Marie Vosiková

Řádové odhady veličin $b_n = E[(X_n - \Theta)^2]$ pro Fabianovu stochastickou aproximační metodu

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 7 (1966), No. 1, 81--86

Persistent URL: `http://dml.cz/dmlcz/142190`

1966 ACTA UNIVERSITATIS CAROLINAE-MATHEMATICA ET PHYSICA No 1 Pag 81-86

Řádové odhady veličin $b_n = \mathsf{E}[(X_n - \Theta)^2]$ pro Fabianovu stochastickou aproximační metodu

MARIE VOSIKOVÁ

Katedra matematické statistiky MFF Karlovy university, Praha

(Došlo 21. prosince 1964)

Nechť R je reálná funkce, nechť Θ je hodnota, v níž R nabývá lokálního maxima. V článku jsou za jistých omezujících předpokladů odvozeny řádové odhady veličin $b_n = \mathsf{E}[(X_n - \Theta)^2]$ pro Fabianovu aproximační metodu formulovanou v práci (1).

Učiňme předpoklady:

(a) Nechť a, c, r, η jsou čísla, pro něž

$$0 < a, \quad 0 < c, \quad 0 < r < 1/2, \quad 0 < \eta .$$

(b) Nechť pro každé přirozené číslo n platí

$$a_n = an^{-1/2}, \quad c_n = cn^{-r} .$$

(c) Budiž X_1 libovolná náhodná veličina; pro $n \geq 1$ položme

$$\begin{aligned} X_{n+1} &= X_n + a_n Y_n , \\ \text{kde} \quad Y_n &= V_{n,1} - V_{n,-1} , \\ Y_{n,i} &= V_{n,i} + R(X_n + ic_n) , \\ i &= -1, 1, \end{aligned}$$

kde $V_{1,-1}$, $V_{1,1}$, $V_{2,-1}$, ... jsou nezávislé, identicky rozdělené spojité náhodné veličiny, nechť hustota f náhodných veličin $V_{n,1} - V_{n,-1}$ je spojitá v nule.

Budeme uvažovat jen tu část funkce R – označme ji S, která je rostoucí pro $x < \Theta$ a klesající pro $x > \Theta$. Budiž S definována na $<A, B>$.
Konečně předpokládejme, že $(X_n + a_n Y_n) \in <A, B>$, $(X_n \pm c_n) \in <A, B>$ pro všechna n.

K řádovým odhadům veličin b_n pro Fabianovo aproximační schéma je zcela obdobně jako v práci [2], kde se vyšetřují řádové odhady b_n pro Kiefer-Wolfowitzovu aproximační metodu, použito lemmatu, které odvodil K. L. Chung [3].

Lemma: Nechť $\{b_n\}$ je posloupnost reálných čísel, pro všechna $n \geq n_0$ nechť platí

$$b_{n+1} \leq \left(1 - \frac{c}{n^s}\right) b_n + \frac{c'}{n^t},$$

kde $0 < s < 1$, $s < t$, $c > 0, c' > 0$. Potom

$$\limsup_{n \to \infty} n^{t-s} b_n \leq \frac{c'}{c}.$$

Věta 1: Nechť platí (*a*) (*b*) (*c*). Nechť S je parabola:
$S(x) = -\alpha (x - \Theta)^2 + \beta$, kde $\alpha > 0$, $A < \Theta < B$. Pak

$$b_n = \mathrm{O}\left(\frac{1}{n^{1/2-r}}\right).$$

Důkaz:

Označme

$$\Delta_n = S(X_n + c_n) - S(X_n - c_n);$$

f budiž hustota $V_{n,1} - V_{n,-1}$; označme $G(x) = \int_0^x f(v)\,\mathrm{d}v$.

Je

$$\mathbf{E}[Y_n/X_n] = 2\int_0^{\Delta_n} f(v)\,\mathrm{d}v = 2G(\Delta_n).$$

Z (*c*) plyne

(1) $$\mathbf{E}[(X_{n+1} - \Theta)^2/X_n] = (X_n - \Theta)^2 + a_n^2 + 4a_n(X_n - \Theta)\,G(\Delta_n).$$

Předpokládáme

$$S(x) = -\alpha(x - \Theta)^2 + \beta, \quad \alpha > 0, \quad A < \Theta < B.$$

Odtud $\Delta_n = -4\,\alpha\, c_n\,(X_n - \Theta)$;

(2) $$G(\Delta_n) = \Delta_n\, G'(\vartheta\Delta_n) = \Delta_n f(\vartheta\Delta_n).$$

Vzhledem k tomu, že $|X_n - \Theta| < B - A$, platí $\Delta_n \to 0$ pro $n \to \infty$ a tedy ze spojitosti f v bodě nula, $f(\vartheta\Delta_n) \to f(0)$. Protože veličiny $V_{n,1}$ a $V_{n,-1}$ jsou spojité, nezávislé a stejně rozložené, je $f(0)$ kladné.
Z toho plyne

(3) $$G(\Delta_n) = -4\,\alpha\, c_n(X_n - \Theta)\{f(0) + \vartheta_{1,n}\,\varepsilon\},$$

kde $|\vartheta_{1,n}| \leq 1$ pro všechna $n > n_0(\varepsilon)$ (a pro libovolné $\varepsilon > 0$).

Dosadíme (3) do (1):

$$\mathbf{E}[(X_{n+1} - \Theta)^2 / X_n] = (X_n - \Theta)^2 [1 - 16\,\alpha\, a_n\, c_n\{f(0) + \vartheta_{1,n}\,\varepsilon\}] + a_n^2.$$

Odtud

$$b_{n+1} \leq b_n \left[1 - \frac{16\,\alpha\, ac\{f(0) - \varepsilon\}}{n^{1/2+r}}\right] + \frac{a^2}{n}$$

pro $n > n_0(\varepsilon)$. Užitím lemmatu dostaneme

$$b_n = \mathrm{O}\left(\frac{1}{n^{1/2-r}}\right),$$

což jsme měli dokázat.

Věta 2: Nechť platí (*a*) až (*e*) a pro S budiž:

(I) $$K_0 \,|\, x - \Theta \,| \leq |\, S'(x) \,| \leq K_1 \,|\, x - \Theta \,|\,, \quad K_0 > 0\,, \quad K_1 > 0\,,$$

(II) $$S^{(2k+1)}(\Theta) = 0\,, \quad |\, S^{(2k+2)}(x) \,| \leq M^{2k+1} (2k+1)\,!$$

pro $A \leq x \leq B$, $k = 1, 2, \ldots$ a nějaké $M > 0$. Pak je

$$b_n = \mathrm{O}\left(\frac{1}{n^{1/2-r}}\right).$$

Důkaz:

Budiž

$$0 < c < \frac{1}{M\sqrt{2}}\,.$$

Potom

$$S(x + c) = \sum_{k=0}^{\infty} \frac{c^k}{k!} S^{(k)}(x)\,,$$

$$S(x - c) = \sum_{k=0}^{\infty} (-1)^k \frac{c^k}{k!} S^{(k)}(x)$$

jsou absolutně konvergentní řady a tedy

$$\Delta = S(x + c) - S(x - c) = 2 \sum_{j=0}^{\infty} \frac{c^{2j+1}}{(2j+1)!} S^{(2j+1)}(x)\,.$$

Označme $\varkappa_\Theta(x) = -\dfrac{S'(x)}{x - \Theta}$ pro $x \neq \Theta$ a $\varkappa_\Theta(\Theta) = K_0$.

Ze (I) plyne

(4) $$K_0 \leq \varkappa_\Theta(x) \leq K_1$$

$$\Delta = \left\{-2\,\varkappa_\Theta(x)\, c + 2 \sum_{j=1}^{\infty} \frac{c^{2j+1}}{(2j+1)!} S^{(2j+2)}\left[\Theta + \vartheta_j\,(x - \Theta)\right]\right\}(x - \Theta)\,,$$

(5) $$2\sum_{j=1}^{\infty} < 2 \sum_{j=1}^{\infty} c^{2j+1} M^{2j+1} = 2\,\frac{c^3 M^3}{1 - c^2 M^2} < 4 M^3 c^3$$

Ze (4) a (5) plyne

$$(X_n - \Theta)\, \Delta_n \leq -(2 - \eta)\, K_0\, c_n\, (X_n - \Theta)^2\,,$$

kde $0 < \eta < 2$ pro $n > n_0(\eta)$.

Je [viz (1) a (2)]:

$$\mathbf{E}[(X_{n+1} - \Theta)^2/X_n] = (X_n - \Theta)^2 + a_n^2 + 4a_n(X_n - \Theta)\Delta_n f(\vartheta\, \Delta_n).$$

Poněvadž $|X_n - \Theta| < B - A$, platí $\Delta_n \to 0$, tedy $f(\vartheta\, \Delta_n) \to f(0)$.

Dále

$$\mathbf{E}[(X_{n+1} - \Theta)^2/X_n] = (X_n - \Theta)^2 + a_n^2 + 4a_n(X_n - \Theta)\, \Delta_n \{f(0) + \vartheta_{1,\, n}\, \varepsilon\} \leq$$

$$\leq (X_n - \Theta)^2 + a_n^2 - 4(2 - \eta)\, K_0\, a_n\, c_n \{f(0) - \varepsilon\}\, (X_n - \Theta)^2 \leq$$

$$\leq (X_n - \Theta)^2 \left(1 - \frac{4(2 - \eta)\, K_0 \{f(0) - \varepsilon\}\, ac}{n^{1/2 + r}}\right) + \frac{a^2}{n}.$$

Z toho

$$b_{n+1} \leq b_n \left(1 - \frac{4(2 - \eta)\, K_0 \{f(0) - \varepsilon\}\, ac}{n^{1/2 + r}}\right) + \frac{a^2}{n}.$$

Použitím uvedeného lemmatu plyne

$$b_n = \mathrm{O}\left(\frac{1}{n^{1/2 - r}}\right),$$

čímž je důkaz proveden.

Věta 3: Nechť platí (*a*) (*b*) (*c*). Nechť pro všechna $x \in \langle A, B \rangle$ je $|S'''(x)| < M_1$ a $|S'(x)| \geq K_0 |x - \Theta|$, kde K_0, M_1 jsou kladné konstanty. Pak

$$b_n = \mathrm{O}\left(\frac{1}{n^{1/2 - r}}\right) \quad \text{pro } r \geq \frac{1}{6},$$

$$= \mathrm{O}\left(\frac{1}{n^{2r}}\right) \quad \text{pro } r < \frac{1}{6}.$$

Důkaz:

Taylorovým rozvojem v bodě X_n dostaneme

$$S(X_n + c_n) = S(X_n) + S'(X_n)c_n + \frac{1}{2!} S''(X_n)\, c_n^2 +$$

$$+ \frac{S'''(X_n + \vartheta_1\, c_n)}{3!}\, c_n{}^3.$$

Položme

$$\frac{S'''(X_n + \vartheta_1\, c_n)}{3!}\, c_n^3 = Z_1.$$

Je

$$|Z_1| \leq \frac{1}{6} M_1\, c_n^3.$$

Obdobně pro

$$S(X_n - c_n).$$

Pak

$$\Delta_n = 2S'(X_n)\, c_n + \mathrm{O}(c_n^3).$$

Odtud a z předpokladu

$$| S'(x) | \geq K_0 \, | x - \Theta |$$

plyne

(6) $$(X_n - \Theta) \, \Delta_n \leq - 2 \, K_0 (X_n - \Theta)^2 \, c_n + O(c_n^3) \, .$$

Z toho, že $| X_n - \Theta | < B - A$ opět plyne $\Delta_n \to 0$

a tedy

$$f(\vartheta \, \Delta_n) \to f(0) \, .$$

Dosazením (6) do (2) dostaneme:

(7) $$(X_n - \Theta) \, G(\Delta_n) \leq - 2 \, K_0 (X_n - \Theta)^2 \, \{ f(0) - \varepsilon \} \, c_n + O(c_n^3) \, .$$

Dosazením (7) do (1) obdržíme

$$\mathbf{E}[(X_{n+1} - \Theta)^2 / X_n] \leq (X_n - \Theta)^2 + a_n^2 - 8 a_n \, K_0 (X_n - \Theta)^2 \, \{ f(0) - \varepsilon \} \, c_n + \\ + O(c_n^3) \, a_n \, ,$$

$$\mathbf{E}[(X_{n+1} - \Theta)^2 / X_n] \leq \left(1 - \frac{8ac \, \{ f(0) - \varepsilon \} \, K_0}{n^{1/2 + r}} \right) (X_n - \Theta)^2 + \\ + O\left(\frac{1}{n^{3r + 1/2}} \right) + \frac{a^2}{n} \, ,$$

$$b_{n+1} \leq b_n \left(1 - \frac{8ac \, \{ f(0) - \varepsilon \} \, K_0}{n^{1/2 + r}} \right) + O\left(\frac{1}{n^{3r + 1/2}} \right) + \frac{a^2}{n} \, .$$

Předchozí vztah splňuje předpoklady lemmatu. Dostáváme požadované tvrzení

$$b_n = O\left(\frac{1}{n^{1/2 - r}} \right) \quad \text{pro} \quad r \geq \frac{1}{6} \, ,$$

$$= O\left(\frac{1}{n^{2r}} \right) \quad \text{pro} \quad r < \frac{1}{6} \, .$$

Literatura

1. V. FABIAN: A new one-dimensional stochastic aproximation method for finding a local minimum of a function, Trans. third Prague conference on information theory, statistical decision function, random processes (1964).
2. V. DUPAČ: O Kiefer-Wolfowitzově aproximační metodě, Časopis pro pěstování matematiky, 82 (1957), 47-75.
3. K. L. CHUNG: On a stochastic approximation method, Annals of Mathematical Statistics 25 (1954), 463-483.

Оценка порядка величин $b_n = \mathbf{E}[(X_n - \Theta)^2]$ для стохастического апроксимационного метода В. Фабиана.

(Резюме)

В работе выведены оценки порядка величин $b_n = \mathbf{E}[(X_n - \Theta)^2]$ для стохастического апроксимационного метода В. Фабиана приведенного в [1]. Пусть X_1 любая случайная величина и для $n \geq 1$ определяем $X_{n+1} = X_n + a_n \, Y_n$, где a_n и y_n удовлетворяет условиям (a),

(b), (c). Θ — это значение, в котором функция S, определенная для $x \in \langle A, B \rangle$, приобретает свой локальный максимум.

Для получения оценки порядка b_n пользуемся леммой К. Л. Чжуна [3]. Если $S(x) = -\alpha(x - \Theta)^2 + \beta$, где $\alpha > 0$, $A < \Theta < B$, то $b_n = O\left(\frac{1}{n^{1/2 - r}}\right)$, $0 < r < \frac{1}{2}$. Если условия (I) и (II), то снова $b_n = O\left(\frac{1}{n^{1/2 - r}}\right)$. Если для всех $x \in \langle A, B \rangle$ $|S'''(x)| < M_1$ и $|S'(x)| \geq K_0 |x - \Theta|$, где K_0 и M_1 положительные постоянные, то

$$b_n = O\left(\frac{1}{n^{1/2 - r}}\right) \quad \text{для} \quad r \geq \frac{1}{6},$$

$$= O\left(\frac{1}{n^{2r}}\right) \quad \text{для} \quad r < \frac{1}{6}.$$

The Order Estimates of $b_n = \mathbf{E}[(X_n - \Theta)^2]$ *for Fabian's Stochastic Approximation Method.*

(Summary)

In the paper the order estimates of $b_n = \mathbf{E}[(X_n - \Theta)^2]$ for Fabian's stochastic approximation method [1] are given. Let X_1 be an arbitrary random variable and for $n \geq 1$ define $X_{n+1} = X_n + a_n Y_n$, where a_n and Y_n satisfy the assumptions (a) (b) (c). Let Θ be the value, in which the functions S, defined on $\langle A, B \rangle$ achieves its maximum. The order estimates of b_n are derived with help of a lemma of K. L. Chung [3].

If $S(x) = -\alpha(x - \Theta)^2 + \beta$ where $\alpha > 0$, $A < \Theta < B$, then $b_n = O\left(\frac{1}{n^{1/2 - r}}\right)$, $0 < r < \frac{1}{2}$. When (I) and (II) are valid, the same estimate holds true. If for all $x \in \langle A, B \rangle$, $|S'''(x)| < M_1$ and $|S'(x)| \geq K_0 |x - \Theta|$ where K_0 and M_1 are positive constants, then

$$b_n = O\left(\frac{1}{n^{1/2 - r}}\right) \quad \text{for} \quad r \geq \frac{1}{6},$$

$$= O\left(\frac{1}{n^{2r}}\right) \quad \text{for} \quad r < \frac{1}{6}.$$