УДК 519.21

В. А. Коваль, канд. физ.-мат. наук (Терноп. приборостроит. ин-т)

О СКОРОСТИ СХОДИМОСТИ ПРОЦЕДУР СТОХАСТИЧЕСКОЙ АППРОКСИМАЦИИ

The rate of convergence of a linear stochastic approximation procedure in R^d is studied with sufficiently general assumptions on the coefficients of the equation.

Досліджується швидкість збіжності лінійної процедури стохастичної апроксимації в R^d при досить загальних припущеннях відносно коефіцієнтів рівняння.

Введение. Рассмотрим в d-мерном евклидовом пространстве R^d с евклидовой нормой $\|\cdot\|$ процедуру стохастической аппроксимации

$$X_n = X_{n-1} - a_n A X_{n-1} + B_n V_n, \quad n = 1, 2, \ldots, \tag{1}$$

где X_0, (V_n) — случайные векторы со значениями в R^d; (a_n) — последовательность положительных чисел таких, что $\lim\limits_{n\to\infty} a_n = 0$, $\sum_{n=1}^{\infty} a_n = \infty$; (B_n) — последовательность неслучайных матриц размера $d \times d;$ A — матрица размера $d \times d$, собственные значения которой $\lambda_1, \ldots, \lambda_s$ удовлетворяют условию $\operatorname{Re}\lambda_i > 0$, $i = 1, \ldots, s$. Наиболее полные результаты, связанные с исследованием скорости сходимости процедуры (1), получены в работах [1, 2]. При этом предполагалось, то последовательности (a_n) и (B_n) имеют вид

$$a_n = n^{-a}, \quad B_n = n^{-b} I, \quad n \geq 1, \quad 0 < a \leq 1, \quad a/2 < b, \tag{2}$$

I — единичная матрица. В настоящей работе скорость сходимости процедуры (1) исследуется для достаточно широких классов последовательностей (a_n) и (B_n), удовлетворяющих условиям, которые являются обобщением условий Льюнга [3].

2. Формулировка основных результатов. Относительно последовательности (B_n) в дальнейшем в некоторых утверждениях будет предполагаться выполненным условие

$$\limsup_{n\to\infty} \left(\rho_{\max}(B_n)/\rho_{\min}(B_n)\right) < \infty, \tag{3}$$

где $\rho_{\max}(\cdot)$ и $\rho_{\min}(\cdot)$ — максимальное и минимальное сингулярные числа матрицы. Данное условие будет, например, выполненным, если $B_n = b_n B$, $n \geq$ ≥ 1, где $(b_n) \subset R^1$, а матрица B — невырождена.

Рассмотрим сначала процедуру (1) при гауссовских возмущениях (V_n). Последовательность независимых гауссовских $N(0, I)$-распределенных случайных векторов в R^d будем называть *стандартной гауссовской последовательностью* (с.г.п.).

Теорема 1. *Пусть в уравнении* (1) (V_n) — *с.г.п., выполнено условие* (3) *и следующие условия*:

$$\|B_n\|^2 / a_n \geq \|B_{n+1}\|^2 / a_{n+1}, \quad n \geq n_0 \geq 1; \tag{4}$$

для некоторого $\nu \in [1, 2)$

$$\|B_{n+1}\|^2 / a_{n+1} \geq (\|B_n\|^2 / a_n)(1 - \sigma a_{n+1})^{\nu}, \quad n \geq n_0, \tag{5}$$

где

$$\sigma = \min_{1 \leq i \leq s} \operatorname{Re}\lambda_i > 0.$$

Тогда

$$\limsup_{n\to\infty}\left(\left(\|B_n\|^2/a_n\right)\log\left(\sum_{j=1}^{n}a_j\right)\right)^{-1/2}\|X_n\| = \alpha \quad п.н., \tag{6}$$

где α *— некоторая неслучайная константа из интервала* $(0,+\infty)$.

Рассмотрим процедуру (1) при субгауссовских возмущениях (V_n). Дадим следующее определение.

Определение 1. *Случайный вектор* $V=(v_1,\ldots,v_d)^{\mathrm{T}}$ (т — *знак транспонирования) называется стандартным субгауссовским вектором, если выполнены следующие условия:* 1) $v_1,\ldots,v_d$ *— независимые субгауссовские случайные величины с гауссовскими штандартами* $\tau(v_1),\ldots,\tau(v_d)$; 2) $Ev_j=0$, $Ev_j^2=1$, $j=1,\ldots,d$ (E *— знак математического ожидания*).

Положим $\tau=\max\limits_{1\le j\le d}\tau(v_j)$.

Определение 2. *Последовательность независимых случайных векторов* (V_n) *называется стандартной субгауссовской последовательностью (с.с.п.), если при каждом* $n\ge 1$ V_n *— стандартный субгауссовский вектор и* $\sup\limits_{n\ge1}\tau_n<\infty$.

Класс с.с.п., кроме с.г.п., включает в себя еще, например, последовательности независимых случайных векторов (V_n), имеющих независимые компоненты, которые удовлетворяют условию 2 определения 1 и таких, что $\|V_n\|\le \mathbb{C}<\infty$ п.н., $n\ge 1$, где $\mathbb{C}$ — неслучайная константа.

Теорема 2. *Пусть в уравнении* (1) (V_n) *— с.с.п. и выполнено условие* (5). *Тогда найдется такая неслучайная константа* $\alpha\in[0,+\infty)$, *что выполняется соотношение* (6).

Замечания. **1.** Пусть в уравнении (1), как и в работе [1], (a_n) и (B_n) имеют вид (2), где $b\le a$ и в случае $a=1$ $\sigma>1/2$. Тогда условия (4), (5) выполнены.

2. Пусть в уравнении (1), как и в работе [2], $a_n=n^{-1}$, $B_n=n^{-(1+\beta)/2}I$, $n\ge 1$, где $\beta\ge 0$ и $\sigma>\beta/2$. Тогда условия (4), (5) выполнены.

3. Вспомогательные результаты. Обозначим через (c_n) произвольную фиксированную последовательность положительных чисел.

Лемма 1. *Пусть в уравнении* (1) (V_n) *— с.г.п., выполнены условия* (3) – (5) *и условие*

$$\varphi_n = c_n^2\|B_n\|/a_n\downarrow 0,\quad n\to\infty. \tag{7}$$

Тогда для того чтобы

$$\lim_{n\to\infty}c_nX_n = 0 \quad п.н., \tag{8}$$

необходимо и достаточно, чтобы

$$\lim_{n\to\infty}\varphi_n\log\left(\sum_{j=1}^{n}a_j\right) = 0. \tag{9}$$

Доказательство. В работе [4] показано, что при любом $\delta\in(0,\sigma)$ для достаточно больших номеров $n\ge i\ge n_1$ выполняются неравенства

$$\prod_{j=i}^{n}(1-\sigma a_j)^2 \le \left\| \prod_{j=i}^{n}(I-a_jA) \right\|^2 \le \mathbb{C}\prod_{j=i}^{n}(1-a_j)^{2(\sigma-\delta)},$$

где $0 < \mathbb{C} < \infty$; $a_i < \min\{1, 1/\sigma\}$, $i \ge n_1$.

Положим $\tau = 2(\sigma-\delta) + (\sigma+\varepsilon)\nu = (2-\nu)\sigma - 2\delta - \nu\varepsilon$. Выберем δ и $\varepsilon > 0$ так, чтобы $\tau \in (0, 1)$. Тогда для $i \ge n_2 \ge n_1$, учитывая условие $\lim\limits_{n\to\infty} a_n = 0$, можем записать

$$\prod_{j=i}^{n}(1-a_j)^{2(\sigma-\delta)} \le \prod_{j=i}^{n}(1-a_j)^{\tau} \exp\left(-\nu(\sigma+\varepsilon)\sum_{j=i}^{n}a_j\right) \le$$

$$\le \prod_{j=i}^{n}(1-a_j)^{\tau}\prod_{j=i}^{n}(1-\sigma a_j)^{\nu}.$$

Кроме того, при $i \ge n_3 \ge n_1$ или

$$\prod_{j=i}^{n}(1-\sigma a_j)^2 \ge \prod_{j=i}^{n}(1-a_j)^{\gamma}, \quad \gamma > \max\{2\sigma, 1\}.$$

Положим $n_4 = \max\{n_0, n_2, n_3\}$, где n_0 — номер из условия теоремы 1. Поскольку исследовать процедуру (1) мы можем начиная с номера $n \ge n_4$ при начальном условии X_{n_4-1}, то для удобства положим $n_4 = 1$. Таким образом, при $1 \le i \le n$ выполнены неравенства

$$\prod_{j=i}^{n}(1-a_j)^{\gamma} \le \left\| \prod_{j=i}^{n}(I-a_jA) \right\|^2 \le \mathbb{C}\prod_{j=i}^{n}(1-a_j)^{\tau}\prod_{j=i}^{n}(1-\sigma a_j)^{\nu}, \quad (10)$$

где $a_i < \min\{1, 1/\sigma\}$, $1 \le i \le n$. Кроме того, условия (4), (5) выполнены для $n_0 = 1$.

Согласно (1) получим

$$X_n = \prod_{j=1}^{n}(I-a_jA)\,X_0 + X_n^0 = x_n + X_n^0, \quad n \ge 1, \quad (11)$$

где (X_n^0) — решение уравнения (1) при $X_0 = 0$. Из неравенства (5) вытекает

$$(\|B_i\|^2/a_i)\prod_{j=i+1}^{n+1}(1-\sigma a_j)^{\nu} \le \|B_{n+1}\|^2/a_{n+1}, \quad 1 \le i \le n. \quad (12)$$

Из (7), (10), (12) следует, что в (11)

$$c_nx_n = c_n\prod_{j=1}^{n}(I-a_jA)\,X_0 \to 0 \quad (n\to\infty) \text{ п.н.}$$

Поэтому можем положить в (1) $X_0 = 0$. Тогда (c_nX_n) является центрированной гауссовской марковской последовательностью в R^d. Обозначим через

$$\Phi(m, n+1) = E\|c_{n+1}X_{n+1} - E(c_{n+1}X_{n+1}\,|\,c_mX_m)\|^2$$

условную дисперсию вектора $(c_{n+1}X_{n+1})$ относительно вектора (c_mX_m), $0 \le$ $\le m \le n$. В силу (1)

$$\Phi(m, n+1) = c_{n+1}^2 \left(\sum_{i=m+1}^{n} \left\| \prod_{j=i+1}^{n+1} (I - a_j A) B_i \right\|^2 + \|B_{n+1}\|^2 \right).$$

Используя правое неравенство (10) и неравенство (12), получаем

$$\Phi(m, n+1) \le \mathbb{C}\, \varphi_{n+1} \left(\sum_{i=m+1}^{n} \prod_{j=i+1}^{n+1} (1 - a_j)^{\tau} + a_{n+1} \right) \le$$

$$\le \mathbb{C}\, \tau^{-1} \varphi_{n+1} \left(\left(1 - a_{n+1}\right)^{\tau} - \prod_{j=m+1}^{n+1} (1 - a_j)^{\tau} + \tau a_{n+1} \right) \le$$

$$\le 2\,\mathbb{C}\, \tau^{-1} \varphi_{n+1} \left(1 - \prod_{j=m+1}^{n+1} (1 - a_j)^{1/2} \right). \tag{13}$$

Согласно условию (3) и неравенствам 14.49 и 20.17 из [5] найдется такая константа $L > 0$, что

$$\left\| \prod_{j=i+1}^{n+1} (I - a_j A) B_i \right\|^2 \ge L \, \| B_i \|^2 \left\| \prod_{j=i+1}^{n+1} (I - a_j A) \right\|^2.$$

Поэтому, используя левое неравенство (10) и неравенство (4), имеем

$$\Phi(m, n+1) \ge$$

$$\ge L \varphi_{n+1} \left[\sum_{i=m+1}^{n} \left(1 - (1 - a_i)^{\gamma}\right)^{-1} \left(\prod_{j=i+1}^{n+1} (1 - a_j)^{\gamma} - \prod_{j=i}^{n+1} (1 - a_j)^{\gamma} + a_{n+1} \right) \right] \ge$$

$$\ge L\gamma^{-1} \varphi_{n+1} \left(1 - \prod_{j=m+1}^{n+1} (1 - a_j)^{1/2} \right). \tag{14}$$

Кроме того, при $m = 0$

$$\Phi(0, n+1) \ge L\gamma^{-1} (1 - \sqrt{1 - a_1})\, \varphi_{n+1}. \tag{15}$$

В силу неравенств (13) – (15) для $0 \le m \le n$ можем записать

$$M_1 \le \Phi(m, n+1) / \left\{ \varphi_{n+1} \left[1 - \left(\prod_{j=1}^{m} (1 - a_j)^{-1/2} \Big/ \prod_{j=1}^{n+1} (1 - a_j)^{-1/2} \right) \right] \right\} \le M_2, \tag{16}$$

где

$$M_1 = L\gamma^{-1} (1 - \sqrt{1 - a_1}), \quad M_1 = 2\,\mathbb{C}\, \tau^{-1}, \quad \prod_{j=1}^{0} (1 - a_j)^{-1/2} = 0.$$

Так как $(c_n X_n)$ — центрированная гауссовская марковская последовательность в R^d, выполнено условие (7), справедливо неравенство (16) и $\prod_{j=1}^{n} (1 - - a_j)^{-1/2} \uparrow \infty$, $n \to \infty$, то в силу [6] (лемма 1) для выполнения соотношения (8) необходимо и достаточно, чтобы

$$\lim_{n \to \infty} \varphi_n \log \left(e + \sum_{j=1}^{n-1} \min \left\{ 1, \log (1 - a_{j+1})^{-1/2} \right\} \right) = 0.$$

Воспользовавшись теоремой 1 [7], получим, что данное соотношение эквивалентно соотношению (9). Лемма 1 доказана.

Замечание 3. Как следует из леммы 1 работы [6], для доказательства достаточной части леммы 1 нужна только правая часть неравенства (16), при доказательстве которой неравенства (3) и (4) не использовались.

Лемма 2. *Пусть в уравнении* (1) (V_n) *— с.г.п. и выполнены условия* (3) – (5), (7). *Тогда для того чтобы*

$$\sup_{n \geq 1} \| c_n X_n \| < \infty \quad \text{п.н.,} \tag{17}$$

необходимо и достаточно, чтобы

$$\sup_{n \geq 1} \left[\varphi_n \log \left(\sum_{j=1}^{n} a_j \right) \right] < \infty. \tag{18}$$

Данная лемма является следствием леммы 1 и леммы 3 из [8]. Доказательство ее аналогично доказательству теоремы 3.1 из [9].

Лемма 3. *Пусть в уравнении* (1) (V_n) *— с.с.п. и выполнены условия* (5), (7). *Тогда для выполнения соотношения* (17) *достаточно, чтобы было выполнено условие* (18).

Данная лемма является непосредственным следствием достаточной части леммы 2 и теоремы редукции [10] (теорема 3.1). При этом нужно воспользоваться представлением (11).

4. Доказательство теорем 1, 2. Положим в леммах 1 – 3

$$c_n = \left(\left(\| B_n \|^2 / a_n \right) \log \left(\sum_{j=1}^{n} a_j \right) \right)^{-1/2}, \quad n \geq 1.$$

Так как $\log \left(\sum_{j=1}^{n} a_j \right) \to \infty, \ n \to \infty$, то

$$\varphi_n = c_n^2 \| B_n \|^2 / a_n \downarrow 0, \quad n \to \infty, \tag{19}$$

и, значит, выполнено условие (7).

Рассмотрим вначале случай, когда (X_n) — решение уравнения (1) при $X_0 = 0$ и (V_n) — с.г.п. Используя неравенство (13), при $m = 0$ получаем

$$E \| c_n X_n \|^2 = \Phi(0, n) \leq 2 \mathbb{C} \tau^{-1} \varphi_n, \quad n \geq 1.$$

Отсюда в силу соотношения (19) следует

$$\lim_{n \to \infty} E \| c_n X_n \|^2 = 0. \tag{20}$$

Для последовательности (X_n) верно представление

$$X_n = \sum_{k=1}^{n} A_{nk} V_k, \quad n \geq 1, \tag{21}$$

где $(A_{nk};\ 1 \leq k \leq n,\ n \geq 1)$ — некоторый набор матриц. Тогда из (20) следует, что при любом $k \geq 1$

$$\lim_{n \to \infty} c_n \| A_{nk} \| = 0. \tag{22}$$

Пусть теперь (X_n) — решение уравнения (1), в котором (V_n) — с.с.п. Тогда для него верно представление (21) и выполнено условие (22). Поэтому в силу леммы 1 из [12, с. 108], которая остается верной и в R^d, найдется такая неслучайная константа $\alpha \in [0, +\infty)$, что

$$\limsup_{n\to\infty} c_n \|X_n\| = \alpha \quad \text{п.н.}$$

Так как

$$\varphi_n \log\left(\sum_{j=1}^{n} a_j\right) = 1, \quad n \geq 1, \tag{23}$$

то в силу достаточных условий лемм 2, 3 $\sup_{n\geq 1} \|c_n X_n\| < \infty$ п.н., и, следовательно, $\alpha < +\infty$.

С другой стороны, если в уравнении (1) (V_n) — с.г.п., то в силу (23) нарушено необходимое условие леммы 1 и, значит, $P\left(\limsup_{n\to\infty} c_n \|X_n\| > 0\right) > 0$. А так как α — неслучайная величина, отсюда следует, что $\alpha > 0$.

Пусть теперь (X_n) — решение уравнения (1) при $X_0 = 0$. Тогда справедливо представление (11) $X_n = x_n + X_n^0$, $n \geq 1$, где $\lim_{n\to\infty}(c_n x_n) = 0$ п.н. В силу уже доказанного

$$\limsup_{n\to\infty} c_n \left\|X_n^0\right\| = \alpha \quad \text{п.н.}$$

Используя неравенство

$$\left| \limsup_{n\to\infty} c_n \|X_n\| - \limsup_{n\to\infty} c_n \left\|X_n^0\right\| \right| \leq$$

$$\leq \limsup_{n\to\infty} c_n \left\| X_n - X_n^0 \right\| = \lim_{n\to\infty} c_n \|x_n\|,$$

заключаем, что $\limsup_{n\to\infty} c_n \|X_n\| = \alpha$ п.н.

Теоремы 1, 2 доказаны.

1. *Коростелев А.П.* Стохастические рекуррентные процедуры (локальные свойства). – М.: Наука, 1984. – 208 с.
2. *Walk H.* Limit behaviour of stochastic approximation processes // Statist. and Decis. – 1988. – **6**, №1, 2. – P. 109 – 128.
3. *Ljung L.* Convergence of recursive stochastic algorithms // Lund Inst. Technology: Rept 7403. Dep. Automat. Contr. 1974. – 134 p.
4. *Walk H., Zsido L.* Convrgence of the Robbins–Monro method for linear problems in a Banach space // J.Math. Anal. and Appl. – 1989. – **139**, №1. – P. 152 – 177.
5. *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. – М.: Наука, 1984. – 320 с.
6. *Коваль В.А.* Принципы сравнения для многомерных гауссовских марковских и гауссовских m-марковских последовательностей // Стохастические системы и их прил. – Киев: Ин-т математики АН УССР, 1990. – С.60 – 65.
7. *Коваль В.А.* Сходимость к нулю гауссовских последовательностей и асимптотическое поведение решений стохастических рекуррентных уравнений// Стохаст. анализ и его прил. – Киев: Ин-т математики АН УССР, 1989. – С.64 – 70.
8. *Егоров В.А.* О законе повторного логарифма// Теория вероятностей и ее применения. – 1969. – **14**, №4. – С.722 – 729.
9. *Булдыгин В.В., Коваль В.А.* Асимптотическое поведение решений стохастических разностных уравнений. – Киев, 1991. – 42 с. – (Препринт/АН УССР. Ин-т математики; №91.24).
10. *Коваль В.А.* Асимптотическое поведение решений стохастических рекуррентных уравнений: Автореф. дис. ... канд. физ.-мат. наук. – Киев, 1991. – 16 с.
11. *Булдыгин В.В., Харазишвили А.Б.* Неравенство Брунна–Минковского и его приложения. – Киев: Наук. думка, 1985. – 197 с.

Получено 16.11.92