

ШВИДКИЙ АЛГОРИТМ ПОШУКУ БАГАТЬОХ МОМЕНТІВ ЗМІНИ

УДК 519.21

Р. Є. МАЙБОРОДА ТА О. В. СУГАКОВА

РЕЗЮМЕ. Для динамічного непараметричного алгоритму апостеріорного пошуку багатьох моментів зміни у послідовності незалежних спостережень знайдено достатні умови збіжності оцінок зі швидкістю $A_N \ln N/N$, де N — кількість спостережень, A_N — будь-яка числова послідовність, що прямує до ∞ .

Задача апостеріорного виявлення моменту зміни розподілу у послідовності незалежних випадкових величин добре вивчена [1, 2]. Менше досліджений випадок, коли таких змін у послідовності може бути декілька. Непараметричні алгоритми для оцінки таких моментів розглядались у [3]. Вони потребують як мінімум порядку N^2 обчислень, де N — кількість елементів послідовності. У [4, 5] розглянутий алгоритм динамічного програмування, що базується на параметричних припущеннях про послідовність і потребує лише порядку KN обчислень, де K — кількість різних можливих розподілів у послідовності. У [6] запропоновано використати цей алгоритм для непараметричного оцінювання і показано, що він дозволяє досягти точності оцінювання моментів зміни порядку $N^{-\alpha}$, де $\alpha < \frac{1}{2}$ (за ймовірністю). У даній роботі ми покажемо, що при виконанні деяких (необтяжливих) додаткових припущень можна досягти швидкості збіжності м.н. порядку $A_N N^{-1} \ln N$, де A_N — довільна послідовність така, що $A_N \rightarrow \infty$.

Нехай $\xi_1, \dots, \xi_N$ — послідовність незалежних випадкових елементів деякого вимірного простору $\mathcal{X}$. Відомо, що $P\{\xi \in A\} = F_{h^0}(A)$, де $F_h \in F = \{F_i, i = 1, \dots, K\}$, $h^0 = \{h_j^0, j = 1, \dots, N\}$ — не випадкова послідовність номерів $h_j^0 \in \{1, \dots, K\}$, яка має вигляд $h_j^0 = \text{const}$ при $k_i = [\theta_i N] \leq j < k_{i+1} = [\theta_{i+1} N]$, де $0 = \theta_0 < \theta_1 < \dots < \theta_R < \theta_{R+1} = 1$ — фіксовані не випадкові числа, які називають моментами зміни, (k_j називають точками зміни). Послідовності номерів розподілів h будемо називати траєкторіями, h^0 — істиною траєкторією послідовності $\xi_1, \dots, \xi_n$.

Нехай $\phi: \mathcal{X} \times \{1, \dots, K\} \rightarrow \mathbf{R}$ — деяка вимірна функція, $\pi_N(g, l) = \pi_N I(g = l)$, де π_N — деяке додатне число.

Розглянемо функціонал

$$J(h) = \sum_{i=1}^N (\pi_N(h_i, h_{i-1}) + \phi(\xi_i, h_i)).$$

(Тут за означенням приймаємо $h_0 = h_1$.) Нехай виконується умова

$$\kappa = \inf_{1 \leq i, k \leq K, i \neq k} (E \phi(\eta_i, k) - E \phi(\eta_i, i)) > 0. \quad (A)$$

1991 AMS Mathematics Subject Classification. Primary 62G20, 93E10; Secondary: 93E24, 60E15.

(Тут η_i — випадковий елемент з розподілом F_i .) Тоді $J(h)$ має характер “штрафу” за відхилення даної траєкторії h від істинної траєкторії h^0 . Справді, завдяки умові (A), мінімум $E \sum_{j=1}^N \phi(\xi_j, h_j)$ досягається при $h = h^0$, а додавання $\pi_N(h_i, h_{i-1})$ має значення штрафування за внесення зайвих змін у траєкторію h . Отже можна сподіватись, що траєкторія

$$\hat{h} = \arg \min_h J(h)$$

буде близькою до істинної. У [4] описано швидкий алгоритм знаходження $\hat{h}$.

Приклад. Нехай $K = 2$, $\mathcal{X} = \mathbf{R}$ і відомо, що $\text{med}(\eta_1) > 0$, $\text{med}(\eta_2) < 0$ ($\text{med}(x)$ — медіана x). Візьмемо $\phi(x, 1) = I(x > 0)$, $\phi(x, 2) = I(x < 0)$. Тоді умова (A) виконується. Інші приклади вибору функцій ϕ , які потребують більшої інформації про розподіли F_k див. [6].

Позначимо

$$\hat{k}_l = \min\{l: \hat{h}_l \neq \hat{h}_j, 1 \leq j < l\},$$

$$\hat{k}_i = \min\{l: \hat{h}_l \neq \hat{h}_j, \hat{k}_{i-1} \leq j < l\}$$

— j -та точка зміни у траєкторії $\hat{h}$, $\hat{R}$ — кількість точок зміни у $\hat{h}$: якщо $l = \hat{k}_{\hat{R}}$, то $\hat{h}_l = \hat{h}_{l+1} = \dots = \hat{h}_N$. Відповідно оцінками моментів зміни є $\hat{\theta}_i = \frac{\hat{k}_i}{N}$

Наша мета — довести спроможність $\hat{\theta}_i$ при $N \rightarrow \infty$ і отримати оцінки швидкості збіжності. При цьому спостереження $\xi_1, \dots, \xi_N$ слід розглядати як елемент схеми серій: $\{\xi_1, \dots, \xi_N\} = \{\xi_1^N, \dots, \xi_N^N\}$, $N \rightarrow \infty$. Для спрощення позначень верхній індекс писати не будемо.

Для формулювання та доведення основного результату нам знадобиться теорія субгаусових випадкових величин [7]. Дійснозначну випадкову величину ζ називають субгаусовою, якщо для всіх $\lambda \in \mathbf{R}$ і деякого C виконується

$$E e^{\lambda \zeta} \leq e^{C\lambda^2/2} \quad (1)$$

Інфімум тих C для яких виконано (1) називають субгаусовим штандартом ζ і позначають $\text{sub}(\zeta)$. Відомо [7] що клас всіх субгаусових випадкових величин з нормою $\|\zeta\| = \sqrt{\text{sub}(\zeta)}$ являє собою банахів простір, причому коли ζ_1 і ζ_2 незалежні, то

$$\text{sub}(\zeta_1 + \zeta_2) \leq \text{sub}(\zeta_1) + \text{sub}(\zeta_2). \quad (2)$$

Справедлива оцінка

$$P\{|\zeta| > x\} \leq 2 \exp\left(-\frac{x^2}{2\text{sub}(\zeta)}\right) \quad (3)$$

З (2) і (3) випливає, що коли $\zeta_1, \dots, \zeta_n$ незалежні субгаусові в.в. і $\text{sub}(\zeta_j) < \sigma^2$ для деякого $\sigma < \infty$, то

$$P\left\{\left|\sum_{j=1}^n \zeta_j\right| > x\right\} \leq 2 \exp\left(-\frac{x^2}{2n\sigma^2}\right). \quad (4)$$

Умова (B): При всіх $i, g = 1, \dots, K$, в.в. $\zeta_{i,g} = \phi(\eta_i, g) - E \phi(\eta_i, g)$ є субгаусовими і $\text{sub}(\zeta_{i,g}) \leq \sigma^2$.

Умова (B) не є занадто обтяжливою, оскільки функції ϕ обирає статистик. Відомо, що коли $|\zeta| < C < \infty$ м.н. і $E \zeta = 0$, то $\text{sub}(\zeta) \leq C$, отже умова (B) виконується для всіх обмежених функцій ϕ , зокрема і для нашого прикладу.

Теорема. Нехай виконані умови (А) і (В) і для послідовностей π_N та d_N виконуються співвідношення:

$$\frac{\pi_N}{\ln N} \rightarrow \infty, \quad \frac{d_N}{\pi_N} \rightarrow 0, \quad \frac{d_N}{N} \rightarrow 0 \quad \text{при } N \rightarrow \infty. \quad (C)$$

Тоді існує таке випадкове $n = n(\omega) < \infty$ м.н., що для всіх $N > n$

$$\hat{R}_N = R \quad (5)$$

$$N \frac{\max_i |\hat{\theta}_i - \theta_i|}{d_N} \rightarrow 0 \text{ м.н. при } N \rightarrow \infty \quad (6)$$

Доведення. Позначимо $\Phi_j(h) = E \phi(\xi_j, h)$. Подію A_N визначимо як

$$A_N = \left\{ \max_{1 \leq l_1 < l_2 \leq N, 1 \leq g \leq K} \sum_{j=l_1}^{l_2} (\phi(\xi_j, h_j^0) - \phi(\xi_j, g)) \leq a_N \right\}.$$

Лема 1. Нехай $a_N > 0$, $a_N / \ln N \rightarrow \infty$ при $N \rightarrow \infty$. Тоді знайдеться випадковий номер n , $n < \infty$ м.н., такий, що A_N виконано для всіх $N > n$.

Доведення. Внаслідок довільності a_N досить довести, що $\forall g = 1, \dots, K$

$$\max_{1 \leq l_1 < l_2 \leq N} \sum_{j=l_1}^{l_2} (\phi(\xi_j, h_j^0) - \phi(\xi_j, g)) \leq a_N. \quad (7)$$

Оскільки, при $h_j^0 = g$, $\phi(\xi_j, h_j^0) - \phi(\xi_j, g) = 0$ і $\sum_{j=l_1}^{l_2} (\phi(\xi_j, h_j^0) - \phi(\xi_j, g))$ можна зобразити у вигляді суми не більше ніж R доданків в яких підсумовування ведеться по інтервалах однорідності h_j^0 , то досить довести (7) для випадку, коли $h_j^0 = h \neq g$, $j = 1, \dots, N$.

Враховуючи умову (В) та (4), для довільних $x > \text{маємо}$

$$P \left\{ \sum_{j=l_1}^{l_2} (\zeta_{jh} - \zeta_{jg}) \geq x \right\} \leq 2 \exp \left(- \frac{x^2}{4(l_2 - l_1 + 1)\sigma^2} \right).$$

Оскільки, за умовою (А),

$$\phi(\xi_j, h) - \phi(\xi_j, g) = \zeta_{jh} - \zeta_{jg} + \Phi_j(h) - \Phi_j(g) \leq \zeta_{jh} - \zeta_{jg} - \kappa,$$

то

$$P \left\{ \sum_{j=l_1}^{l_2} (\phi(\xi_j, h) - \phi(\xi_j, g)) \geq x - \kappa(l_2 - l_1 + 1) \right\} \leq 2 \exp \left(- \frac{x^2}{4(l_2 - l_1 + 1)\sigma^2} \right).$$

Поклавши $x = a_N + \kappa(l_2 - l_1 + 1)$, отримуємо

$$\begin{aligned} p(l_1, l_2, N) &= P \left\{ \sum_{j=l_1}^{l_2} (\phi(\xi_j, h) - \phi(\xi_j, g)) \geq a_N \right\} \\ &\leq 2 \exp \left(- \frac{(a_N + \kappa(l_2 - l_1 + 1))^2}{4(l_2 - l_1 + 1)\sigma^2} \right). \end{aligned}$$

Враховуючи, що для будь-якого $y > 0$

$$\exp \left(- \frac{(a_N + \kappa y)^2}{4\sigma^2 y} \right) \leq \exp \left(- \frac{a_N \kappa}{\sigma^2} \right)$$

маємо $p(l_1, l_2, N) \leq 2 \exp(-a_N \kappa / \sigma^2)$ і

$$\begin{aligned} \bar{p}(N) &= P \left\{ \max_{1 \leq l_1 < l_2 \leq N} \sum_{j=l_1}^{l_2} (\phi(\xi_j, h) - \phi(\xi_j, g)) \geq a_N \right\} \\ &\leq \sum_{1 \leq l_1 < l_2 \leq N} p(l_1, l_2, N) \leq 2N^2 \exp(-a_N \kappa / \sigma^2). \end{aligned}$$

Позначимо $f_N = a_N / \ln N$. За умовою леми, $f_N \rightarrow \infty$, отже

$$\sum_{N=1}^{\infty} \bar{p}(N) \leq \sum_N 2N^2 \exp(-a_N \kappa / \sigma^2) = 2 \sum_N N^{2-f_N \kappa / \sigma^2} < \infty.$$

За теоремою Бореля–Кантеллі отримуємо твердження леми 1. $\square$

Подію B_N визначимо як

$$B_N = \left\{ \max_{h \in K, 1 \leq l_1 < l_2 \leq N} \left| \sum_{j=l_1}^{l_2} \phi(\xi_j, h) - \Phi_j(h) \right| \leq b_N (l_2 - l_1 + 1)^{1/2} \right\}.$$

Лема 2. Нехай $b_N / \ln N \rightarrow \infty$ при $N \rightarrow \infty$. Тоді знайдеться випадковий номер $n = n(\omega) < \infty$ м.н., такий, що для всіх $N > n$ виконується подія B_N .

Доведення. Внаслідок умови (B) та нерівності (4) $\forall x > 0$

$$P \left\{ \left| \sum_{j=l_1}^{l_2} (\phi(\xi_j, h) - \Phi_j(h)) \right| > x \right\} \leq 2 \exp \left(- \frac{x^2}{2(l_2 - l_1 + 1)\sigma^2} \right).$$

Поклавши $x = b_N (l_2 - l_1 + 1)^{1/2}$, отримуємо

$$P \left\{ \left| \sum_{j=l_1}^{l_2} \phi(\xi_j, h) - \Phi_j(h) \right| > b_N (l_2 - l_1 + 1)^{1/2} \right\} \leq 2 \exp(-b_N^2 / 2\sigma^2).$$

Звідси, так само як у доведенні леми 1, отримуємо твердження леми 2. $\square$

Лема 3. Якщо подія A_N відбулась і $a_N < \pi_N / (R + 2)$, то $\hat{R}_N \leq R$.

Доведення. Оскільки $\hat{h} = \arg \min_h J(h)$, то $J(\hat{h}) \leq J(h^0)$. Отже

$$\hat{R}_N \pi_N + \sum_{j=1}^N \phi(\xi_j, \hat{h}_j) \leq R \pi_N + \sum_{j=1}^N \phi(\xi_j, h_j^0),$$

тобто

$$\pi_N (\hat{R}_N - R) \leq \sum_{l=1}^{\hat{R}_N+1} \sum_{j=\hat{k}_{l-1}+1}^{\hat{k}_l} (\phi(\xi_j, h_j^0) - \phi(\xi_j, \hat{h}_j)).$$

Оскільки A_N відбулась, права частина нерівності не перевищує $a_N (\hat{R}_N + 1)$ і

$$\pi_N (\hat{R}_N - R) \leq a_N (\hat{R}_N + 1),$$

звідки, враховуючи умову леми, маємо

$$\hat{R}_N \leq \frac{\pi_N}{\pi_N - a_N} R + \frac{a_N}{\pi_N - a_N} < R + 1.$$

Оскільки $\hat{R}_N$ та R цілі числа, то $\hat{R}_N \leq R$. $\square$

Для довільного j визначимо $l(j)$ як максимум тих l , для яких при всіх i , $j \leq i \leq l+j$, $\hat{h}_i \neq h_i^0$. (Якщо $\hat{h}_j = h_j^0$, то $l(j) = 0$.) Тобто $l(j)$ це довжина інтервалу праворуч від j , на якому $\hat{h}_j$ відрізняється від h_j^0 .

Лема 4. Якщо відбулись події A_N та B_N , то для будь-якого j

$$l(j) \leq \left(\frac{b_N + \sqrt{b^2 + \kappa \pi_N (R+2)}}{\kappa} \right)^2.$$

Доведення. Розглянемо траекторію $h' = \{h'_i\}$, визначену наступним чином: $h'_i = \hat{h}_i$ при $i < j$ та $i > l(j)$, $h'_i = h_i^0$ при $j \leq i \leq l(j)$. Тоді з умови $J(\hat{h}) \leq J(h')$ отримуємо

$$\hat{R}_N \pi_N + \sum_{i=1}^N \phi(\xi_i, \hat{h}_i) \leq R'_N \pi_N + \sum_{i=1}^N \phi(\xi_j, h'_i), \quad (8)$$

де R'_N — кількість змін у траекторії h' , $R'_N \leq 2 + R_N + R \leq 2(R+1)$ за лемою 3. Оскільки $h'_i = \hat{h}_i$ при $i \notin [j, l(j)]$, то з (8) маємо

$$S = \sum_{i=j}^{j+l(j)} \left(\phi(\xi_i, \hat{h}_i) - \phi(\xi_i, h_i^0) \right) \leq (R+2)\pi_N. \quad (9)$$

Представимо S у вигляді

$$S = \sum_{i=j}^{j+l(j)} \left(\phi(\xi_i, \hat{h}_i) - \Phi(\hat{h}_i) \right) + \sum_{i=j}^{j+l(j)} \left(\Phi(\hat{h}_i) - \Phi(h_i^0) \right) + \sum_{i=j}^{j+l(j)} \left(\Phi(h_i^0) - \phi(\xi_i, h_i^0) \right).$$

Оскільки B_N виконано, то перша і друга суми за модулем не перевищують $b_N(l(j))^{1/2}$. Друга сума більше $\kappa l(j)$ за умовою (А). Отже

$$S \geq \kappa l(j) - 2b_N(l(j))^{1/2}. \quad (10)$$

Поклавши $z = (l(j))^{1/2}$ з (9) та (10) отримуємо

$$\kappa z^2 - 2b_N z - (R+2)\pi_N < 0$$

звідки маємо

$$z < \frac{b_N + \sqrt{b_N^2 + \kappa(R+2)\pi_N}}{\kappa}. \quad \square$$

Закінчення доведення теореми. Виберемо a_N та b_N так, щоб

$$a_N / \ln N \rightarrow \infty, \quad a_N / \pi_N \rightarrow 0, \quad b_N^2 / \ln N \rightarrow \infty, \quad b_N^2 / \pi_N \rightarrow 0$$

(це можна зробити внаслідок умови (С)). Тепер, за лемами 1, 2 досить довести (5) та (6) при достатньо великих N за умови виконання A_N та B_N . З леми 4 випливає, що для $l_N = \max_{1 \leq j \leq N} |l(j)|$, $l_N / d_N \rightarrow 0$. Оскільки $d_N / N \rightarrow 0$, то при великих N інтервали $[k_i - d_N, k_i + d_N]$ при різних i не перетинаються. Але, за означенням l_N , для кожного k_i існує $\hat{k}_m$, таке, що

$$|k_i - \hat{k}_m| \leq l_N. \quad (11)$$

Отже в кожному з R інтервалів $[k_i - d_N, k_i + d_N]$ повинно бути хоча б одне $\hat{k}_m$. Тому $\hat{R}_N \geq R$. За лемою 3, $\hat{R}_N \leq R$. Таким чином $\hat{R}_N = R$ і у (11) $m = i$. Враховуючи $l_N / d_N \rightarrow \infty$ з (11) отримуємо (6). $\square$

ЛІТЕРАТУРА

1. Б. С. Дарховский, Б. Е. Бродский, *Асимптотический анализ некоторых оценок в апостериорной задаче о разлажке*, Теор. вероятность. и примен. **35** (1990), № 3, 551-557.
2. L. Horvath, *The maximum likelihood ratio method for testing changes in the parameters of normal observations*, Ann. Statist. **21** (1993), № 2, 671-680.
3. Б. Е. Бродский, Б. С. Дарховский, *Алгоритм апостериорного обнаружения многократных разладок случайной последовательности*, Автоматика и телемеханика (1993), № 1, 62-67.
4. В. В. Моттль, И. Б. Мучник, В. Г. Яковлев, *Оптимальная сегментация экспериментальных кривых*, Автоматика и телемеханика **8** (1983), 84-95.
5. Т. О. Иванова, В. В. Моттль, И. Б. Мучник, *Оценивание параметров скрытых марковских моделей шумоподобных сигналов со скачкообразно изменяющимися вероятностными свойствами. I*, Автоматика и телемеханика (1994), № 9, 75-97.
6. О. В. Сугакова, *Пошук моментів зміни у потоці незалежних спостережень*, Теор. ймовірност. та матем. статист. **55** (1996), 167-172.
7. В. В. Булдыгин, Ю. В. Козаченко, *О субгауссовских случайных величинах*, Украин. матем. журн. **32** (1980), № 6, 723-730.

252127 КИЇВ, ПРОСПЕКТ АКАДЕМІКА ГЛУШКОВА, 6, МЕХАНІКО-МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ, КАФЕДРА ТЕОРІЙ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

КИЇВСЬКИЙ ЕКОНОМІЧНИЙ ІНСТИТУТ МЕНЕДЖМЕНТУ

Надійшла 20.08.97