УДК 519.21

Р. Е. МАЙБОРОДА, канд. физ.-мат. наук (Киев. ун-т)

Непараметрическое обнаружение разладки по наблюдениям с погрешностью

Рассмотрена задача обнаружения разладки в последовательности независимых случайных величин по наблюдениям, которые представляют собой сумму исследуемых данных с независимым неоднородным шумом. Построены сильно состоятельная оценка момента разладки и доверительный интервал.

Розглянуто задачу виявлення розладки у послідовності незалежних випадкових величин за спостереженнями, що являють собою суму досліджуваних даних з незалежним неоднорідним шумом. Побудовані сильно слушна оцінка моменту розладки і надійний проміжок.

Задача непараметрического обнаружения разладки последовательности независимых случайных величин (с. в.) хорошо исследована [1—3]. Обычно предполагают, что наблюдения $\zeta_1, \ldots, \zeta_N$ представляют собой точно измеренные значения исследуемой последовательности. В данной работе рассматривается случай, когда в наблюдениях присутствует неоднородная аддитивная погрешность (шум) с известным распределением.

Пусть $\xi_j, \eta_j, \varepsilon_j, j \in N$,— независимые в совокупности с. в., а $\zeta_j = \xi_j + \varepsilon_j$ при $1 \leqslant j \leqslant \vartheta N$ и $\zeta_j = \eta_j + \varepsilon_j$ при $\vartheta N \leqslant j \leqslant N$, где $\vartheta \in (0, 1)$. Число ϑ назовем моментом разладки выборки ζ_j, а ε_j — погрешностями наблюдений ζ_j. Предполагается, что функции распределения $F(x) = \mathbf{P}\{\xi_j < x\}$ и $H(x) = \mathbf{P}\{\eta_j < x\}$ неизвестны, $F \neq H$. Производящие функции моментов погрешностей $\Phi_j(u) = \mathbf{E}\exp(u\varepsilon_j) < \infty$ известны. Требуется оценить значение момента разладки ν по наблюдениям.

Обозначим через $f_\xi(u) = \mathbf{E}\exp(u\xi)$ производящую функцию моментов величины ξ,

$$R_N(u, t) = N^{-1}\left((1-t)\sum_{j=1}^{tN}\exp(u\zeta_j)/\Phi_j(u) - t\sum_{j=tN}^{N}\exp(u\zeta_j)/\Phi_j(u)\right).$$

В качестве оценки для ϑ рассмотрим $\widehat{\vartheta}_N = \widehat{\vartheta}_N(U) = \operatorname{argmax}_t \sup_{u\in U} |R_N(u, t)|$, где U — некоторое подмножество $\mathbb{R}$. В дальнейшем будут использоваться два вида множеств U : U — конечное множество точек $U = S = \{u_1\}_{i=1}^n$ или U — отрезок с центром в 0, $U = B = [-b, b]$, $b > 0$. Субгауссовским штандартом (нормой с. в. ε) назовем $\operatorname{sub}(\varepsilon) = \inf\{\tau > 0 : \exp(u\varepsilon) \leqslant \exp(\tau^2u^2/2)\ \forall u \in \mathbb{R}\}$ (ср. с. [4]).

Теорема 1. *Пусть существует $a \in \mathbb{R}$, такое, что $\forall\ u \in U\ |2u| < a$ и $f_\xi(v) < \infty, f_\eta(v) < \infty$ при $|v| < a$. Тогда*

1) *если $f_\xi(v) \neq f_\eta(v)$ при некоторых $v \in S$ и $\sum_{j=1}^{\infty} j^{-2}\Phi_j(2u) < \infty$ для всех $u \in S$, то $\widehat{\theta}_N(S) \to \theta$ п. н. при $N \to \infty$;*

2) *если $\operatorname{sub}^2(\varepsilon_j) = o(\ln j)$ $(j \to \infty,\ N \to \infty)$, то $\widehat{\theta}_N(B) \to \theta$ п. н.*

Обозначим

$$X_j^*(u) = \exp(u(\xi_j + \varepsilon_j))/\Phi_j(u),$$

$$Y_j^*(u) = \exp(u(\eta_j + \varepsilon_j))/\Phi_j(u), \quad X_j = X_j^* - \mathbf{E}X_j^*, \quad Y_j = Y_j^* - \mathbf{E}Y_j^*,$$

$$Z_j = X_j \text{ при } j \leqslant \vartheta N, \quad Z_j = Y_j \text{ при } j > \vartheta N,$$

$$I_N(u, t) = N^{-1}\sum_{j=1}^{tN} Z_j(u), \quad \alpha(N, U) = \sup_{u\in U}\max\left\{N^{-1}\sum_{j=1}^{N} X_j,\ N^{-1}\sum_{j=1}^{N} Y_j\right\}.$$

Лемма 1. *Если выполнены условия п. 1 теоремы, то $\alpha(N, S) \to 0$ п. н., $N \to \infty$.*

Доказательство. Достаточно проверить, что $N^{-1}\sum_{j=1}^{N} X_j(u) \to 0$ п. н. при любом $u \in S$. Поскольку $\mathbf{E}X_j(u) = 0$, то по закону больших чисел [5, с. 96] сходимость следует из конечности

$$\sum_{j=1}^{N} j^{-2}\mathbf{E}(X_j^*(u))^2 = \sum_{j=1}^{N} j^{-2} f_\xi(2u)\,\Phi_j(2u)/\Phi_j^2(u) < \infty,$$

поскольку из $\mathbf{E}\varepsilon_j = 0$ следует $\Phi_j(u) \geqslant 1\ \forall u$.

Лемма 2. *Если выполнены условия п. 2 теоремы 1, то $\alpha(N, B) \to 0$ п. н. $N \to \infty$.*

Доказательство см. в [6].

Лемма 3. *Если $\alpha(N, U) \to 0$ п. н. при $N \to \infty$, то*

$$\sup_{t\in[0,1]}\sup_{u\in U}|I_N(u, t)| \to 0 \text{ п. н.}$$

Доказательство. Если $0 < t \leqslant \theta$, то $|I_N(u,t)| < t\alpha(tN, U)$ а при

$$\theta < t < 1, \quad |I_N(u,t)| \leqslant \theta\alpha(\theta N, U) + N^{-1}\Big|\sum_{j=1}^{Nt} Y_j - \sum_{j=1}^{\theta N} Y_j\Big| \leqslant$$

$$\leqslant 2\theta\alpha(\theta N, U) + t\alpha(tN, U).$$

Для произвольного $\delta > 0$ обозначим $N_0(\delta) = \inf\{m : \alpha(k) < \delta \,\forall k \geqslant m\}$ и $C = \sup_k \alpha(k, U)$. Поскольку $\alpha(k, U) \to 0$ п. н., то $N_0(\delta) < \infty$ и $C < \infty$ п. н. Положим $t_0 = \delta/C$. Тогда при $0 < t < t_0$ $t\alpha(tN, U) < \delta$ для любых N, а при $t_0 < t < 1$ $t\alpha(tN, U) < \delta$ для $N > N_0(\delta)/t_0$. Таким образом, при $0 < t < 1$ и $N > N_0(\delta)/t_0$, $t\alpha(tN, U) < \delta$. Следовательно, $|I_N(u,t)| \leqslant 3\delta$ при $N > N_0(\delta)/t_0$. В силу произвольности δ $\sup_{u,t}|I_N(u,t)| \to 0$ п. н., $N \to \infty$.

Доказательство теоремы. Легко видеть, что $\mathrm{E}R_N(u,t) = (f_\xi(u) - f_\eta(u))\, g(t,\theta)$, где $g(t,\theta) = t(1-\theta)$ при $t < \theta$, $g(t,\theta) = (1-t)\theta$ при $t \geqslant \theta$. Заметим, что $\sup_{u \in U}|f_\xi(u) - f_\eta(u)| \neq 0$. Действительно, при $U = S$ это верно по условию теоремы, а при $U = B$ следует из аналитичности f_ξ и f_η в окрестности 0 и $F \not\equiv H$. Поскольку максимум $g(t,\theta)$ достигается в единственной точке $t = \theta$, то достаточно доказать, что $\sup_{u,t}|R_N(u,t) - \mathrm{E}R_N(u,t)| \to 0$ п. н. Однако

$$J_N = |R_N(u,t) - \mathrm{E}R_N(u,t)| = |I_N(u,t) - tI_N(u,1)| \leqslant 2\sup_{u,s}|I_N(u,s)|, \qquad (1)$$

так что утверждение теоремы следует из леммы 3.

Используя (1), можно построить грубый доверительный интервал для момента разладки θ. Пусть $f_\xi(u) < \infty$, $f_\eta(u) < \infty$ при $u \leqslant a$, $2b < a$. Обозначим

$$f_0 = \max\{f_\xi(+2a), f_\eta(\pm 2a)\},$$

$$\Phi_j = \max_{|u| \leqslant a}\{\Phi_j(u), \Phi_j', \Phi_j''(u)\}, \quad \beta_2 = 2(2^{1/3} + 2^{-2/3})/\pi,$$

$$\beta_3 = (3^{-3/4} + 3^{1/4})/\pi^2, \quad \Delta = a - 2b,$$

$$r_j = \Phi_j(1 + 2\beta_2 a^{1/3}\Delta^{-4/3} + 3\Phi_j + \beta_3 a^{1/4}\Delta^{-9/4} + 2\beta_2 a^{1/3}\Delta^{-4/3}\Phi_j),$$

$$\delta = \max_{|u| < b}|f_\xi(u) - f_\eta(u)|, \quad K_N = \sum_{j=1}^{N} r_j.$$

Теорема 2. *Пусть*

$$t = \hat\theta([-b, b]), \quad p \in (0, 1),$$

$$\varepsilon = \varepsilon_p = (32 f_0 b K_N/p)^{1/2}/(N\delta), \quad t_- = (t + (t^2 - 4\varepsilon_p)^{1/2})/2,$$

$$t_+ = (1 + t - ((1-t)^2 - 4\varepsilon_p)^{1/2})/2.$$

Тогда $\mathrm{P}\{t_- \leqslant \theta \leqslant t_+\} \leqslant p$.

Доказательство. Пусть x, y — гладкие функции, отображающие $[-b, b]$ в $\mathbb{R}$. Обозначим

$$(x, y) = x(0)\, y(0) + \int_{-b}^{b} x'(u)\, y'(u)\, du, \quad \|x\| = (x, x)^{1/2},$$

H — замыкание класса гладких функций в норме $\|\cdot\|$. Тогда H — сепарабельное гильбертово пространство и $\sup_{|u|<b}|x(u)| \leqslant \|x\|$ [6]. Оценим

$$\mathbf{E}\|X_j\|^2 \leqslant \mathbf{E}\|X_j^*\|^2 \leqslant \int_{-b}^{b}\mathbf{E}(\exp(u\zeta_j)\,\zeta_j/\Phi_j(u) + \exp(u\zeta_j)\,\Phi_j'/\Phi_j^2(u))^2 du.$$

Учитывая, что $\Phi_j(u) > 1$ и используя для оценки $f'_\xi(u)$, $f''_\xi(u)$ формулу Коши, можно получить $\mathbf{E}\| X_j \|^2 \leqslant 2bf_0 r_j$.

Таким образом, применяя неравенство Колмогорова [5, с. 454], получаем

$$\mathbf{P}\{\sup_{t,u} | \mathrm{I}_N(u, t) | > \lambda\} \leqslant \mathbf{P}\{\sup_t \| \mathrm{I}_N(\cdot, t) \| \geqslant \lambda\} \leqslant$$

$$\leqslant \mathbf{E} \| I_N(\cdot, 1) \|^2/\lambda^2 \leqslant 2bf_0 K_N/(N\lambda)^2.$$

Из (1) имеем $\mathbf{P}\{J_N > \delta\varepsilon/2\} < p$. Пусть событие $J_N < \delta\varepsilon/2$ наступило. Заметим, что t может быть равно $\hat{\theta}_N$ ($[-b, b]$) только в том случае, если $2J_N > \delta(g(\theta, \theta) - g(t, \theta))$, так что $\varepsilon > \theta(1-\theta) - g(t, \theta)$. Из этого неравенства получаем $t(1-\theta) > \theta(1-\theta) - \varepsilon$ при $t > \theta$ и $(1-t)\theta > \theta(1-\theta) - \varepsilon$ при $t > \theta$. Решая эту систему неравенств относительно θ, получаем $t_- < \theta < t_+$. Теорема доказана.

Замечания. 1. В условиях п. 2 теоремы 1, при $N \to \infty$ $K_N = o(N^{1+\alpha})$ для любого $\alpha > 0$, так что $t_+ - t_- = o(N^{-1/2+\alpha})$. Если для всех j $\operatorname{sub}(\varepsilon_j) < C < \infty$, то $t_+ - t_- = O(N^{-1/2})$.

2. При построении t_+ и t_- используются величины f_0 и δ, связаные с неизвестными функциями f_ξ и f_η. Поэтому необходимо либо иметь оценки сверху для f_0 и снизу для δ, либо оценивать их по выборке. В последнем случае можно использовать эмпирическую производящую функцию моментов [6].

1. *Дарховский Б. С.* Непараметрический метод оценивания интервалов однородности случайной последовательности // Теория вероятностей и ее применения.— 1985.— 30, вып. 4.— С. 795—799.
2. *Майборода Р. Е.* Обнаружение изменений свойств случайных процессов // Статист. пробл. управления.— 1984.— Вып. 65.— 244 с.
3. *Майборода Р. Е.* Квантильные методы поиска разладки // Там же.— 1988.— Вып. 83.— С. 105—109.
4. *Булдыгин В. В.*, *Козаченко Ю. В.* О субгауссовских случайных величинах // Укр. мат. журн.— 1980.— 32, № 6.— С. 723—730.
5. *Гихман И. И.*, *Скороход А. В.* Теория случайных процессов: В 3-х т.— М. : Наука, 1971.— Т. 1.— 664 с.
6. *Майборода Р. Е.* Оценивание производящей функции моментов по наблюдениям с погрешностью // Теория вероятностей и мат. статистика.— 1988.— Вып. 39.— С. 87—90.

Получено 30,10.90