УДК 519.21

Н. Н. Леоненко, д-р физ.-мат. наук,
Ле Фе До, асп. (Киев. ун-т)

Об оценках параметров регрессии случайных полей. I

Исследованы асимптотические свойства оценок модифицированного метода наименьших квадратов параметров регрессии случайных полей.

Досліджені асимптотичні властивості оцінок модифікованого методу найменших квадратів параметрів регресії випадкових полів.

В настоящей статье изучаются оценки параметров регрессии случайных полей по наблюдениям их независимых копий в случае, когда «случайный шум» представляет собой случайное поле с независимыми приращениями. Для случайных процессов подобные вопросы изучались в статье [1]. При других предположениях относительно «случайного шума» теория оценивания параметров регрессии случайных полей изложена в гл. 3 монографии [2].

1. Статистическая модель. Пусть $Y_1(t), \ldots, Y_N(t)$; $t \in I = [0, 1]^n \subseteq R^n$; $n \geqslant 1$, $N \geqslant 2$,— независимые случайные поля такие, что

$$Y_i(t) = (\alpha + \beta x_i)\, g(t) + \varepsilon_i(t), \quad t \in I, \tag{1}$$

где $g(t)$, $t \in I$,— известная неслучайная функция такая, что

$$0 \leqslant \| g \|_2 < \infty, \tag{2}$$

x_i, $i = \overline{1, N}$, — известные регрессионные константы, а $(\alpha, \beta) \in R^2$ — неизвестные коэффициенты регрессии, которые необходимо оценить на основании результатов наблюдения $Y_i(t)$, $t \in I$, $i = \overline{1, N}$. Здесь и далее

$$\|\cdot\|_2 = \left\{ \int_I (\cdot(t))^2\, dt \right\}^{1/2}, \tag{3}$$

а $\varepsilon_i(t)$, $t \in I$, $i = \overline{1, N}$, — независимые случайные поля, причем

$$\varepsilon_i(0) = 0, \quad \| \varepsilon_i(\cdot) \|_2 < \infty \text{ с вероятностью } 1.$$

Предположим, кроме того, что выполнено условие

$$M\varepsilon_i(t) = 0, \quad \operatorname{cov}(\varepsilon_i(s), \varepsilon_i(t)) = \prod_{i=1}^{n} \min\{t_i, s_i\} \tag{4}$$

или

$$M\varepsilon_i(t) = 0, \quad \operatorname{cov}(\varepsilon_i(s), \varepsilon_i(t)) = \frac{1}{2}\{|t| + |s| - |t - s|\}, \tag{5}$$

где $|t| = \sqrt{\sum_{i=1}^{n} t_i^2}$ — евклидова норма в R^n.

Замечание. Если $\varepsilon_i(t)$ — гауссовские случайные поля (этот случай представляет наибольший интерес для приложений), то в случае, когда выполнено условие (4), говорят о винеровском случайном поле в смысле Н. Н. Ченцова [3], а в случае, когда выполнено условие (5), говорят о винеровском случайном поле в смысле П. Леви [4]. При $n \geqslant 2$ свойства этих полей существенно различны.

Отметим лишь, что винеровское случайное поле $\varepsilon_i(t)$, $t \in I$, в смысле Ченцова имеет независимые приращения: случайные величины $\Delta[a, b)$ и $\Delta[c, d)$ независимы, если параллелепипеды $[a, b) = \{x \in I : a_i \leqslant x_i < b_i;\ i = \overline{1, n}\}$ и $[c, d) = \{x \in I : c_i \leqslant x_i < d_i;\ i = \overline{1, n}\}$ не пересекаются, где

$$\Delta[a, b) = \sum_{\nu} \varepsilon_i(b_1 - \nu_1(b_1 - a_1), \ldots, b_n - \nu_n(b_n - a_n)),$$

$\sum_{\nu} \ldots$ означает, что суммирование распространяется на все 2^n возможных значений булевого вектора $\nu = (\nu_1, \ldots, \nu_n)$.

В дальнейшем функция $g(t)$, $t \in I$, имеет вид

$$g(t) = g_{\mathrm{I}}(t) = \prod_{j=1}^{n} g_j(t_j), \quad t \in I, \tag{6}$$

или

$$g(t) = g_{\mathrm{II}}(t) = \sum_{i=1}^{n} g_j(t_j), \quad t \in I, \tag{7}$$

где $(g_1(t_1), \ldots, g_n(t_n))$ — известные функции.

2. Оценки модифицированного метода наименьших квадратов для схемы регрессии с по-

лем с независимыми приращениями. При каждом t рассмотрим функции

$$L(\alpha, \beta, t) = \sum_{i=1}^{N} [Y_i(t) - (\alpha + \beta x_i) g(t)]^2,$$

и предположим, что

$$L(\alpha, \beta) = \int_I L(\alpha, \beta, t)\, dt,$$

где здесь и далее

$$\int_I f(\cdot, t)\, dt = \int_0^1 \overset{n}{\dots} \int_0^1 f(\cdot, t)\, dt_1 \dots dt_n.$$

Оценками модифицированного метода наименьших квадратов (ОММНК) называются такие случайные величины $\check{\alpha}$, $\check{\beta}$, для которых

$$L(\check{\alpha}, \check{\beta}) = \min_{\alpha, \beta} L(\alpha, \beta).$$

Нетрудно показать, что ОММНК имеют вид

$$\check{\beta} = \| g \|_2^{-2} \Big[\sum_{i=1}^{N} (x_i - \bar{x})^2 \Big]^{-1} \sum_{i=1}^{N} (x_i - \bar{x}) \int_I Y_i(t) g(t)\, dt, \tag{8}$$

$$\check{\alpha} = N^{-1} \| g \|_2^{-2} \sum_{i=1}^{N} \int_I Y_i(t) g(t)\, dt - \check{\beta}\bar{x} = \Big[N \| g \|_2^2 \sum_{i=1}^{N} (x_i - \bar{x})^2 \Big]^{-1} \times$$

$$\times \sum_{i=1}^{N} \Big(\sum_{j=1}^{N} x_j^2 - N x_i \bar{x} \Big) \int_I Y_i(t) g(t)\, dt, \tag{9}$$

где $\bar{x} = (x_1 + \dots + x_N)/N$.

Рассмотрим свойства ОММНК в случае, когда выполнено условие (4). Случай, когда выполнено условие (5), рассмотрим во второй части работы.

Теорема 1. *При условиях* (1) — (4):

i) $M\check{\alpha} = \alpha$, $M\check{\beta} = \beta$ *для любых* α, β;

ii) *оценки* $\check{\alpha}$, $\check{\beta}$ *инвариантны относительно сдвигов, т. е. для любых действительных чисел* c *и* d

$$\check{\beta}(Y + (c + dx) g) = \check{\beta}(Y) + d, \tag{10}$$

$$\check{\alpha}(Y + (c + dx) g) = \check{\alpha}(Y) + c, \tag{11}$$

где символом $\check{\theta}(Y + (c + dx) g)$ *обозначены ОММНК параметра* θ ($\theta = \alpha$ *или* $\theta = \beta$), *полученные на основании наблюдений случайных полей* $Y_i(t) + (c + dx_i) g(t)$, $i = \overline{1, N}$.

При условиях (1) — (4), (6), (7) *справедливо условие*:

iii) *для* $k, l \in \{1, 2, \dots, n\}$ *пусть*

$$G_k(u) = \int_0^u g_k(s)\, ds, \quad 0 \leqslant u \leqslant 1; \qquad \bar{G}_k = \int_0^1 G_k(u)\, du,$$

$$\bar{G}_k(v) = \int_0^v G_k(u)\, du, \quad \bar{\bar{G}}_k = \int_0^1 \bar{G}(v)\, dv,$$

$$\sigma_k^2 = \sigma_{G_k}^2 = (G_k(1) - \bar{G}_k)^2 + \| G_k - \bar{G}_k \|_2^2; \qquad k = \overline{1, n},$$

$$\sigma_I^2 = \prod_{k=1}^{n} \sigma_k^2,$$

$$\sigma_{\mathrm{II}}^2 = \frac{1}{3^{n-1}} \sum_{k=1}^{n} \sigma_k^2 + \frac{2}{3^{n-2}} \sum_{k,l=1}^{n} \left(\frac{1}{2} G_k(1) - \overline{\overline{G}}_k\right)\left(\frac{1}{2} G_l(1) - \overline{\overline{G}}_l\right).$$

Тогда

$$D\check{\alpha} = \sigma^2 \sum_{i=1}^{N} x_i^2 \left[\| g \|_2^4 N \sum_{i=1}^{N} (x_i - \bar{x})^2\right]^{-1},$$

$$D\check{\beta} = \sigma^2 \left[\| g \|_2^4 \sum_{i=1}^{N} (x_i - \bar{x})^2\right]^{-1},$$

где $\sigma^2 = \sigma_{\mathrm{I}}^2$, *если выполнено условие* (6), *и* $\sigma^2 = \sigma_{\mathrm{II}}^2$, *если выполнено условие* (7).

Следовательно, ОММНК $\check{\alpha}$ *и* β *являются* L_2 *- состоятельными оценками* α *и* β *при условиях*

$$\lim_{N\to\infty} \sum_{i=1}^{N} (x_i - \bar{x})^2 = \infty; \qquad \lim_{N\to\infty} \sum_{i=1}^{N} x_i^2 \left[N \sum_{i=1}^{N} (x_i - \bar{x})^2\right]^{-1} = 0; \tag{12}$$

iv) пусть существуют такие $0 < \lambda < \infty$ *и* $-\infty < \mu < \infty$, *что при* $N \to \infty$

$$\max_{1\leqslant i\leqslant N} (x_i - \bar{x})^2 \left[\sum_{i=1}^{N} (x_i - \bar{x})^2\right]^{-1} \to 0; \qquad \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 \to \lambda^2; \qquad \bar{x} \to \mu. \tag{13}$$

Тогда последовательность случайных векторов $N^{1/2}(\check{\alpha} - \alpha, \check{\beta} - \beta)$ *слабо сходится к двумерному нормальному случайному вектору с нулевым средним и матрицей ковариаций* $\| g \|_2^{-4} \sigma^2 \Sigma$, *где матрица имеет вид*

$$\Sigma = \begin{pmatrix} (\lambda^2 + \mu^2)/\lambda^2 & -\mu/\lambda^2 \\ -\mu/\lambda^2 & 1/\lambda^2 \end{pmatrix}. \tag{14}$$

Доказательство. При условиях (2) — (7)

$$M \int_I \varepsilon_i(t) g(t)\, dt = \int_I M\varepsilon_i(t) g(t)\, dt = 0, \tag{15}$$

$$M\left(\int_I \varepsilon_i(t) g_{\mathrm{I}}(t)\, dt\right)^2 = \int_I \int_I M\varepsilon_i(t) \varepsilon_i(s) \prod_{k=1}^{n} g_k(t_k) g_k(s_k)\, dt_1 \dots dt_n ds_1 \dots ds_n =$$

$$= \int_I \int_I \prod_{k=1}^{n} \min(t_k, s_k) g_k(t_k) g_k(s_k)\, dt_k ds_k = \prod_{k=1}^{n} \left\{\int_0^1 g_k(t_k) \int_0^{t_k} s_k g(s_k)\, ds_k dt_k +\right.$$

$$\left. + \int_0^1 t_k g(t_k) (G_k(1) - G_k(t_k))\, dt_k\right\} = \prod_{k=1}^{n} \left\{(G_k(1) - \overline{G}_k)^2 + \| G_k - \overline{G}_k \|_2^2\right\} =$$

$$= \prod_{k=1}^{n} \sigma_k^2 = \sigma_{\mathrm{I}}^2;$$

и

$$M\left(\int_I \varepsilon_i(t) g_{\mathrm{II}}(t)\, dt\right)^2 = \int_I \int_I M\varepsilon_i(t) \varepsilon_i(s) \sum_{k,l=1}^{n} g_k(t_k) g_k(s_k)\, dt_1 \dots dt_n ds_1 \dots ds_n =$$

$$= \sum_{k=1}^{n} \int_0^1 \int_0^1 \min(t_k, s_k) g_k(t_k) g(s_k)\, dt_k ds_k \iint_{I^2 \setminus I_k^2} \prod_{\substack{l=1 \\ k\neq l}}^{n} \min(t_l, s_l)\, dt_l ds_l +$$

$$+2\prod_{k,l=1}^{n}\int_0^1\int_0^1 \min(t_k, s_k)\, g_k(t_k)\, g_k(s_k)\, dt_k ds_k \int_0^1\int_0^1 \min(t_l, s_l)\, g_l(t_l)\, g_l(s_l)\, dt_l ds_l \times$$

$$\times \iint\limits_{I^2\setminus I_k^2\times I_l^2} \prod_{\substack{q=1\\ q\neq k,l}}^{n} \min(t_q, s_q)\, dt_q ds_q = \sum_{k=1}^{n} \frac{1}{3^{n-1}}\sigma_k^2 +$$

$$+\frac{2}{3^{n-2}}\sum_{k,l=1}^{n}\left(\frac{1}{2}G_k(1)-\bar{\bar{G}}_k\right)\left(\frac{1}{2}G_l(1)-\bar{\bar{G}}_l\right);$$

где

$$I^2\setminus I_k^2 = \underbrace{[0,\ 1]^2}_{1}\times\ldots\times\underbrace{[0,\ 1]^2}_{k-1}\times\underbrace{[0,\ 1]^2}_{k+1}\times\ldots\times\underbrace{[0,\ 1]^2}_{n},$$

$$I^2\setminus I_k^2\times I_l^2 = \underbrace{[0,\ 1]^2}_{1}\times\ldots\times\underbrace{[0,\ 1]^2}_{k-1}\times\underbrace{[0,\ 1]^2}_{k+1}\times\ldots\times\underbrace{[0,\ 1]^2}_{l-1}\times\underbrace{[0,\ 1]^2}_{l+1}\times\ldots\times\underbrace{[0,\ 1]^2}_{n}.$$

Отсюда непосредственно вытекают свойства (i), (ii), (iii). Докажем теперь (iv).

Предположим, что $\check{\delta} = \check{\alpha} + \bar{x}\check{\beta}$. Тогда $D(N^{1/2}\check{\delta}) = \|g\|_2^{-4}\sigma^2$ и $M\check{\delta} = \alpha + \bar{x}\beta$.

Из (iii) и (13) имеем

$$\lim_{N\to\infty} D(N^{1/2}\check{\beta}) = \|g\|_2^{-4}\sigma^2\lambda^{-2}; \qquad \operatorname{cov}(\check{\delta},\ \check{\beta}) = 0$$

для всех $N \geqslant 2$. С другой стороны, рассмотрим поле

$$\check{\gamma} = a_1 N^{1/2}(\check{\delta} - M\check{\delta}) + a_2 N^{1/2}(\check{\beta} - \beta).$$

Ясно, что $M\check{\gamma} = 0$, $D(\check{\gamma}) = (a_1^2 + a_2^2\lambda^{-2})\|g\|_2^{-4}\sigma^2$, где a_1 и a_2 — произвольные действительные константы. Запишем $\check{\gamma}$ в более конкретном виде:

$$\check{\gamma} = a_1 N^{-1/2}\|g\|_2^{-2}\sum_{i=1}^{N}\int_I Y_i(t)\, g(t)\, dt - a_1 N^{1/2} M\check{\delta} + a_2 N^{1/2}\|g\|_2^{-2}\times$$

$$\times\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^{-1}\sum_{i=1}^{N}(x_i-\bar{x})\int_I Y_i(t)\, g(t)\, dt - a_2 N^{1/2}\beta = \|g\|_2^{-2}\times$$

$$\times\sum_{i=1}^{N}\left\{a_1 N^{1/2} + a_2 N^{1/2}\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^{-1}(x_i-\bar{x})\right\}\int_I Y_i(t)\, g(t)\, dt -$$

$$-(a_1 N^{1/2} M\check{\delta} + a_2 N^{1/2}\beta).$$

Из (12) и (13) легко показать, что $\check{\gamma}$ удовлетворяет условию теоремы V.1.2 из [5]. Действительно,

$$a_i^* = a_1 N^{-1/2} + a_2 N^{1/2}\left[\sum_{j=1}^{N}(x_j-\bar{x})^2\right]^{-1}(x_i-\bar{x}),$$

$$\max(a_i^*)^2 = \max\left\{a_1^2 N^{-1} + a_2^2 N\left[\sum_{j=1}^{N}(x_j-\bar{x})^2\right]^{-2}(x_i-\bar{x})^2 + 2a_1a_2\times\right.$$

$$\left.\times\left[\sum_{j=1}^{N}(x_j-\bar{x})^2\right]^{-1}(x_i-\bar{x})\right\} \leqslant a_1^2 N^{-1} + a_2^2 N\left[\sum_{j=1}^{N}(x_j-\bar{x})^2\right]^{-1}\max(x_i-\bar{x})^2\times$$

$$\times\left[\sum_{j=1}^{N}(x_j-\bar{x})^2\right]^{-1} + 2a_1a_2\left[\sum_{j=1}^{N}(x_j-\bar{x})^2\right]^{-1}\max((x_j-\bar{x}),\ 1)\to 0; \qquad N\to\infty,$$

и следовательно,

$$\sum_{i=1}^{N} a_i^2 = a_1^2 + a_2^2 N \left[\sum_{i=1}^{N} (x_i - \overline{x})^2\right]^{-1} \to a_1^2 + a_2^2\lambda^{-2}, \quad N \to \infty.$$

Из теоремы V. 1. 2 [5] следует, что $a_1 N^{1/2}(\check{\delta} - M\check{\delta}) + a_2 N^{1/2}(\check{\beta} - \beta)$ имеет асимптотическое нормальное распределение $(0, (a_1^2 + a_2^2\lambda^{-2}) \| g \|_2^{-4}\sigma^2)$ для всех a_1 и a_2.

Используя лемму 5—1 [6], можно доказать, что $N^{1/2}(\check{\delta} - M\check{\delta}, \check{\beta} - \beta)$ имеет асимптотическое нормальное распределение $((0, 0), \| g \|_2^{-4}\sigma^2\Delta)$, где $\Delta = \begin{pmatrix} 1 & 0 \\ 0 & 1/\lambda^2 \end{pmatrix}$.

Применив лемму 5—1 [6] для поля $N^{1/2}\{[\check{\delta} - M\check{\delta} + \overline{x}(\check{\beta} - \beta)], \check{\beta} - \beta\}$, получим (iv). Теорема 1 доказана.

1. *Wu T. J., Wasan N. T.* Time integration least squares estimators of regression parameters of independent stochastic processes // Stoch. Proc. Appl.— 1990.— 35.— P. 141—148.
2. *Леоненко Н. Н., Иванов А. В.* Статистический анализ случайных полей.— Киев : Вища шк. Изд-во при Киев. ун-те, 1986.— 216 с.
3. *Ченцов Н. Н.* Винеровские случайные поля от нескольких параметров // Докл. АН СССР.— 1956.— **106**.— С. 155—161.
4. *Леви П.* Стохастические процессы и броуновскоедвижение.— М. : Наука, 1972.— 376 с.
5. *Гаек Я., Шидак З.* Теория ранговых критериев.— М. : Наука, 1971.— 376 с.
6. *Adichie J. N.* Estimates of regression parameters based on rank tests // Ann. Statist.— 1967.— 38.— P. 894—904.

Получено 12.02.91