$h_2 : [a,b] \to [c,d]$ такий, що $h_2(t) = c + \frac{t-a}{b-a}(d-c)$ і який зберігає орієнтацію. Із описаного вище випливає, що $f = h_2^{-1} \circ g \circ \tilde{\varphi}_2^{-1} \circ \tilde{\varphi}_1$ та $S_r^1 = \tilde{\varphi}_1^{-1}[\tilde{\varphi}_2(g^{-1}(h_2(f(S_r^1))))], r = \overline{1,n}$, згідно до побудови інваріанту. Теорему доведено.

Нагадаємо, що для коректності задачі необхідне співпадання порядків обходу на букетах.

Узагальнюючи букет із декількох кіл до випадку, коли з'являються сингулярні точки (точки перетину кіл), для прикладу розглянемо такі топологічні простори, як на рис. 2

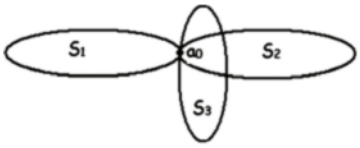

Рис. 2

Неважко переконатись, що побудований вище інваріант також підходить для класифікації функцій, заданих на таких топологічних просторах, потрібно лише додати декілька умов, а саме: 1) кількість критичних точок функцій на таких колах між відповідними сингулярними точками має співпадати, а також, 2) звуження функцій у сингулярних точках мають співпадати (аби функції були неперервними).

4. Висновок

Для букета із *n* кіл, побудовано комбінаторний інваріант $Sp^n(f)$ для неперервної функції f, заданої на цьому букеті, яка має скінчену кількість точок локального екстремуму. За допомогою цього інваріанту виведено критерій *PPO*-еквівалентності неперервних функцій із скінченим числом локальних екстремумів, що задані на букеті з n кіл. Він зводиться до перевірки порядку обходу кіл у заданих топологічних просторах і перевірки на збіжність комбінаторних інваріантів $Sp^n(f)$ та $Sp^n(g)$ функцій $f : S \to \mathrm{R}$ та $g : S \to \mathrm{R}$, ($S = \bigcup_{r=1}^{s} S_r^1$), тобто перевірки на збіжність узагальнених змій, що відповідають цим функціям, а саме $G^f(2m-1, k-1)$ із $G^g(2m-1, k-1)$. Випадок букету кіл узагальнено до випадку, коли з'являються сингулярні точки і доведено, що даний інваріант можна використовувати і для класифікації функцій, заданих на таких узагальнених топологічних просторах.

1 *Андріюк О.П.* Функції на одновимірних многовидах: автореф. дис. на здобуття наук. ступеня канд. фіз.-мат. наук: спец.01.01.01 "Математичний аналіз" / О.П. Андріюк. – Київ,2006. – 19с. 2 *Арнольд В.И.* Исчисление змей и комбинаторика чисел Бернулли, Эйлера и Спрингера групп Кокстера / В.И. Арнольд // Успехи мат.наук. – 1992. – Т.47, №1(283). – С.3-45 3 *Шарко В.В.* Гладкая и топологическая эквивалентность функций на поверхностях / В.В. Шарко // Укр.мат.жур – 2003. – Т.55, №5 – С.687-700 4 *Юрчук І.А.* Комбінаторні аспекти топологічної класифікацій функцій на колі / Юрчук І.А. / Укр..мат.жур.– 2008.–Т60, №6. – С.829-836.

Надійшла до редколегії 30.11.09

УДК 519.21

Г. Репетацька

ПОКРАЩЕНА ОЦІНКА ОРТОГОНАЛЬНОЇ РЕГРЕСІЇ ДЛЯ НЕЯВНОЇ ФУНКЦІОНАЛЬНОЇ МОДЕЛІ З ПОХИБКАМИ В ЗМІННИХ

Для неявної функціональної моделі регресії з похибками в змінних розглядається оцінка ортогональної регресії. Встановлено умови неконзистентності, знайдено перший член розкладу асимптотичного відхилення оцінки по дисперсії похибок, та запропонована покращена оцінка, яка має менше асимптотичне відхилення. Отримані результати розглянуті на прикладі еліпса.

For implicit functional errors-in-variables model the orthogonal regression estimator is considered. The conditions for the inconsistency are presented, and the first term of the asymptotic expansion of its deviation is derived. An improved estimator is proposed, that has smaller asymptotic deviation from the true value. As an example the estimation of the ellipse is considered.

1. Вступ

Розглянемо неявну функціональну модель регресії з похибками у змінних

$$\begin{cases} G(\xi_i, \beta_0) = 0, \\ x_i = \xi_i + \varepsilon_i, \ i = \overline{1,n}, \end{cases} \tag{1}$$

де $G(\xi, \beta)$ – скалярна функція від векторних змінних $\xi \in \mathbb{R}^q$, $q \ge 2$ та $\beta \in \Theta \subset \mathbb{R}^p$, β_0 – параметр, що оцінюється, x_i – дані спостережень, ξ_i – невідомі детерміновані величини, ε_i – похибки спостережень.

Для моделі (1) розглянемо оцінку ортогональної регресії (о.о.р.) $\hat{\beta}_n$, встановимо, що вона неконзистентна, та запропонуємо покращену оцінку, яка має менше асимптотичне відхилення при малих похибках та великих n. Раніше аналогічні результати були отримані в [1,3] для явних моделей регресії.

Відомо, що для лінійної функції регресії о.о.р. консистентна (див. [4]), та при нормально розподілених похибках з однаковою дисперсією є оцінкою максимальної правдоподібності. Але в роботі [3] доведено, що для явної нелінійної моделі регресії о.о.р. неконзистентна і навіть відділена від істинного значення β_0. Також в [3] знайдено перший член розкладу її асимптотичного відхилення по дисперсії похибок та запропоновано покращену оцінку, яка має менше асимптотичне відхилення від β_0. Згодом в [1] ці результати були поширені з двовимірної моделі на багатовимірну. В даній статті ми поширюємо отримані результати на неявну модель регресії, що дає змогу використовувати покращену оцінку в оцінюванні параметрів неявних моделей, наприклад, для оцінювання еліпсоїда за спостереженнями з адитивною похибкою точок на його поверхні.

Перехід до неявної моделі вимагає постановки умов, адаптованих до неявної моделі, а також змін в обчисленні асимптотичного відхилення. Але деякі теореми не потребують змін, оскільки при нових умовах хід доведення не змінюється. Для таких теорем наводимо тільки формулювання з посиланням на [1,3,6].

2. Модель та умови

Надалі I_n позначає одиничну матрицю n-го порядку, $\|\cdot\|$ – евклідову норму. Векторні величини будемо записувати векторами-стовпцями, а похідні – векторами-рядками; змінну, по якій береться похідна – верхнім індексом, наприклад, G^{ξ} – це вектор-рядок частинних похідних за компонентами ξ; $\lambda_{\min}(V)$ позначає найменше власне значення симетричної матриці V; $U_r(x_0)$ – відкрита куля з центром x_0 і радіусом r, $\bar{U}_r(x_0)$ – її замикання.

Означимо також множину $\Gamma_\beta := \{u \in \mathbb{R}^q : G(u,\beta) = 0\}$, що є графіком неявно заданої функції G з параметром β. Будемо позначати через ξ_β довільну точку графіка Γ_β.

Нехай для моделі (1) виконуються наступні умови:

(i) $\beta_0 \in \operatorname{int}\Theta$, $\Theta \subset \mathbb{R}^p$ – деякий компакт; $\Gamma_\beta \neq \varnothing$ для всіх $\beta \in \Theta$, та $\Gamma_{\beta_0} = \partial A_{\beta_0}$, де $A_{\beta_0} \in \mathbb{R}^q$ – відкрита однозв'язна множина.

(ii) $\|\xi_i\| \le A,\ i \ge 1,$ де A – деяка стала.

(iii) $G(\xi,\beta) \in C^3(\mathbb{R}^q \times U)$ – тричі неперервно диференційовна функція, де $U \supset \Theta$ – відкрита, та $G^{\xi}(\xi_\beta,\beta) \neq 0$ для довільних $\beta \in \Theta,\ \ \xi_\beta$.

(iv) Похибки $\{\varepsilon_i, i \ge 1\}$ незалежні однаково розподілені, $\varepsilon_i \sim N(\vec{0},\sigma^2 I_q)$, де $\sigma > 0$ – невідомий параметр.

Зауваження: умова нормальності похибок не принципова, але спрощує доведення.

Для будь-якої функції $F(\xi,\beta)$ означимо $F_n := \frac{1}{n}\sum_{i=1}^{n} F(\xi_i,\beta_0)$ – середнє значення функції в дійсних точках спостереження. Означимо величини k_n, V_n для наступних функцій:

$$k(\xi,\beta) = \operatorname{tr}\left[G^{\xi\xi}\left(I_q - \frac{G^{\xi\mathrm{T}} G^{\xi}}{\|G^{\xi}\|^2}\right)\right]\frac{G^{\beta}}{\|G^{\xi}\|^2}, \quad V(\xi,\beta) = \frac{(G^{\beta})^{\mathrm{T}} G^{\beta}}{\|G^{\xi}\|^2}.$$

Для виконання подальших теорем 2-4 суттєві умови відділеності k_n, V_n від нуля:

(v) $$\varliminf_{n\to\infty} \|k_n\| > 0;$$

(vi) $\varliminf_{n\to\infty} \lambda_{\min}(V_n) > 0$.

Надалі будемо вимагати виконання умов *(i)-(iii)*.

Важливий наслідок з *(ii), (iii)* : якщо $K \subset \mathbb{R}^q$ – компакт, то для всіх (ξ_β,β) таких, що $\beta \in \Theta$, $\xi_\beta \in K$, похідна $G^{\xi}(\xi_\beta,\beta)$ відділена від нуля: $\|G^{\xi}(\xi_\beta,\beta)\| \ge \delta_0$, де $\delta_0 > 0$ – деяка стала.

Також виконання цих умов забезпечує обмеженість k_n, V_n зверху.

3. Оцінка ортогональної регресії

Нехай $q(x,\beta) := \min_{u\in\Gamma_\beta} \|x-u\|^2$ – квадрат найменшої відстані від точки x до графіка Γ_β;

$$Q(\beta) := \frac{1}{n}\sum_{i=1}^{n} q(x_i,\beta) \text{ – цільова функція.}$$

Означимо оцінку ортогональної регресії $\hat\beta$ як борелеву функцію від спостережень x_i, $i = \overline{1,n}$, в якій цільова функція досягає найменшого значення:

$$\hat\beta \in \arg\min_{\beta\in\Theta} Q(\beta). \tag{2}$$

Для довільного набору $\{x_i\}_{i=1}^n$ точка, в якій $Q(\beta)$ досягає мінімуму, може бути не єдина, але завжди можна вибрати таку, щоб функція $\hat\beta = \hat\beta\left(\{x_i\}_{i=1}^n\right)$ була борелевою (див. [5]). Нехай $h(x,\beta) \in \arg\min_{u\in\Gamma_\beta} \|x-u\|^2$ – точка найменшої відстані від x до графіка Γ_β. Розглянемо питання існування та єдиності $h(x,\beta)$.

Існування. Для довільних $\beta \in \Theta$ та $x \in \mathbb{R}^q$ завжди існує хоча б одна точка з Γ_β, найближча до x, внаслідок замкненості множини Γ_β. Позначимо через $h(x,\beta)$ будь-яку з таких точок.

Єдиність. Оскільки функція G є неперервно диференційою, то $h(x,\beta)$ є точкою мінімімуму функції Лагранжа за змінною u:

$$\begin{cases} L(u;x,\beta) = \| x-u \|^2 + 2a(x,\beta)G(u,\beta) \to \min, \\ G(u,\beta)=0, \end{cases}$$

де $a = a(x,\beta)$ – відповідний скалярний множник, а h та a задовільняють систему рівнянь

$$\begin{cases} L^u(u;x,\beta)\big|_{u=h} = 0 \\ G(h,\beta)=0 \end{cases}, \quad \text{àáî} \quad \begin{cases} x-h-a(G^u(u,\beta)|_{u=h})^{\mathrm{T}} = 0 \\ G(h,\beta)=0 \end{cases}. \tag{3}$$

Система рівнянь (3) визначає h як основу перпендикуляра, опущеного з x на Γ_β: $(x-h) \,\|\, (G^u(h,\beta))^{\mathrm{T}}$, де $h \in \Gamma_\beta$; a – коефіцієнт пропорційності між перпендикуляром і вектором нормалі $(G^u)^{\mathrm{T}}$. Система (3) неявно задає векторну функцію $\alpha(x,\beta) := \begin{pmatrix} h(x,\beta) \\ a(x,\beta) \end{pmatrix}$ як розв'язок векторного рівняння $F(\alpha;x,\beta) = \vec{0}$, де $F = \left(L^u;G\right)^{\mathrm{T}}$ – $(q+1)$-вимірна функція.

Лема 1. *Нехай для моделі (1) виконуються умови (i)-(iii). Тоді векторна функція* $h(x,\beta)$ *та скалярна* $a(x,\beta)$ *однозначно визначаються з системи (3) і є двічі неперервно диференційовними в околі кожної з точок* (ξ_β,β)*, таких що* $\beta \in \Theta$ *та* $\|\xi_\beta\| \le A$*, причому радіус околу можна вибрати єдиним для всіх точок:*

$$\begin{pmatrix} h \\ a \end{pmatrix} : \bar{U}_{\nu_0}(\xi_\beta) \times \bar{U}_{\nu_0}(\beta) \to \mathbb{R}^{q+1}.$$

Доведення. Легко бачити, що $F(\alpha(\xi_\beta,\beta);\xi_\beta,\beta) = \vec{0}$, $F^\alpha\left(\alpha(\xi_\beta,\beta);\xi_\beta,\beta\right) = \begin{pmatrix} -I_q - aG^{\xi\xi}(\xi_\beta,\beta) & G^{\xi\mathrm{T}}(\xi_\beta,\beta) \\ G^\xi(\xi_\beta,\beta) & 0 \end{pmatrix}$.

В деякому околі $U_{\nu(\xi_\beta,\beta)}(\xi_\beta,\beta) := U_{\nu(\xi,\beta)}(\xi_\beta) \times U_{\nu(\xi_\beta,\beta)}(\beta)$ точки (ξ_β,β), матриця F^α буде невиродженою. Обґрунтуємо це. З (3) та умов (ii), (iii) випливає, що $\| a G^\xi(h,\beta) \| = \| x - h(x,\beta) \| \le \| x-\xi_\beta \|$, а з відділеності $\| G^\xi \|$ від нуля,

$$| a(x,\beta) | \le \delta_0^{-1} \cdot \| x-\xi_\beta \|. \tag{4}$$

З врахуванням (4) та означення визначника як алгебраїчної суми добутків елементів в різних рядках і ствопчиках,

$$\det(-F^\alpha) = \det\begin{pmatrix} I_q & -G^{\xi\mathrm{T}} \\ -G^\xi & 0 \end{pmatrix} + O(a) = -\| G^\xi \|^2 + O\left(\| x-\xi_\beta \|\right) \le -\delta_0^2 + O\left(\| x-\xi_\beta \|\right) -$$

визначник відділений від нуля, при x близьких до ξ_β. *(Легко переконатись, що визначник в останньому рядку дорівнює* $-\| G^\xi \|^2$*, якщо розкласти його за першим стовпчиком і далі знайти другий доданок, розклавши його за першим рядком.)* Тоді, за теоремою про неявну функцію, в деякому околі $U_{\nu(\xi_\beta,\beta)}(\xi_\beta,\beta)$ будуть існувати: однозначна двічі неперервно диференційовна векторна функція $h(x,\beta)$ та скалярна функція $a(x,\beta)$. Внаслідок компактності множини всіх можливих значень (ξ_β,β), можна вибрати єдине значення ν_0 замість усіх можливих $\nu(\xi_\beta,\beta)$. Для зручності в подальших дослідженнях, ν_0 вибираємо меншим за максимально можливе, для того щоб неявно задані функції були визначені і диференційовні в **замкненому** околі точок спостереження.

Лема доведена.

Зауваження: Зокрема, ξ_{β_0} може бути будь-якою з точок ξ_i, $i = \overline{1,n}$, тобто неявно задана функція однозначна та диференційовна в околі істинних точок спостережень.

Дана лема дозволяє стверджувати, що для всіх x_i, відхилення яких від істинного значення ξ_i менші за нормою від ν_0, та $\beta \in U_{\nu_0}(\beta_0)$, існує тільки один перпендикуляр, опущений з x на Γ_β. Ми розглядаємо модель при малих σ^2, але якщо відхилення буде більшим за ν_0, перпендикуляр h може бути не єдиним. Тому розділимо спостереження на дві частини. Виділимо в спостереженнях множину індексів B_n, для елементів якої буде виконуватись лема 1: $B_n := \left\{ i = \overline{1,n} : \| \varepsilon_i \| < \nu_0 \right\}$. Цільову функцію розділимо на дві частини:

$$Q(\beta) = Q_1(\beta) + Q_2(\beta) := \tfrac{1}{n}\textstyle\sum_{i \in B_n} q(x_i,\beta) + \tfrac{1}{n}\sum_{i \notin B_n} q(x_i,\beta),$$

де $Q_1(\beta)$ буде основною частиною, тричі неперервно диференційовною та обмеженою, а $Q_2(\beta)$ будемо розглядати як залишок.

4. Неконзистентність оцінки ортогональної регресії

Наступні позначення будемо використовувати для послідовностей випадкових векторів, що залежать від параметрів $\beta\in\Theta$ та $\sigma>0$:

Означення 1. Послідовність $\eta_n=\eta_n(\beta,\sigma)$ називається рівномірною стохастично обмеженою, якщо $\lim\limits_{c\to\infty}\sup\limits_{n\ge1;\beta\in\Theta;\sigma>0}\mathrm{P}\left(\|\eta_n(\beta,\sigma)\|>c\right)=0$. Позначення: $\eta_n=O_{\mathrm{P}}(1)$.

Означення 2. $\eta_n(\beta)=o_{\mathrm{P}}(1)$, якщо для довільних $\beta\in\Theta$, $\eta_n(\beta)\to0,\ n\to\infty$, за ймовірністю.

Означення 3. $\eta_n(\beta,\sigma)=o_{\sigma\mathrm{P}}(1)$, якщо для всіх $c>0$ $\lim\limits_{\sigma\to0+}\sup\limits_{n\ge1}\mathrm{P}\left(\sup\limits_{\beta\in\Theta}\|\eta_n(\beta,\sigma)\|>c\right)=0$.

У наступній теоремі отримано асимптотичні розклади $Q(\beta)$ та похідних $Q_1(\beta)$ в точці β_0, які використовуються для встановлення умов неконзистентності та знаходження асимптотичного відхилення.

Теорема 1. *Нехай для моделі (1) виконуються умови (i)–(iv). Тоді*

$$Q(\beta)=Q_1(\beta)+\sigma^4 o_{\sigma P}(1), \tag{5}$$

$$Q(\beta_0)=\sigma^2+\left(\sigma^3+\frac{\sigma^2}{\sqrt{n}}\right)O_{\mathrm{P}}(1)+\sigma^4 o_{\sigma\mathrm{P}}(1), \tag{6}$$

$$Q_1^{\beta}(\beta_0)=\sigma^2 k_n+\left(\sigma^3+\frac{\sigma}{\sqrt{n}}\right)O_{\mathrm{P}}(1)+\sigma^4 o_{\sigma\mathrm{P}}(1), \tag{7}$$

$$Q_1^{\beta\beta}(\beta_0)=2V_n+\sigma O_{\mathrm{P}}(1)+\sigma^4 o_{\sigma\mathrm{P}}(1). \tag{8}$$

Наступна лема використовується при доведенні теореми 1 та подальшої леми 4.

Лема 2. *Нехай $\{a_i : i\ge1\}$ – обмежена числова послідовність, ζ_i – незалежні однаково розподілені випадкові вектори зі скінченними другими моментами. Тоді*

$$\frac{1}{n}\sum_{i=1}^{n}a_i\zeta_i=\frac{\mathrm{E}\,\zeta_1}{n}\sum_{i=1}^{n}a_i+\frac{\sqrt{\mathrm{D}\,\zeta_1}}{\sqrt{n}}O_{\mathrm{P}}(1).$$

Лема легко доводиться за нерівністю Чебишова.

Доведення теореми 1. Для величин ξ_i, x_i, ε_i, $h_i:=h(x_i,\beta_0)$ та аналогічних величин, якщо розглядається випадок конкретного i, індекс i будемо упускати. Наприклад, замість $q(x_i,\beta)$ писатимемо $q(x,\beta)$.

Доведення формули (5). Доведемо, що $Q_2(\beta)=\sigma^4 o_{\sigma\mathrm{P}}(1)$. Розглянемо $q(x,\beta)$. Враховуючи, що $h(\xi,\beta)\in\Gamma_\beta$, маємо

$$q(x,\beta)=\|x-h(x,\beta)\|^2\le\|x-h(\xi,\beta)\|^2\le\left(\|x-\xi\|+\|\xi-h(\xi,\beta)\|\right)^2\le2\|\varepsilon\|^2+D,$$

де D – стала для даної моделі. Тут ми використали компактність Θ, неперервність G та умову (ii).

Далі, $Q_2(\beta)=\frac{1}{n}\sum_{i\notin B_n}q(x_i,\beta)\le\frac{1}{n}\sum_{i=1}^{n}\left(2\|\varepsilon_i\|^2+D\right)I_{\|\varepsilon_i\|\ge\nu_0}$. Нехай $\tilde\varepsilon_i:=\varepsilon_i/\sigma$, $\tilde\varepsilon_i\sim N(\vec{0};I_q)$ – величини зі стандартним нормальним розподілом. Оцінимо математичні сподівання доданків:

$$\mathrm{E}\|\varepsilon\|^2 I_{\|\varepsilon\|\ge\nu_0}=\frac{\sigma^4}{\nu_0^2}\mathrm{E}\frac{\nu_0^2}{\sigma^2}\|\tilde\varepsilon\|^2 I_{\|\tilde\varepsilon\|\ge\nu_0^2/\sigma^2}\le\frac{\sigma^4}{\nu_0^2}\mathrm{E}\|\tilde\varepsilon\|^4 I_{\|\tilde\varepsilon\|\ge\nu_0/\sigma}=\sigma^4 o(1),\ \sigma\to0+.$$

$$\mathrm{E}\,I_{\|\varepsilon\|\ge\nu_0}=\frac{\sigma^4}{\nu_0^4}\mathrm{E}\frac{\nu_0^4}{\sigma^4}I_{\|\tilde\varepsilon\|\ge\nu_0/\sigma}=\frac{\sigma^4}{\nu_0^4}\mathrm{E}\|\tilde\varepsilon\|^4 I_{\|\tilde\varepsilon\|\ge\nu_0/\sigma}=\sigma^4 o(1),\ \sigma\to0+.$$

Отже, $\mathrm{E}\,Q_2(\beta)=\sigma^4 o(1),\ \sigma\to0+$. За нерівністю Чебишова,

$\mathrm{P}(\sigma^{-4}Q_2(\beta)>C)\le\dfrac{\mathrm{E}\,\sigma^{-4}Q_2(\beta)}{C}=\dfrac{o(1)}{C}\to0,\ \sigma\to0+$. Це означає, що $Q_2(\beta)=\sigma^4 o_{\sigma\mathrm{P}}(1)$. Формула доведена.

Знаходження основних доданків формул (6)-(8).

Для отримання решти формул будемо окремо розглядати кожен доданок $q(x_i,\beta)$ цільової функції $Q(\beta)$ та її похідних в точці β_0. Доданки, для яких $i\notin B_n$, будуть становити достатньо малу частину загальної суми при малих σ, не більшу за $\sigma^4 o_{\sigma P}(1)$, що доводиться аналогічно до попереднього пункту.

Розглянемо фіксоване $i\in B_n$. Тоді ε_i – обмежена випадкова величина, та існує однозначна двічі неперервно диференційовна функція $h_i:=h(\xi_i+\varepsilon_i;\beta)$ та $q(x_i,\beta)=\|x_i-h_i\|^2$, де ξ_i – фіксований параметр. Знайдемо розклад за формулою Тейлора функції q (та її похідних) відносно ε, та математичні сподівання членів розкладу.

Означимо величину Δ з рівності $h=\xi+\Delta$.

Зауважимо, що $\|\Delta\|^2=\|\xi-h\|^2\le\|\xi-x\|^2+\|x-h\|^2\le2\|\xi-x\|^2=2\|\varepsilon\|^2$. Отже, $\Delta=O(\|\varepsilon\|),\varepsilon\to0$. Тоді з першого рівняння (3) також випливає, що $a=O(\|\varepsilon\|),\varepsilon\to0$.

Для будь-якої величини P, залежної від похибки ε, через P_1, P_2 позначаємо лінійну та квадратичну частини розкладу P за формулою Тейлора в околі $\vec{0}$. Знайдемо розклад Δ та a за степенями ε у вигляді $\Delta = \Delta_1 + \Delta_2 + O(\|\varepsilon\|^3)$, $a = a_1 + a_2 + O(\|\varepsilon\|^3)$. Для цього функції в системі (3) розкладемо за формулою Тейлора в околі точки $(\xi;\beta_0)$. Маємо

$$\begin{cases} \varepsilon - \Delta = G^{\xi\mathrm{T}}(\xi,\beta_0)a + G^{\xi\xi}(\xi,\beta_0)\Delta\cdot a + O(\|\varepsilon\|^3) \\ G^{\xi}(\xi,\beta_0)\Delta + \frac{1}{2}\Delta^{\mathrm{T}}G^{\xi\xi}(\xi,\beta_0)\Delta = O(\|\varepsilon\|^3) \end{cases}$$

Через $O(\cdot)$ позначаємо рівномірно обмежену величину: $\|O(t)\| \le C\|t\|$, для довільних β,ξ,σ,n,i.

Значення функцій при фіксованих ξ,β_0 будемо записувати без аргумента. Далі отримуємо

$$\begin{cases} \varepsilon - \Delta_1 = G^{\xi\mathrm{T}}a_1 \\ G^{\xi}\Delta_1 = 0 \end{cases}, \text{ звідки } \begin{cases} \Delta_1 = \varepsilon - G^{\xi\mathrm{T}}a_1 \\ a_1 = G^{\xi}\cdot\varepsilon/\|G^{\xi}\|^2 \end{cases}, \text{ та } \begin{cases} -\Delta_2 = G^{\xi\mathrm{T}}a_2 + G^{\xi\xi}\Delta_1 a_1 \\ G^{\xi}\Delta_2 + \frac{1}{2}\Delta_1^{\mathrm{T}}G^{\xi\xi}\Delta_1 = 0 \end{cases}, \text{ звідки } \begin{cases} \Delta_2 = -\left(G^{\xi\mathrm{T}}a_2 + G^{\xi\xi}\Delta_1 a_1\right) \\ a_2 = \dfrac{\frac{1}{2}\Delta_1^{\mathrm{T}}G^{\xi\xi}\Delta_1 - a_1 G^{\xi}G^{\xi\xi}\Delta_1}{\|G^{\xi}\|^2} \end{cases}.$$

Знайдемо розклад $q(x;\beta_0)$ за степенями ε та математичне сподівання його доданків. Маємо:

$$q = \|x-h\|^2 = \|\varepsilon-\Delta\|^2 = \|\varepsilon-\Delta_1\|^2 + O(\|\varepsilon\|^3) = a_1^2\|G^{\xi}\|^2 + O(\|\varepsilon\|^3);$$

$$q_0 = q_1 = 0;\ \ \mathrm{E}\,a_1^2 = \frac{G^{\xi}\varepsilon\cdot\varepsilon^{\mathrm{T}}G^{\xi\mathrm{T}}}{\|G^{\xi}\|^4} = \frac{\sigma^2}{\|G^{\xi}\|^2};\ \ \mathrm{E}\,q_2 = \|G^{\xi}\|^2\cdot\mathrm{E}\,a_1^2 = \|G^{\xi}\|^2\cdot\frac{\sigma^2}{\|G^{\xi}\|^2} = \sigma^2.$$

Далі знаходимо розклади за степенями ε величин $q^{\beta}(x,\beta_0)$ та $q^{\beta\beta}(x,\beta_0)$. Спочатку знайдемо похідні:

$$q^{\beta}(x;\beta) = \frac{\partial}{\partial\beta}L\left(h(x,\beta);x;\beta\right) = L^{u}h^{\beta} + L^{\beta} = L^{\beta} = 2aG^{\beta}(h,\beta) + 2a^{\beta}G(h,\beta) = 2a\,G^{\beta}(h,\beta);$$

$$q^{\beta\beta}(x;\beta) = \frac{\partial}{\partial\beta}\left(2G^{\beta\mathrm{T}}(h,\beta)\cdot a\right) = 2G^{\beta\mathrm{T}}(h,\beta)a^{\beta} + 2a\left(G^{\beta\beta}(h,\beta) + G^{\beta\xi}(h,\beta)h^{\beta}\right).$$

Знайдемо $q_0^{\beta\beta}$. Для того щоб знайти a^{β}, продиференціюємо (3) за β:

$$\begin{cases} -h^{\beta} = G^{\xi\mathrm{T}}a^{\beta} + a\left(G^{\xi\xi}h^{\beta} + G^{\xi\beta}\right) \\ G^{\xi}h^{\beta} + G^{\beta} = 0 \end{cases}, \text{ звідки знаходимо } a^{\beta} = \frac{G^{\beta} - aG^{\xi}(G^{\xi\xi}h^{\beta} + G^{\xi\beta})}{\|G^{\xi}\|^2}, \text{ та } q_0^{\beta\beta} = 2G^{\beta\mathrm{T}}a^{\beta} = \frac{G^{\beta\mathrm{T}}G^{\beta}}{\|G^{\xi}\|^2};$$

Знайдемо розклад $q^{\beta}(x;\beta_0)$ за степенями ε та математичне сподівання його доданків. Маємо:

$$q^{\beta} = 2a\,G^{\beta}(h,\beta_0) = 2\left(a_1 + a_2 + O(\|\varepsilon\|^3)\right)\left(G^{\beta} + G^{\beta\xi}\Delta_1 + O(\|\varepsilon\|^2)\right), \text{ звідки } q_0^{\beta} = 0,\ \ \mathrm{E}\,q_1^{\beta} = 0;\ \ q_2^{\beta} = 2a_1G^{\beta\xi}\Delta_1 + 2a_2G^{\beta}.$$

Знайдемо математичні сподівання компонент q_2^{β}: $\mathrm{E}\,a_1^2 = \dfrac{\sigma^2}{\|G^{\xi}\|^2}$; $\mathrm{E}\,a_1\Delta_1 = \vec{0}$; $\mathrm{E}\,\Delta_1\Delta_1^{\mathrm{T}} = \sigma^2\left(I_q - \dfrac{G^{\xi\mathrm{T}}G^{\xi}}{\|G^{\xi}\|^2}\right)$.

Звідси отримуємо $\mathrm{E}\,q_2^{\beta} = \dfrac{\sigma^2}{\|G^{\xi}\|^2}\mathrm{tr}\left[G^{\xi\xi}\left(I_q - \dfrac{G^{\xi\mathrm{T}}G^{\xi}}{\|G^{\xi}\|^2}\right)\right]G^{\beta}$.

Доведення (6)-(8). Ми знайшли основні величини, необхідні для доведення формул. Для прикладу розглянемо (7), а решта формул доводяться аналогічно.

Розглянемо $Q_1^{\beta}(\beta_0)$. Його розклад за Тейлором $Q_1^{\beta}(\beta_0) = \frac{1}{n}\sum_{i\in B_n} q(x_i,\beta_0) = A_1 + A_2 + \frac{1}{n}\sum_{i\in B_n} O(\|\varepsilon\|^3)$, де $A_k = \frac{1}{n}\sum_{i\in B_n} q_k^{\beta}(\varepsilon_i)$, $k = 1,2,$ – лінійна та квадратична частини.

Розглянемо доданки A_1. $q_1^{\beta}(\varepsilon)$ є лінійною формою від $\tilde{\varepsilon} = \varepsilon/\sigma$ з обмеженими коефіцієнтами для всіх $i = \overline{1,n}$: $q_1^{\beta} = 2a_1G^{\beta} = 2\sigma\left(G^{\xi}\tilde{\varepsilon}/\|G^{\xi}\|^2\right)G^{\beta}$. Отже, A_1 – сума лінійних форм від $\tilde{\varepsilon}_i$ з нульовими математичними сподіваннями. Виділимо основну частину та залишок: $A_1 = S_1 - S_2$, де $S_1 := \frac{\sigma}{n}\sum_{i=1}^{n} q_1^{\beta}(\tilde{\varepsilon}_i)$, $S_2 := \frac{\sigma}{n}\sum_{i\notin B_n} q_1^{\beta}(\tilde{\varepsilon}_i)$. За лемою 2 отримуємо $S_1 = \vec{0} + \sigma/\sqrt{n}\cdot O_{\mathrm{P}}(1)$. Те, що $S_2 = \sigma^4 o_{\sigma\mathrm{P}}(1)$, легко довести за допомогою міркувань, аналогічних (5). Отже, $A_1 = \sigma/\sqrt{n}\cdot O_{\mathrm{P}}(1) + \sigma^4 o_{\sigma\mathrm{P}}(1)$. Так само знаходимо $A_2 = \sigma^2 k_n + \sigma^2/\sqrt{n}\cdot O_{\mathrm{P}}(1) + \sigma^4 o_{\sigma\mathrm{P}}(1)$.

Нам залишається розглянути залишок $R := \frac{1}{n}\sum\limits_{i\in B_n} O(\|\varepsilon\|^3)$: маємо $\|R\| \le const\dfrac{\sigma^3}{n}\sum\limits_{i=1}^{n}\|\tilde{\varepsilon}_i\|^3 = \sigma^3 O_{\mathrm{P}}(1)$.

Формула (7) доведена. Аналогічно доводяться рівності (6) та (8). □

Сформулюємо теорему про неконзистентність, і навіть асимптотичну відокремленість оцінки $\hat{\beta}_n$ параметра β_0 від істинного значення, при відокремленому від нуля k_n:

Теорема 2. *Нехай для моделі (1) виконуються умови (i)-(v). Тоді для довільного* $\varepsilon > 0$ *знайдуться такі* $\tau > 0$ *та* $\sigma_\varepsilon > 0$ *, що для всіх* $\sigma \in (0;\sigma_\varepsilon]$

$$\varliminf_{n\to\infty} \mathrm{P}\left(\|\hat\beta_n - \beta_0\| > \sigma^2\tau\right) > 1-\varepsilon .$$

Доведення даної теореми грунтується на формулах (5), (7) і цілком аналогічне доведенню відповідного твердження в [1–3].

5. Асимптотичне відхилення

Для характеризації асимптотичної поведінки послідовності випадкових векторів будемо використовувати наступні позначення.

Означення 4. $\eta_n(\sigma) = \tilde O_{\sigma\mathrm{P}}(1)$ *, якщо* $\lim\limits_{\sigma\to 0+}\varlimsup\limits_{n\to\infty}\mathrm{P}\left(\|\eta_n(\sigma)\| > C\right)\to 0,\ C\to\infty$.

Означення 5. $\eta_n(\sigma) = \tilde o_{\sigma\mathrm{P}}(1)$ *, якщо* $\forall C > 0:\ \lim\limits_{\sigma\to 0+}\varlimsup\limits_{n\to\infty}\mathrm{P}\left(\|\eta_n(\sigma)\| > C\right) = 0$.

Зауваження: якщо $\eta_n(\sigma) = o_{\sigma\mathrm{P}}(1)$, то $\eta_n(\sigma) = \tilde o_{\sigma\mathrm{P}}(1)$ (обернене твердження хибне).

Нехай $\rho(x,\Gamma_\beta)$ – відстань від точки x до Γ_β . Сформулюємо умову контрастності, або асимптотичної відділеності точки β_0 :

$$\textit{(con)}\ \ \forall\delta>0:\ \varliminf_{n\to\infty}\ \inf_{\|\beta-\beta_0\|>\delta}\frac{1}{n}\sum_{i=1}^{n}\rho^2(\xi_i,\Gamma_\beta) > 0.$$

Дана умова забезпечує можливість існування консистентної оцінки при $\sigma\to 0$.

Лема 3. *Нехай для моделі (1)* $G\in C(\mathbb{R}^q,\Theta)$ *та виконується умова контрастності (con). Тоді* $\hat\beta_n - \beta_0 = \tilde o_{\sigma\mathrm{P}}(1)$.

Це твердження сформульоване в [3]. Наводимо його без доведення.

Наступна теорема, за якою знаходиться перший член розкладу асимптотичного відхилення оцінки по σ^2 , також не потребує окремого доведення (див. [1,3,6]).

Теорема 3. *Нехай для моделі (1) виконуються умови (i)-(vi) та (con). Тоді*

$$\hat\beta_n = \beta_0 - \frac{\sigma^2}{2}V_n^{-1}k_n^{\mathrm{T}} + \sigma^2\tilde o_{\sigma\mathrm{P}}(1).$$

6. Покращена оцінка ортогональної регресії

Теорема 3 дозволяє оцінити перший член розкладу асимптотичного відхилення по σ , при умові, що його складові σ^2, k_n, V_n відомі з достатньою точністю. Подамо оцінки, які використовувались для оцінювання даних параметрів у статтях [1,3,6]. У вирази для k_n, V_n замість невідомих параметрів (ξ_i,β_0) підставимо наближені значення $(x_i,\hat\beta_n)$, та означимо покращену оцінку параметра β_0 :

$$\tilde\beta_n := \hat\beta_n + \frac{\hat\sigma^2}{2}\hat V_n^{-1}\hat k_n , \tag{9}$$

де $\hat k_n := \frac{1}{n}\sum\limits_{i=1}^{n}k(x_i,\hat\beta_n)$, $\hat V_n := \frac{1}{n}\sum\limits_{i=1}^{n}V(x_i,\hat\beta_n)$, $\hat\sigma^2 := Q(\hat\beta_n)$.

Наступні леми забезпечують достатню точність оцінок $\hat\sigma^2, \hat k_n, \hat V_n$.

Лема 4. *Нехай виконуються умови (i)-(iv) та умова контрастності (con). Тоді*

$$\sigma^2 - \hat\sigma^2 = \sigma^2\tilde o_{\sigma\mathrm{P}}(1) .$$

Доведення леми безпосередньо випливає з формули (6) та співвідношення, яке виникає при доведенні Теореми 3: $Q(\hat\beta) - Q(\beta_0) = \sigma^4\tilde O_{\sigma\mathrm{P}}(1)$ (див. [1,6]).

Лема 5. *Нехай для моделі (1) виконуються умови (i)-(iv) та умова контрастності (con);* $F\in C^1(\mathbb{R}^q\times U)$ *, де* $U\supset\Theta$ *– відкрита, та для деяких сталих* λ, C *виконується нерівність* $\|F^\xi(\xi,\beta)\|\le Ce^{\lambda\|\xi\|}, \xi\in\mathbb{R}^q, \beta\in\Theta.$ *Тоді*

$$\frac{1}{n}\sum_{i=1}^{n}F(\xi_i,\beta_0) = \frac{1}{n}\sum_{i=1}^{n}F(x_i,\hat\beta_n) + \tilde o_{\sigma\mathrm{P}}(1) .$$

Ця лема використовувалась для оцінки k_n, V_n в явних моделях регресії (див. [1,3,6]). Обмеження на функцію F накладаються для того, щоб обмежити доданки з великими похибками ε_i .

Якщо умови леми 5 виконуються для функцій $k(\xi,\beta)$ та $V(\xi,\beta)$, то оцінки $\hat k_n$, $\hat V_n$ будуть достатньо точними. Але функції можуть бути необмеженими в точках $(x_i,\hat\beta_n)$, оскільки знаменник $\|G^\xi(x_i,\hat\beta_n)\|^2$ може бути близьким до нуля або навіть перетворюватись в 0 (у випадку явної моделі це неможливо: $\|G^\xi\|^2 = 1 + (g^\xi)^2 \ge 1$). Звичайно, в малому околі дійсних точок спостереження, при виконанні умов (i)-(iii), знаменник буде відділеним від нуля, але при великих похибках умова може порушуватись (див. далі приклад). Такої проблеми не виникає, якщо в функції $k(\xi,\beta)$ та $V(\xi,\beta)$ підставити значення $(\hat{h}_i,\hat{\beta}_n)\in\Gamma_{\hat{\beta}_n}$, де $\hat{h}_i := h(x_i,\hat{\beta}_n)$. Означимо альтернативну оцінку $\tilde{\beta}'_n$:

$$\tilde{\beta}'_n := \hat{\beta}_n + \frac{\hat{\sigma}^2}{2}\hat{V}_n'^{-1}\hat{k}'_n, \tag{10}$$

де $\hat{k}'_n := \frac{1}{n}\sum_{i=1}^{n} k(\hat{h}_i,\hat{\beta}_n)$, $\hat{V}'_n := \frac{1}{n}\sum_{i=1}^{n} V(\hat{h}_i,\hat{\beta}_n)$, $\hat{\sigma}^2 := Q(\hat{\beta}_n)$.

Означимо також многовид $W := \left\{(\xi;\beta)\,|\,G(\xi,\beta)=0,\ \beta\in\Theta\right\} = \bigcup_{\beta\in\Theta}\left(\Gamma_\beta\times\{\beta\}\right)\subset\mathbb{R}^q\times\Theta$.

Лема 6. *Нехай для моделі (1) виконуються умови (i)-(iv) та умова контрастності (con); $F\in C^1(U)$, де $U\supset W$ – відкрита множина; функція F експоненційно обмежена: при деяких сталих λ, C $\|F(\xi_\beta,\beta)\|\le Ce^{\lambda\|\xi_\beta\|}$, $(\xi_\beta,\beta)\in W$. (Функцією F буде вектор-функція $k(\xi,\beta)$ або матричнозначна функція $V(\xi,\beta)$). Тоді*

$$\tfrac{1}{n}\sum_{i=1}^{n} F(\xi_i,\beta_0) - \tfrac{1}{n}\sum_{i=1}^{n} F(\hat{h}_i,\hat{\beta}_n) = \tilde{o}_{\sigma P}(1).$$

Доведення. Як і при доведенні Теореми 1, окремо будемо окремо розглядати доданки для $i\in B_n$ та $i\notin B_n$. Індекси i будемо упускати. Наприклад, $\hat{h}=h(x,\hat{\beta}_n)$ – точка найменшої відстані від x до $\Gamma_{\hat{\beta}_n}$, $\xi\equiv h(\xi,\beta_0)$ – дійсна точка спостереження.

1. Розглянемо випадок $i\in B_n$. З Леми 2 випливає, що для довільного $\varepsilon>0$, для всіх $\sigma\in(0,\sigma_\varepsilon]$, при $n\ge n_{\sigma\varepsilon}$, $\hat{\beta}\in\bar{U}_{\nu_0}(\beta_0)$ зі ймовірністю не меншою за $1-\varepsilon$. Надалі будемо вважати що $\hat{\beta}\in\bar{U}_{\nu_0}(\beta_0)$.

Означимо H як функцію від x та β: $H(x,\beta):=F(h(x,\beta),\beta)$, де $x\in\mathbb{R}^q$, $\beta\in\Theta$, $h(x,\beta)\in W$. Тоді

$$\begin{aligned}\tfrac{1}{n}\sum\nolimits_{i\in B_n}\left[F(\hat{h}_i,\hat{\beta})-F(\xi_i,\beta_0)\right] &= \tfrac{1}{n}\sum\nolimits_{i\in B_n}\left[F(h(x_i,\hat{\beta}),\hat{\beta})-F(h(\xi_i,\hat{\beta}),\hat{\beta})+F(h(\xi_i,\hat{\beta}),\hat{\beta})-F(h(\xi_i,\beta_0),\beta_0)\right]=\\ &=\tfrac{1}{n}\sum\nolimits_{i\in B_n}\left[H(x_i,\hat{\beta})-H(\xi_i,\hat{\beta})\right]+\tfrac{1}{n}\sum\nolimits_{i\in B_n}\left[H(\xi_i,\hat{\beta})-H(\xi_i,\beta_0)\right].\end{aligned}$$

Розглянемо кожен доданок окремо. Знайдемо похідні за x,β, та застосуємо теорему про середнє. Але спочатку доведемо обмеженість похідних $H^x(x,\beta)=F^\xi\left(h(x,\beta),\beta\right)h^x$ та $H^\beta(x,\beta)=F^\xi\left(h(x,\beta),\beta\right)h^\beta+F^\beta\left(h(x,\beta),\beta\right)$.

Щоб знайти h^x та h^β, диференціюємо (3) за x,β. Отримаємо $h^x=T^{-1}\operatorname{Pr}_{(G^u)^\perp}$, $h^\beta=-T^{-1}\left(\dfrac{G^{u\mathrm{T}}G^\beta}{\|G^u\|^2}+a\operatorname{Pr}_{(G^u)^\perp}G^{u\beta}\right)$, де $\operatorname{Pr}_{(G^u)^\perp}=I_q-G^{u\mathrm{T}}G^u/\|G^u\|^2$, $T=I_q+a\operatorname{Pr}_{(G^u)^\perp}G^{uu}$, всі похідні функції G обчислені в $(h(x,\beta),\beta)$.

З (4) та Леми 1 випливає обмеженість $a(x,\beta)$, $h(x,\beta)$, а також невиродженість та відділеність від нуля власних чисел матриці T для довільного $\beta\in\Theta$. Тоді h^β,h^x, а також H^β,H^x, будуть обмеженими. Далі, $(\xi_i,\beta)\in W$, причому точки $\xi_i\in\bar{U}_A(\vec{0})$, $x_i\in\bar{U}_{\nu_0}(\xi_i)$ належать компакту. Всі точки $\left\{(h(x,\beta);\beta)\,|\,x\in\bar{U}_{\nu_0}(\xi_i), i\in B_n;\beta\in\Theta\right\}$ також належать компакту з множини W, в тому числі $\left(h(\xi_i,\beta),\beta\right)$ та $\left(h(x_i,\beta),\beta\right)$.

В околі точок W похідна $G^\xi(\xi,\beta)$ ненульова, а якщо ξ належить компакту, – відділена від нуля за нормою. Отже, функція F неперервно диференційовна в околі точок (h,β). А функція h неперервно диференційовна в околі $U_{\nu_0}(\xi_\beta,\beta)$, в тому числі в околі ξ_i та в точках x_i, де $i\in B_n$. Отже, H неперервно диференційовна як суперпозиція функцій F і h, а її похідні обмежені для всіх $i\in B_n$. Застосуємо до доданків теорему про середнє:

$$\left|\tfrac{1}{n}\sum\nolimits_{i\in B_n}\left[H(x_i,\hat{\beta})-H(\xi_i,\hat{\beta})\right]\right|=\left|\tfrac{1}{n}\sum\nolimits_{i\in B_n}H^\xi(\bar{\xi}_i,\hat{\beta})\sigma\tilde{\varepsilon}_i\right|\le\sup_{\xi\in\bar{U}_{\nu_0}(\xi_i),i\in B_n}\|H^\xi(\xi,\hat{\beta})\|\cdot\tfrac{\sigma}{n}\sum\nolimits_{i=1}^{n}\|\tilde{\varepsilon}_i\|=o_{\sigma P}(1),$$

$$\left|\tfrac{1}{n}\sum\nolimits_{i\in B_n}\left[H(\xi_i,\hat{\beta})-H(\xi_i,\beta_0)\right]\right|=\left|\tfrac{1}{n}\sum\nolimits_{i\in B_n}H^\beta(\xi_i,\bar{\beta})(\hat{\beta}-\beta_0)\right|\le\sup_{i\in B_n;\,\beta\in\Theta}\|H^\beta(\xi_i,\bar{\beta})\|\cdot\tilde{o}_{\sigma P}(1)=\tilde{o}_{\sigma P}(1).$$

2. Розглянемо випадок $i\notin B_n$. Доведемо, що $\tfrac{1}{n}\sum\nolimits_{i\notin B_n}\left[F(\hat{h}_i,\hat{\beta})-F(\xi_i,\beta_0)\right]=\tilde{o}_{\sigma P}(1)$.

а) Оскільки $\|\hat{h}_i\|\le\|x_i\|+\rho(x_i,\Gamma_{\hat{\beta}})\le\|x_i\|+\|x_i-\xi_i\|+\rho(\xi_i,\Gamma_{\hat{\beta}})\|\le\|\xi_i\|+2\|\varepsilon_i\|+const\le const+2\|\varepsilon_i\|$, то

$$\tfrac{1}{n}\sum\nolimits_{i\notin B_n}\|F(\hat{h}_i,\hat{\beta})\|\le\tfrac{C}{n}\sum\nolimits_{i\notin B_n}e^{\lambda\|\hat{h}_i\|}\le\tfrac{C}{n}\sum\nolimits_{i\notin B_n}e^{\lambda(const+2\|\varepsilon_i\|)}\le\tfrac{C_1}{n}\sum\nolimits_{i\notin B_n}e^{2A\|\varepsilon_i\|}=\tfrac{C_1}{n}\sum\nolimits_{i=1}^{n}e^{2\lambda\|\varepsilon_i\|}I_{\|\varepsilon_i\|\ge\nu_0}.$$

Для нормальних похибок $\mathrm{E}\, e^{2\lambda\|\varepsilon_i\|}<\infty$, отже $\mathrm{E}\, e^{2\lambda\sigma\|\tilde\varepsilon_i\|} I_{\|\tilde\varepsilon_i\|\ge \nu_0/\sigma}=o(1),\ \sigma\to 0+$. Звідси за нерівністю Чебишева легко отримати $\frac{1}{n}\sum_{i\notin B_n}\|F(\hat h_i,\hat\beta)\|=o_{\sigma\mathrm{P}}(1)$.

б) внаслідок компактності множини значень ξ_i та неперервності функції F на W маємо

$$\left\|\frac{1}{n}\sum_{i\notin B_n}F(\xi_i,\beta_0)\right\|\le \frac{D}{n}\sum_{i\notin B_n}1=o_{\sigma\mathrm{P}}(1).$$

Лема доведена. □

Кінцевим результатом всіх попередніх тверджень є оцінка, яка має менше асимптотичне зміщення.

Теорема 4. *Нехай для моделі (1) виконуються умови (i)-(vi), умова контрастності (con) та для деяких сталих* λ, C *виконується нерівність* $\|G^{\xi\xi}(\xi,\beta)\|+\|G^{\beta}(\xi,\beta)\|\le Ce^{\lambda\|\xi\|},\ \xi\in\mathbb{R}^q,\beta\in\Theta$. *Тоді*

$$\tilde\beta_n'-\beta_0=\sigma^2\tilde o_{\sigma\mathrm{P}}(1).$$

Якщо додатково $\|G^{\xi}(\xi,\beta)\|\ge\delta_0>0$ *для всіх* $\xi\in\mathbb{R},\ \beta\in\Theta$*, то також* $\tilde\beta_n-\beta_0=\sigma^2\tilde o_{\sigma\mathrm{P}}(1)$.

7. Приклад

Розглянемо о.о.р. та її асимптотичне відхилення на прикладі еліпса:

$$\begin{cases} x_i=\xi_i+\varepsilon_i, & i=\overline{1,n},\\ (\xi_i-c_0)^{\mathrm{T}}A_0(\xi_i-c_0)=1. \end{cases}$$

Симетрична додатно визначена матриця $A=(a_{ij})_{i,j=1}^{2}$ задає напрямки та величини осей, вектор $c=(c_1;c_2)^{\mathrm{T}}$ – центр, а $\xi_i\in\mathbb{R}^2$ – невідомі точки на еліпсі; $\beta=\left(a_{11};a_{12};a_{22};c_1;c_2\right)^{\mathrm{T}}$ – вектор параметрів,

$G(\xi,\beta)=\frac{1}{2}((\xi-c)^{\mathrm{T}}A(\xi-c)-1)$ – функція регресії. На параметри накладаємо наступні обмеження:

Власні числа λ_1,λ_2 *матриці* A *обмежені та відділені від нуля:* $0<C_1\le\lambda_1\le\lambda_2\le C_2$; $\|c\|\le R$.

Тоді виконуються умови (i), (ii). *Для похибок будемо вимагати виконання умови (iv).*

Перевіримо виконання умов (iii) та контрастності (con), та обмеження (v), (vi) на k_n,V_n. Маємо

$$G^{\xi}(\xi,\beta)=(\xi-c)^{\mathrm{T}}A;\quad G^{\xi\xi}(\xi,\beta)=A;\quad G^{\beta}(\xi,\beta)=\left[G^{A};G^{c}\right]=\left[\tfrac{1}{2}(\xi-c)(\xi-c)^{\mathrm{T}};-(\xi-c)^{\mathrm{T}}A\right].$$

Оскільки матриця A невироджена, $G^{\xi}=\vec{0}$ тоді і тільки тоді, коли $\xi-c=\vec{0}$, але в такому випадку $G=-\frac{1}{2}\ne 0$.

Отже, $\|G^{\xi}\|\ne 0$ для всіх допустимих β, $\xi\in\Gamma_\beta$, і *умова (iii) виконується.*

Розглянемо k_n. Нехай $\vec{u}=G^{\xi}/\|G^{\xi}\|$ – нормований вектор похідної G^{ξ}. Тоді

$$\mathrm{tr}\left(G^{\xi\xi}\left(I_q-\frac{G^{\xi\mathrm{T}}G^{\xi}}{\|G^{\xi}\|^2}\right)\right)=\mathrm{tr}\,A-\mathrm{tr}\left(A\cdot\vec{u}\cdot\vec{u}^{\mathrm{T}}\right)=\lambda_1+\lambda_2-\vec{u}^{\mathrm{T}}A\vec{u}\ge\lambda_1+\lambda_2-\lambda_2=\lambda_1>0. \tag{11}$$

Окремо розглянемо відділеність від нуля тих складових, що відповідають параметрам A і c:

$$k_n=\left[\frac{1}{n}\sum_{i=1}^{n}t\,G^{A}\Big|_{(\xi_i,\beta_0)};\frac{1}{n}\sum_{i=1}^{n}t\,G^{c}\Big|_{(\xi_i,\beta_0)}\right]=:\left[k_n';k_n''\right],$$

де t – ліва частина (11).

В похідній $G^{A}=\left((\xi^{(1)}-c_0^{(1)})^2;2(\xi^{(1)}-c_0^{(1)})(\xi^{(2)}-c_0^{(2)});(\xi^{(2)}-c_0^{(2)})^2\right)$ принаймні 2 компоненти відділені від нуля. Отже, k_n' *також відділене від нуля, що забезпечує виконання умови (v).*

За теоремою 3, перший член розкладу матриці $\hat A_n-A_0$ завжди буде ненульовим і навіть додатно визначеним (з врахуванням виконання умови контрастності і нерівності Коші-Буняковського) . А k_n'' може прямувати до нуля при $n\to\infty$, наприклад, якщо точки ξ_i (або їх міра μ) симетричні відносно центра c_0. Зокрема, при рівномірному розподілі ξ_i по еліпсу $k_n''\approx\vec{0}$, перший член розкладу $\hat c_n-c_0$ буде нульовим (ймовірно, тоді оцінка центра $\hat c_n$ виявиться консистентною з міркувань симетрії).

Умова контрастності *(con)* вимагає асимптотичної відділеності β_0 від інших точок параметичної множини – для того щоб при нульових похибках оцінка була єдиною і співпадала з β_0. У даному випадку різні значення параметра β задають різні криві. Але для знаходження еліпса потрібні чотири різні точки. Якщо ξ_i будуть приймати не більше трьох різних значень, ми не зможемо знайти β_0 навіть по відомим ξ_i. Також ми не знайдемо асимптотичне зміщення $\hat{\beta}_n$, якщо множина дійсних точок спостереження $\{\xi_i, i \geq 1\}$ має не більше трьох граничних точок. Нехай μ – борельова міра на еліпсі Γ_{β_0}, породжена послідовністю $\{\xi_i, i \geq 1\}$.

Для виконання умови (con) необхідно і достатньо, щоб міра μ була зосереджена більш ніж в трьох точках.

Залишилось перевірити виконання умови (vi) про невиродженість та відділеність від нуля власних чисел матриці V_n – суми n невід'ємих матриць рангу 1. Її можна сформулювати так: середнє арифметичне $x^T V_n x$, де x належить одиничній сфері з центром в початку координат, що дорівнює квадрату скалярного добутку $(G^\beta x)^2 / \| G^\xi \|^2$, відділене від нуля. Необхідно принаймні 5 різних матриць.

Для виконання умови (vi) необхідно і достатньо, щоб міра μ була зосереджена не менш ніж в п'яти точках.

У підсумку, для виконання умов (i)-(vi), міра μ повинна бути зосереджена не менш ніж в п'яти точках, наприклад, розподілена з деякою додатною щільністю по еліпсу або його дузі.

Умови леми 5 не виконуються, оскільки G^ξ перетворюється в $\vec{0}$ у центрі еліпса. А умови леми 6 – виконуються, отже, краще застосувати оцінки (10).

8. Висновки

Розглянуто оцінку ортогональної регресії у випадку неявної функціональної моделі з невідомим параметром σ, встановлено умови її неконзистентності, та знайдено перший член розкладу її асимптотичного відхилення. Оскільки для побудови покращеної оцінки не завжди можна беззастережно використовувати ті оцінки, що застосовувалися у випадку явної моделі, запропоновано альтернативні оцінки. В подальшому цікаво було б порівняти, які з них кращі для практичних застосувань. Також на прикладі моделі еліпсоїда проілюстровано значення умов (i)-(vi). Для подальшого уточнення оцінки потрібно поширити результати [6] на неявну модель – знайти другий член асимптотичного відхилення оцінки, та кращі оцінки σ^2, k_n, V_n.

1. *Репетацька Г.С.* Неконзистентність оцінки ортогональної регресії у векторній нелінійній моделі з похибками в змінних // Теорія ймовірності та математична статистика. – 2005. – Вип. 73 – С. 146-160. 2. *Fazekas I., Kukush A.G., Zwanzig S.* On inconsistency of the least squares estimator in nonlinear functional relations. Preprint. – Department of Statistics and Demography, Odense University, Denmark, 1998. 3. *Fazekas I., Kukush A., Zwanzig S.* Bias correction of nonlinear orthogonal regression // Ukrainian Mathematical Journal – 2001. – Vol. 56, № 8. – P. 1101-1118. 4. *Fuller W.A.* Measurement errors models. – Wiley, New York, 1987. 5. *Phanzagl J.* On the measurability and consistency of minimum contrast estimates // Metrika. –1969. –14. – P. 249-273. 6. *Repetatska G.* Modified orthogonal regression estimator in the quadratic errors-in-variables model // Theory of Stochastic Processes. – 2005. – Вип. 11(27), № 3-4. – P. 110-120.

Надійшла до редколегії 19.10.09

УДК 512.552

Н. Кайдан, асп.

САГАЙДАКИ СКІНЧЕННИХ КІЛЕЦЬ

В цій статті ми будемо розглядати правий та лівий сагайдаки скінченного комутативного кільця. Буде сформульовано та доведено теорему про ланцюгові скінченні кільця.

In this article we consider the right and left quivers finite commutative ring. Will be formulated and prove a theorem about chain finite rings

1. Вступ

Ми розглядаємо скінченні кільця з $1 \neq 0$. Скінченне кільце A називається розкладним, якщо A є прямим добутком двох кілець відмінних від нуля. В протилежному випадку кільце A називається нерозкладним.

Позначимо через R радикал скінченного кільця A. Будь-яке скінченне кільце A є артіновим з двох сторін кільцем і тому є напівдосконалим кільцем. Тоді правий (лівий) регулярний A-модуль A_A $({}_A A)$ має наступний розклад в пряму суму нерозкладних правих (лівих) проективних A-модулів: $A_A = P_1^{n_1} \oplus ... \oplus P_s^{n_s}$ $({}_A A = Q_1^{n_1} \oplus ... \oplus Q_s^{n_s})$, де X^n пряма сума n копій модуля X. Модулі $P_1,...,P_s$ $(Q_1,...,Q_s)$ попарно неізоморфні.

Нехай $1 = f_1 + ... + f_s$ буде відповідним розкладом одиниці кільця A в сумму попарно ортогональних ідемпотентів, тобто $f_i A = P_i^{s_i}$. Тоді ${}_A A = Af_1 \oplus ... \oplus Af_s = Q_1^{n_1} \oplus ... \oplus Q_s^{n_s}$ – розклад напівдосконалого кільця A в пряму суму нерозкладних проективних правих (лівих) A – модулів, тобто $Af_i = Q_i^{n_i}$, де Q_i – нерозкладний проективний правий A – модуль ($i = 1,...,s$). Розглянемо двосторонній розклад Пірса кільця A:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \dots & A_{ss} \end{pmatrix}$$

Двосторонній розклад Пірса радикала R кільця A має вигляд: $R = \bigoplus\limits_{i,j} f_i R f_j, \quad i, j = 1,...,s$.