv), оцінка $\hat\beta_M$ асимптотично еквівалентна $\hat\beta_A$, тобто $\sqrt{n}(\hat\beta_M - \hat\beta_A) \xrightarrow{\mathsf{P}} 0$ (див. теорему 2 з роботи [5]). Тоді за теоремою Слуцького [1] теорема 9 залишається справедливою для $\hat\beta_M$ за умови невиродженості Ω. За цієї умови справедливий аналог теореми 10, в якому у виразі для $\hat U$ ми заміняємо оцінку $\hat\beta_A$ на $\hat\beta_M$.

5. Висновки

У поліноміальній функціональній моделі з похибками вимірювання ми отримали слабкіші умови консистентності ALS-оцінок, ніж в роботі [4], а також навели умови асимптотичної нормальності оцінки $\hat\beta_A$, чого не було зроблено у згаданій роботі. Для малої та середньої вибірки краще користуватись модифікованими оцінками $\hat\beta_M$, $\hat\sigma^2_{\varepsilon,M}$ з роботи [5], причому $\hat\beta_M$ має таку ж асимптотичну коваріаційну матрицю, що і $\hat\beta_A$; оцінка $\hat\beta_M$ асимптотично нормальна за тих же умов, що і $\hat\beta_A$, якщо матриця (1) невироджена.

У подальшому ми плануємо з'ясувати, чи будуть асимптотично еквівалентними оцінки $\hat\sigma^2_{\varepsilon,A}$ та $\hat\sigma^2_{\varepsilon,M}$. Цікаво також розглянути структурну нормальну модель (2), в якій ξ_i, ε_i, δ_i нормально розподілені. У такій моделі, звичайно, також працює ALS-оцінка, але можна побудувати і квазівірогідну оцінку $\hat\beta_Q$. Хотілось би порівняти асимптотичну ефективність оцінок $\hat\beta_A$ та $\hat\beta_Q$. Для некорельованих похибок ($\sigma_{\delta\varepsilon}=0$) це зроблено у [9]; крім того, див. огляд [10], де згадуються також інші моделі регресії з позибками вимірювання. У роботі [8] встановлено асимптотичну оптимальність

ALS-оцінки в лінійній функціональній моделі з похибками вимірювання; цікаво було б отримати схожий результат для поліноміальної моделі.

Література

1. П. Биллингсли, *Сходимость вероятностных мер*, "Наука", Москва, 1977.
2. В. С. Королюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин, *Справочник по теории вероятности и математической статистике*, "Наука", Москва, 1985.
3. В. В. Петров, *Предельные теоремы для сумм независимых случайных величин*, "Наука", Москва, 1987.
4. C.-L. Cheng and H. Schneeweiss, *Polynomial regression with errors in the variables*, Journal of the Royal Statistical Society B **60** (1998), no. 1, 189–199.
5. C.-L. Cheng, H. Schneeweiss, and M. Thamerus, *A small sample estimator for a polynomial regression with errors in the variables*, Journal of the Royal Statistical Society **62** (2000), no. 4, 699–709.
6. C.-L. Cheng and A. Kukush, *Goodness-of-fit test in a polynomial errors-in-variables model*, Ukrainian Mathematical Journal **56** (2004), no. 4, 641-661.
7. P. Hall and Y. Ma, *Testing the suitability of polynomial models in errors-in-variables problems*, The Annals of Statistics **35** (2007), no. 6, 2620–2638.
8. A. Kukush and E.-O. Maschke, *The efficiency of adjusted least squares in the linear functional relationship*, Journal of Multivariate Analysis **87** (2003), no. 2, 261–274.
9. A. Kukush, A. Malenko, and H. Schneeweiss, *Optimality of the quasi-score estimator in a mean-variance model with applications to measurement error models*, Journal of Statistical Planning and Inference **139** (2009), no. 10, 3461–3472.
10. H. Schneeweiss and A. Kukush, *Comparing the efficiency of structural and functional methods in measurement error models*, Theory Probab. Math. Statist. **80** (2010), 131–142.
11. I. O. Sen'ko, *Consistency of an adjusted least-squares estimator in a vector linear model with measurement errors*, Ukrainian Mathematical Journal **64** (2013), no. 11, 1739–1751.
12. I. O. Sen'ko, *The asymptotic normality of an adjusted least squares estimator in a multivariate vector errors-in-variables regression model*, Theory Probab. Math. Statist. **88** (2014), 175–190.

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

Адреса електронної пошти: alexander_kukush@univ.kiev.ua, 777Tsar777@mail.ru

Надійшла 23/12/2014