ՀԱՅԱՍՏԱՆԻ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱՅԻ ԶԵԿՈՒՅՑՆԵՐ
ДОКЛАДЫ НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК АРМЕНИИ
Том 94 1993 № 2

МАТЕМАТИКА

УДК 519.21

К. В. Гаспарян

О процедурах стохастической аппроксимации. Мартингальный подход

(Представлено академиком НАН Армении Р. В. Амбарцумяном 10.III 1992)

Хорошо известны (см., например, ([1])) стохастические процедуры Роббинса—Монро и Кифера—Вольфовица для нахождения корня уравнения регрессии и экстремума функции регрессии. В данной работе рассматриваются общие модели этого типа и получаются оценки неизвестных параметров по семимартингальным наблюдениям. В отличие от работы ([2]), где рассмотрены аналогичные вопросы, стохастический базис здесь предполагается произвольным, а рассматриваемые на нем процессы—нерегулярными (не принадлежат пространству Скорохода càdlàg-процессов, т. е. с непрерывными справа и имеющими пределы слева траекториями).

При получении оценок используются мартингальные методы, введенные в данной общей ситуации Л. И. Гальчуком в ([3]) и развитые далее автором в ([4, 5])):

Введем следующие обозначения (см. ([3])):

P и O — σ-алгебры предсказуемых и опциональных множеств (а также множества соответствующих измеримых процессов);

P_s — множество строго предсказуемых процессов $(a_t)_{t\geq 0}\in P_s$, если $(a_t)_{t>0}\in \mathrm{P}$ и $(a_{t+})_{t>0}\in O$);

$\mathrm{A}_{\mathrm{loc}}$, V^+ и $\mathrm{A}^+_{\mathrm{loc}}$ — множества процессов локально интегрируемой вариации, возрастающих и локально интегрируемых возрастающих процессов;

$\mathrm{M}_{\mathrm{loc}}$ и $\mathrm{M}^2_{\mathrm{loc}}$ — множества опциональных локальных и локально квадратично интегрируемых мартингалов;

S — множество опциональных семимартингалов.

Пусть имеется некоторый стохастический базис $(\Omega,\ \mathrm{F},\ (\mathrm{F}_t)_{t\geq 0},\ \mathrm{P})$, где F — полная σ-алгебра (по вероятностной мере P), $(\mathrm{F}_t)_{t\geq 0}$ — произвольный неубывающий поток σ-алгебр ($\mathrm{F}_s\subseteq \mathrm{F}_t\subseteq \mathrm{F}$, $s<t$). Траектории всех рассматриваемых на нем процессов $X=(X_t)_{t\geq 0}$ предполагаются имеющими двусторонние пределы X_{t-} и X_{t+} для любого $t\geqslant 0$.

Любой процесс $a = (a_t)_{t\geq 0} \in A_{loc}$ имеет представление

$$a = a^r + a^g,$$

где $a^r = a^c + \sum_{s<\cdot} \Delta a_s \in A_{loc}$, $a^g = \sum_{s<\cdot} \Delta^+ a_s \in A_{loc}$, $a^c \in A_{loc}$ — процесс с непрерывными траекториями, $\Delta a_s = a_s - a_{s-}$, $\Delta^+ a_s = a_{s+} - a_s$, $s \geq 0$.

Всякий процесс $m = (m_t)_{t\geq 0} \in M_{loc}$ допускает (см. ([3])) разложение (единственное с точностью до неотличимости)

$$m = m^r + m^g,$$

где $m^r = m^c + m^d$, $m^c \in M_{loc}$ — непрерывный локальный мартингал, m^d и $m^g \in M_{loc}$ — „чисто разрывные" càdlàg и càglàd локальные мартингалы, соответственно.

Для процессов $X, Y \in S$ определим квадратическую ковариацию

$$[X, Y]_t = \langle X^c, Y^c \rangle_t + \sum_{s\leq t} \Delta X_s \Delta Y_s + \sum_{s<t} \Delta^+ X_s \Delta^+ Y_s,$$

где $\langle X^c, Y^c \rangle$ — взаимная квадратическая характеристика непрерывных мартингальных составляющих X^c и Y^c процессов X и Y.

1. *Процедура Роббинса—Монро*:

Пусть процесс $X = (X_t)_{t\geq 0}$ удовлетворяет стохастическому уравнению

$$X_t = -\int\limits_{]0,t]} \gamma_s^{(1)} dY_s^r - \int\limits_{[0,t[} \gamma_s^{(2)} dY_{s+}^g, \tag{1}$$

где

$$Y_t^r = \int\limits_{]0,t]} R^1(X_{s-})\, da_s^r + m_t^r, \qquad Y_t^g = \int\limits_{[0,t[} R^2(X_s)\, da_{s+}^g + m_t^g,$$

$Y = Y^r + Y^g \in S$ — наблюдаемый процесс, $m = m^r + m^g \in M_{loc}$, $a = a^r + a^g \in A_{loc}^+$, $R^i(x) = \beta_i(x - \theta)$, $\beta_i \in R_+ \setminus \{0\}$, $i = 1, 2$, $x \in R^1$, $\theta \in R^1$ — неизвестный оцениваемый параметр, положительные функции $\gamma^{(1)} \in P$ и $\gamma^{(2)} \in O$ такие, что

$$(\gamma^{(1)} \cdot a_t^r) = \Bigl(\int\limits_{]0,t]} \gamma_s^{(1)} da_s^r\Bigr) \in A_{loc}^+, \quad (\gamma^{(2)} \cdot a_t^g) = \Bigl(\int\limits_{[0,t[} \gamma_s^{(2)} da_{s+}^g\Bigr) \in A_{loc}^+.$$

Условие (A):

$$\gamma^{(1)} \cdot a_\infty^r = \gamma^{(2)} \cdot a_\infty^g = \infty, \quad |\gamma^{(1)}|^2 \cdot a_\infty^r < \infty, \quad |\gamma^{(2)}|^2 \cdot a_\infty^g < \infty \text{ п. н.}$$

Имеет место следующая (ср. ([2]))

Теорема 1. *Пусть* $m \in M_{loc}^2$ *и для всех* $t \geq 0$ *п. н. имеем*

a) $\dfrac{d\langle m^r \rangle_t}{da_t^r} \leq \xi^1$, $\dfrac{d\langle m^g \rangle_t}{da_t^g} \leq \xi^2$ — *где* $\langle m^r \rangle \in P \cap A_{loc}^+$ *и* $\langle m^g \rangle \in P_s \cap A_{loc}^+$ — *квадратические характеристики* $m^r \in M_{loc}^2$ *и* $m^g \in M_{loc}^2$, $0 < \xi^i < \infty$ *п. н. — случайные величины*, $i = 1, 2$;

b) $0<\eta^1<1-\bar\beta_1\gamma_t^{(1)}\Delta a_t$, $0<\eta^2<1-\bar\beta_2\gamma_t^{(2)}\Delta^+ a_t$, *где* $0<\beta_i<$ $<\bar\beta_i<\infty$, $\bar\beta_i$ — *заданные константы, а* $\eta^i<\infty$ *п. н.—с. в.,* $i=1, 2$. *Тогда при выполнении условия* (A) *имеем*

$$X_t\to\theta \text{ п. н. при } t\to\infty.$$

Доказательство. Запишем уравнение (1) в виде ($t\geqslant 0$)

$$X_t-\theta=N_t-\theta+\int_{]0,t]}(X_s-\theta)\,dZ_s^r+\int_{]0,t[}(X_s-\theta)\,dZ_{s+}^g, \tag{2}$$

где

$$N_t=N_t^r+N_t^g,\quad N_t^r=-\gamma^{(1)}\cdot m_t^r=-\int_{]0,t]}\gamma_s^{(1)}dm_s^r,$$

$$N_t^g=-\gamma^{(2)}\cdot m_t^g=-\int_{]0,t[}\gamma_s^{(2)}dm_{s+}^g\in M_{loc}^2$$

(определение интегралов см. в ([3])),

$$Z_t^r=-\beta_1(\gamma^{(1)}\cdot a_t^r),\quad Z_t^g=-\beta_2(\gamma^{(2)}\cdot a_t^g).$$

Тогда, так как для любого $t\geqslant 0$ имеем

$$\Delta Z_t>-1,\quad \Delta^+Z_t>1 \text{ и}$$

$$\sum_{s<t}\frac{|\Delta N_s\Delta Z_s|}{1+\Delta Z_s}\leqslant\frac{\bar\beta_1}{\eta^1}[N^r,N^r]_t^{1/2}(\gamma^{(1)}\cdot a_t^r)<\infty,$$

$$\sum_{s>t}\frac{|\Delta^+N_s\Delta^+Z_s|}{1+\Delta^+Z_s}\leqslant\frac{\bar\beta_2}{\eta^2}[N^g,N^g]_t^{1/2}(\gamma^{(2)}\cdot a_t^g)<\infty,$$

то (см. ([4]), теорема 3.3.2) имеем

$$X_t-\theta=B_t^{-1}W_t,$$

где $B_t=\varepsilon(Z)_t^{-1}$, $\varepsilon(Z)_t$ — экспонента Долеан, т. е. решение уравнения

$$Y_t=1+Y_-\cdot Z_t^r+Y\cdot Z_t^g \text{ (см. ([6])), а}$$

$$W_t=W_t^r+W_t^g,\quad W_t^r=B_-\cdot\left(N^r-\frac{[Z^r,N^r]}{1+\Delta Z}\right)_t,$$

$$W_t^g=B\cdot\left(N^g-\frac{[Z^g,N^g]}{1+\Delta^+Z}\right)_t.$$

Далее из представления (см. ([6]))

$$\varepsilon(Z)_t=\exp Z_t\prod_{s<t}(1+\Delta Z_s)e^{-\Delta Z_s}\prod_{s<t}(1+\Delta^+Z_s)e^{-\Delta^+Z_s}$$

следует, что $\varepsilon(Z)>0$ п. н., $B\in V^+$ и, согласно условию (A), имеем $B_\infty=\infty$ п. н.

Кроме того, так как

$$\varepsilon(Z)_t = \varepsilon(Z)_{t-}(1-\beta_t\gamma_t^{(1)}\Delta a_t), \qquad \varepsilon(Z)_{t+} = \varepsilon(Z)_t(1-\beta_2\gamma_t^{(2)}\Delta^+ a_t),$$

то $\varepsilon(Z)\in \mathrm{P}$, $\varepsilon(Z)_+\in O$, т. е. $\varepsilon(Z)\in \mathrm{P}_s$.

С другой стороны, имеем

$$B_-\cdot\frac{[Z^r, N^r]}{1+\Delta Z} = B\cdot[Z^r, N^r]\in \mathrm{M}^2_{\mathrm{loc}},$$

$$B\cdot\frac{[Z^g, N^g]}{1+\Delta^+ Z} = B_+\cdot[Z^g, N^g]\in \mathrm{M}^2_{\mathrm{loc}},$$

так как

$$Z^r\in \mathrm{P}\cap V, \quad Z^g\in \mathrm{P}_s\cap V \quad \text{и} \quad [Z^r, N^r] = (\Delta Z)\cdot N^r, \quad [Z^g, N^g] = (\Delta^+ Z)\cdot N^g$$

(см. [7]). Таким образом $W\in \mathrm{M}^2_{\mathrm{loc}}$.

Рассмотрим теперь следующий процесс:

$$\overline{Y}_t = (1+B)^{-1}\cdot W_t^r + (1+B_+)^{-1}\cdot W_t^g\in \mathrm{M}^2_{\mathrm{loc}}.$$

Отсюда, согласно условию (A), имеем

$$\langle \overline{Y}^r\rangle_\infty \leqslant 2(1+B)^{-2}(B_-^2 + B^2(\Delta Z)^2)\cdot\langle N^r\rangle_\infty \leqslant 4\langle N^r\rangle_\infty < \infty. \quad (3)$$

Аналогично

$$\langle \overline{Y}^g\rangle_\infty \leqslant 2(1+B_+)^{-2}(B^2 + B_+^2(\Delta^+ Z)^2)\cdot\langle N^g\rangle_\infty \leqslant 4\langle N^g\rangle_\infty < \infty. \quad (4)$$

Таким образом

$$B\in V^+\cap \mathrm{P}_s, \quad B_\infty = \infty \quad \text{и} \quad \langle \overline{Y}\rangle_\infty < \infty \text{ п. н.}$$

Отсюда следует (см. [5]), что

$$B_t^{-1} W_t \to 0 \text{ п. н. при } t\to\infty.$$

2. *Процедура Кифера—Вольфовица*:

Пусть процесс $X=(X_t)_{t\geqslant 0}$ удовлетворяет стохастическому уравнению

$$X_t = -\int_{]0,t]}(4c_s^{(1)})^{-1}\gamma_s^{(1)}\,dY_s^r - \int_{[0,t[}(4c_s^{(2)})^{-1}\gamma_s^{(2)}\,dY_{s+}^g, \quad (5)$$

где

$$Y_t^r = \int_{]0,t]}[J^1(X_{s-}+c_s^{(1)}) - J^1(X_{s-}-c_s^{(1)})]\,da_s^r + m_t^r,$$

$$Y_t^g = \int_{[0,t[}[J^2(X_s+c_s^{(2)}) - J^2(X_s-c_s^{(2)})]\,da_{s+}^g + m_t^g,$$

$Y = Y^r + Y^g\in S$ — наблюдаемый процесс, $m = m^r + m^g\in \mathrm{M}_{\mathrm{loc}}$, $a = a^r + a^g\in \mathrm{A}^+_{\mathrm{loc}}$, $J^i(x) = \beta_i(x-\theta)^2$ — функция регрессии, $\beta_i\in R_+\setminus\{0\}$.

$i=1, 2$, $x\in R^1$, $\theta\in R^1$ — неизвестный параметр, положительные функции $\gamma^{(1)}$, $c^{(1)}\in \mathrm{P}$ и $\gamma^{(2)}$, $c^{(2)}\in O$ такие, что $\gamma^{(1)}\cdot a^r$, $\gamma^{(2)}\cdot a^g\in \mathrm{A}^+_{loc}$, $\gamma^{(1)}_t\downarrow 0$, $c^{(1)}_t\downarrow 0$ п. н. при $t\to\infty$.

Условие (B):

$$\gamma^{(1)}\cdot a^r_- = \gamma^{(g)}\cdot a^g_- = \infty,\ |\gamma^{(1)}(c^{(1)})^{-1}|^2\cdot a^r_- < \infty,\ |\gamma^{(2)}(c^{(2)})^{-1}|^2\cdot a^g_- < \infty \text{ п. н.}$$

Теорема 2 (ср. ([2])). *Пусть* $m\in \mathrm{M}^2_{loc}$ *и выполняются условия* a) и b) *теоремы* 1 *и условие* (B). *Тогда*

$$X_t\to\theta \text{ п. н. при } t\to\infty.$$

Доказательство. Запишем уравнение (5) в виде (2), где

$$N^r_t = -(4c^{(1)})^{-1}\gamma^{(1)}\cdot m^r_t,\quad N^g_t = -(4c^{(2)})^{-1}\gamma^{(2)}\cdot m^g_t,$$

$$Z^r_t = -\beta_1(\gamma^{(1)}\cdot a^r_t),\quad Z^g_t = -\beta_2(\gamma^{(2)}\cdot a^g_t).$$

Далее доказательство проводится так же, как в теореме 1.

Замечание 1. Предположим, что в моделях 1 и 2 имеем $m\in \mathrm{M}_{loc}$. Тогда условие a) в теоремах 1 и 2 надо заменить на следующее:

$$\frac{d[m^r, m^r]_t}{da^r_t}\leqslant \xi^1<\infty,\qquad \frac{d[m^g, m^g]_t}{da^g_t}\leqslant \xi^2<\infty\quad (t>0). \tag{6}$$

Для модели 1 согласно условию (A) и (6) получим по аналогии с (3) и (4)

$$[\overline{Y}^r, \overline{Y}^r]_- \leqslant 4[N^r, N^r]_- < \infty,\quad [\overline{Y}^g, \overline{Y}^g]_- \leqslant 4[N^g, N^g]_- < \infty.$$

С другой стороны, в силу условия (6) для любого F-момента остановки T и F_+-момента остановки U имеем $(F_+ = (\mathrm{F}_{t+})_{t\geq 0}$, $\mathrm{F}_{t+} = \bigcap\limits_{s>t}\mathrm{F}_s)$

$$E|\Delta W_T|I_{T<\infty}\leqslant EB_T|\Delta N_T|I_{T<\infty}\leqslant (EB^2_T)^{1/2}(E[N^r, N^r]^2_T)^{1/2}I_{T<\infty}<\infty,$$

$$E|\Delta^+ W_U|I_{U<\infty}\leqslant EB_U|\Delta^+ N_U|I_{U<\infty}\leqslant (EB^2_U)^{1/2}(E[N^g, N^g]^2_U)^{1/2}I_{U<\infty}<\infty,$$

так как процессы B^2 $[N^r, N^r]$, $[N^g, N^g]\in \mathrm{P}_s\cap \mathrm{V}^+ = \mathrm{P}_s\cap \mathrm{A}^+_{loc}$ (см. ([6])).

Отсюда следует (см. ([5]))

$$X_t\to\theta \text{ п. н. при } t\to\infty.$$

Аналогично проводится доказательство и для модели 2.

Замечание 2. Рассмотрим дискретную процедуру Роббинса—Монро (см. ([1]))

$$x_n = x_{n-1} - \gamma_n y_n,\quad n\geqslant 1,$$

где $y_n = \beta(x_{n-1}-\theta)+\varepsilon_n$, $(\varepsilon_n, \mathrm{F}_n)_{n\geq 1}$ — мартингал-разность с $E(|\varepsilon_n|^2|\mathrm{F}_{n-1})\leqslant\xi<\infty$, последовательность $(\gamma_n, \mathrm{F}_{n-1})_{n\geq 1}$ такая, что

$$0 < \gamma_n < \hat{\beta}^{-1} \quad \text{и} \quad \sum_{n=1}^{\infty} \gamma_n = \infty, \quad \sum_{n=1}^{\infty} \gamma_n^2 < \infty \quad \text{п. н.}$$

Применяя тогда теорему 1 с $a_t = [t]$ и $F_t = F_{[t]}$, получим

$$x_n \to \theta \ \text{п. н.} \quad \text{при} \quad n \to \infty.$$

Ереванский государственный университет

Կ. Վ. ԳԱՍՊԱՐՅԱՆ

Ստոխաստիկ ապրոկսիմացիայի պրոցեդուրաների մասին: Մարտինգալային մոտեցում

Հոդվածում ստացված են անհայտ պարամետրի գնահատականները Ռոբինս-Մոնրոյի և Կիֆեր-Վոլֆովիցի պրոցեդուրաների համար ըստ սեմիմարտինգալային տիպի դիտարկումների։

Ստոխաստիկ բազիսը այստեղ ենթադրվում է կամայական, իսկ բոլոր դիտարկվող պրոցեսները ոչ ռեգուլյար (երկկողմանի սահմաններ ունեցող հետագծերով)։

ЛИТЕРАТУРА—ԳՐԱԿԱՆՈՒԹՅՈՒՆ

[1] *М. Б. Невельсон, Р. З. Хасьминский.* Стохастическая аппроксимация и рекуррентное оценивание, Наука, М., 1983. [2] *А. В. Мельников,* УМН, т. 43, № 4 (1988). [3] *Л. И. Гальчук.* Мат. сб., т. 112 (154), № 4(8) (1980). [4] *К. В. Гаспарян,* Канд. дис., Ереван, 1988. [5] *К. В. Гаспарян,* 5 Междунаролн. Вильнюсская конфер. по теории вероятн. и мат. статист. Тезисы докладов, т. I (1989). [6] *Л. И. Гальчук.* Теория вероятн. и ее примен., т. 29, № 1 (1984). [7] *К. В. Гаспарян,* в. кн.: Труды ВЦ АН АрмССР и ЕГУ, т. 15 (1988).