A STOCHASTIC APPROXIMATION WITH A SEQUENCE OF DEPENDENT RANDOM VARIABLES

Watanabe, Masafumi
Department of Applied Mathematics, Fukuoka University

https://doi.org/10.5109/13146

出版情報：統計数理研究. 19 (3/4), pp.25-42, 1981-03. Research Association of Statistical Sciences
バージョン：
権利関係：

A STOCHASTIC APPROXIMATION WITH A SEQUENCE OF DEPENDENT RANDOM VARIABLES

By

Masafumi WATANABE

(Received September 20, 1980)

Summary

Let $\{Y_n\}$ be a sequence of dependent random variables and $\{\Phi_n(\cdot,\cdot)\}$ be a sequence of Borel functions. Let θ_n be a solution of the equation $M_n(x)=0$ for each $n\geqq 1$, where $M_n(x)=\mathrm{E}\Phi_n(x, Y_n)$. A Robbins-Monro type stochastic approximation procedure $X_{n+1}=X_n-a_n\Phi_n(X_n, Y_n)$ is considered for estimating θ_n for n sufficiently large. Under some assumptions about $\{a_n\}$, $\{\theta_n\}$, $\{Y_n\}$ and $\{\Phi_n(\cdot,\cdot)\}$ which may not include the fundamental condition $\mathrm{E}[\Phi_n(X_n, Y_n)\mid X_1, \cdots, X_n]=M_n(X_n)$ a.s., the a.s. convergence and in mean-square convergence of $|X_n-\theta_n|$ to zero are studied.

1. Introduction.

This paper is a continuation of our previous paper [7] and is concerened with the following Robbins-Monro type stochastic approximation method with a sequence of dependent random variables.

Let $\{Y_n\}$ be a sequence of $\boldsymbol{R}^M$-valued random vectors and $\{\Phi_n(\cdot,\ \cdot)\}$ be a sequence of $\boldsymbol{R}^N$-valued Borel functions defined on $\boldsymbol{R}^N\times\boldsymbol{R}^M$. Assume that $\mathrm{E}\Phi_n(x,\ Y_n)$ exists for all $x\in\boldsymbol{R}^N$ and $n\geqq 1$, and it is unknown to us. Assuming that the equation

$$\mathrm{E}\Phi_n(x,\ Y_n)=\mathbf{0} \qquad (\mathbf{0} \text{ denotes the zero vector of } \boldsymbol{R}^N) \tag{1.1}$$

has a solution $x=\theta_n$ for each $n\geqq 1$, it is desired to estimate θ_n for n sufficiently large on the basis observed values $\Phi_1(X_1,\ Y_1)$, $\Phi_2(X_2,\ Y_2)$, $\cdots$ at the points $(X_1,\ Y_1)$, $(X_2,\ Y_2)$, $\cdots$ where X_1, X_2, $\cdots$ are produced by the following recurrence relation;

$$\begin{cases} X_1=\text{an arbitrary constant vector of } \boldsymbol{R}^N \\ X_{n+1}=X_n-a_n\Phi_n(X_n,\ Y_n), \end{cases} \tag{1.2}$$

where $\{a_n\}$ is a decreasing sequence of positive numbers which converges to zero. The above situation will be concerned throughout this paper.

In the above problem, if the condition

Department of Applied Mathematics, Fukuoka University, Fukuoka 814-01, Japan

$$\mathrm{E}[\Phi_n(X_n, Y_n) | X_1, X_2, \cdots, X_n] = M_n(X_n) \quad \text{a.s.} \tag{1.3}$$

where $M_n(x) = \mathrm{E}\Phi_n(x, Y_n)$, is satisfied then this is considered as the usual Robbins-Monro stochastic approximation method. For example, if $\{Y_n\}$ is a sequence of independent random vectors then (1.3) is automatically satisfied under the procedure (1.2). Hence this is the case of the usual Robbins-Monro stochastic approximation. If $\{Y_n\}$ is a sequence of dependent random vectors then (1.3) may not be satisfied.

Thus, in this paper we shall study the above stochastic approximation under the situation that (1.3) does not hold. In [7], we considered the above problem under the assumption that $\Phi_n(x, y)$ can be expressed in the form of $(x-\theta_n)A_n(y)+\Gamma_n(y)$. And the a.s. convergence of the process (1.2) was investigated. In this paper we consider the same problem under some assumptions which are more general than ours in [7].

This paper consists of five sections. In Section 2 we give notation and four lemmas to be used throughout the paper. In Section 3 the conditions that have to impose on the procedure (1.2) are described. In Section 4 we shall give results about the a.s. convergence and in mean-square convergence of (1.2). In Section 5 we shall treat two examples of results. One of them was discussed in our previous paper [7].

2. Notations and lemmas.

The conventions introduced here hold throughout. Let $\boldsymbol{R}^k$ be k-dimensional Euclidian space. If $a, b \in \boldsymbol{R}^N$ and $c, d \in \boldsymbol{R}^{N_0}$, $\langle a, b\rangle$ and $\langle c, d\rangle_0$ denote their inner products, respectively. The Euclidian norms of $a \in \boldsymbol{R}^N$ and $c \in \boldsymbol{R}^{N_0}$ are denoted by $|a|$ and $|c|_0$, respectively, of course, $|a| = \langle a, a\rangle^{1/2}$ and $|c|_0 = \langle c, c\rangle_0^{1/2}$.

Let $(\Omega, \mathcal{A}, P)$ be a probability space. Define the dependent coefficient of σ-fields $\mathcal{A}'$ and $\mathcal{A}''$ which are sub-σ-fields of $\mathcal{A}$ by the relation;

$$\phi(\mathcal{A}', \mathcal{A}'') = \sup_{A \in \mathcal{A}''} \left(\operatorname*{ess\,sup}_{\omega \in \Omega} |P(A|\mathcal{A}')(\omega) - P(A)| \right). \tag{2.1}$$

For the above definition of the dependent coefficient we refer to Iosifescu and Theodorescu [1]. Let $\mathcal{A}_i$, $i \in T$ be sub-σ-fields of $\mathcal{A}$. Then $\bigvee_{i \in T} \mathcal{A}_i$ denotes the σ-fields generated by the join of the σ-fields $\mathcal{A}_i$, $i \in T$.

Next we shall give four lemmas to be used throughout the paper. The following Lemma 1 is given in [4].

LEMMA 1. *Let $\{\alpha_n\}$ and $\{v_n\}$ be sequences of non-negative numbers such that*

(i) $\sum v_n < \infty$, $\quad \sum \alpha_n < \infty$.

If $\{x_n\}$ is a sequence of non-negative numbers such that, for some integer n_0 and for all $n \geqq n_0$,

(ii) $x_{n+1} \leqq \max\{h, (1+\alpha_n)x_n + v_n\}$

where $h > 0$, then $\{x_n\}$ is bounded.

Using Lemma 2.3 of [6], we obatin the following lemma.

LEMMA 2. *Let $\{h_n\}$, $\{a_n\}$, $\{b_n\}$, $\{d_n\}$, $\{u_n\}$, $\{v_n\}$ and $\{w_n\}$ be sequences of real numbers such that*

(i) $h_n \geqq 0$, $\lim_{n\to\infty} h_n = 0$

(ii) $a_n \geqq a_{n+1} \geqq 0$, $\lim_{n\to\infty} a_n = 0$, $\sum a_n = \infty$

(iii) $u_n > 0$, $\sup_n u_n < \infty$, $\sup_n \{u_n | a_n^{-1} - a_{n+1}^{-1}|\} < \infty$

(iv) $\lim_{n\to\infty} b_n = 0$, $\sum |v_n| < \infty$

(v) $\lim_{n\to\infty} a_n u_n^{-1} \sum_{k=1}^{n} w_k = 0$

(vi) $d_n \geqq 0$, $\sum d_n < \infty$.

If $\{x_n\}$ *is a sequence of non-negative numbers such that, for some positive integer* n_0 *and for all* $n \geqq n_0$,

(vii) $x_{n+1} \leqq \max\{h_n, (1-a_n+d_n)x_n + a_n b_n + a_n w_n + v_n\}$,

then it holds that $\lim_{n\to\infty} x_n = 0$.

PROOF. It involves no loss of generality to assume

$$1 - a_n \geqq \frac{1}{2} \quad \text{for all} \quad n \geqq n_0 . \tag{2.2}$$

Let N and n be arbitrary positive integers such that $n_0 < N < n$. Set

$$y_n = a_n b_n + a_n w_n + v_n .$$

The iterate (vii) back to N. This yields

$$x_{n+1} \leqq \max\left\{\max_{N \leq k \leq n}\left\{T_{k+1}^{(n)} h_k + \sum_{i=k+1}^{n} T_{i+1}^{(n)} y_j\right\},\ T_N^{(n)} x_N + \sum_{i=N}^{n} T_{i+1}^{(n)} y_i\right\}$$

where

$$T_k^{(n)} = \begin{cases} \prod_{j=k}^{n} (1 - a_j + d_j) & \text{if} \quad 1 \leqq k \leqq n \\ 1 & \text{if} \quad k = n+1 . \end{cases}$$

Hence we obtain

$$x_{n+1} \leqq \max_{N \leq k \leq n} \{T_{k+1}^{(n)} h_k\} + T_N^{(n)} x_N + 2 \max_{N \leq k \leq n} \left| \sum_{i=k}^{n} T_{i+1}^{(n)} y_i \right| . \tag{2.3}$$

Now from (i) and (v), for any $\varepsilon > 0$ we can chose N so that

$$\sup_{N \leq n} h_n < \varepsilon , \tag{2.5}$$

$$\sup_{N-1 \leq n} \left| a_n u_n^{-1} \sum_{k=1}^{n} w_k \right| < \varepsilon . \tag{2.5}$$

Since $1-a_n+d_n=(1-a_n)(1-(1-a_n)^{-1}d_n)$, (2.2) implies

$$1-a_n+d_n\leqq(1-a_n)(1+2d_n)\leqq 1+2d_n\,.$$

Hence, (vi) and (2.4) imply that there exists a positive constant $K_1=\prod_{n=1}^{\infty}(1+2d_n)$ such that

$$\max_{N\leqq k\leqq n} T_{k+1}^{(n)}h_k<K_1\varepsilon\,. \tag{2.6}$$

And it is easily seen from (ii) and (vi) that

$$\lim_{n\to\infty} T_N^{(n)}x_N\leqq\lim_{n\to\infty}K_1x_N\prod_{k=N}^{n}(1-a_k)=0\,. \tag{2.7}$$

Note that

$$\left|\sum_{i=k}^{n}T_{i+1}^{(n)}y_i\right|\leqq\sum_{i=n_0}^{n}T_{i+1}^{(n)}(a_i|b_i|+v_i|)+\left|\sum_{i=k}^{n}a_iw_iT_{i+1}^{(n)}\right|\,. \tag{2.8}$$

First, we shall prove that

$$\lim_{n\to\infty}\sum_{i=n_0}^{n}T_{i+1}^{(n)}(a_i|b_i|+|v_i|)=0\,. \tag{2.9}$$

This is obtained with the aid of Lemma 2.3 of [6], which states that if $\{A_n\}$ is a sequence of non-negative numbers satisfying

$$A_{n+1}\leqq(1-a_{n+1})A_n+a_{n+1}b_{n+1}+v_{n+1}$$

where $a_n\geqq 0$, $\lim_{n\to\infty}a_n=0$, $\sum a_n=\infty$, $b_n\geqq 0$, $\lim_{n\to\infty}b_n=0$, $v_n\geqq 0$ and $\sum v_n<\infty$, then $\lim_{n\to\infty}A_n=0$. Since $(1-a_n+d_n)\leqq(1-a_n)(1+2d_n)$ and $\sum d_n<\infty$, it follows that

$$\sum_{i=n_0}^{n}T_{i+1}^{(n)}(a_i|b_i|+|v_i|)\leqq K_1\sum_{i=n_0}^{n}(a_i|b_i|+|v_i|)\prod_{k=i+1}^{n}(1-a_k)$$

where $K_1=\prod_{n=1}^{\infty}(1+2d_n)<\infty$. Define $A_n=\sum_{i=n_0}^{n}(a_i|b_i|+|v_i|)\times\prod_{k=i+1}^{n}(1-a_k)$. Then Lemma 2.3 of [6] gives (2.9).

Next, let us define $u_0=a_1=1$, $s_0=0$ and for $n\geqq k\geqq 1$,

$$s_n=a_nu_n^{-1}\sum_{k=1}^{n}w_k\,,\qquad s_k^{(n)}=\sum_{i=k}^{n}a_iw_iT_{i+1}^{(n)}\,.$$

Since $w_i=a_i^{-1}u_is_i-a_{i-1}^{-1}u_{i-1}s_{i-1}$, we obtain

$$s_k^{(n)}=u_ns_n-a_ka_{k-1}^{-1}u_{k-1}s_{k-1}T_{k+1}^{(n)}+\sum_{j=k}^{n-1}a_j^{-1}u_js_j(a_jT_{j+1}^{(n)}-a_{j+1}T_{j+2}^{(n)})$$

$$=u_ns_n-a_ka_{k-1}^{-1}u_{k-1}s_{k-1}T_{k+1}^{(n)}+\sum_{j=k}^{n-1}a_{j+1}s_ju_jT_{j+2}^{(n)}(a_{j+1}^{-1}-a_j^{-1}-1)\,.$$

Therefore it follows from (ii), (iii) and (vi) that there exist positive constants K_2 and K_3 such that

(2.10) $$|s_k^{(n)}| \leqq K_2 \sup_{N-1 \leqq n} |s_n| + K_3 \sum_{j=n_0}^{n} a_j |s_{j-1}| T_{j+1}^{(n)} .$$

Applying Lemma 2.3 of [6] to the second term in the righthand side of (2.10) we obtain

(2.11) $$\lim_{n\to\infty} \sum_{j=n_0}^{n} a_j |s_{j-1}| T_{j+1}^{(n)} = 0 .$$

Therefore (2.3) and (2.5) to (2.11) imply

$$\limsup_{n\to\infty} x_n \leqq (K_1 + 2K_2)\varepsilon .$$

Because ε is arbitrary, this concludes the proof of Lemma 2.

REMARK. Examples of two sequences $\{a_n\}$ and $\{u_n\}$ satisfying (ii) and (iii) are easy to obtain. For example, $a_n = n^{-\alpha}$ $(0<\alpha\leqq 1)$ and $u_n = n^{-\beta}$ $(0\leqq\beta)$ satisfy (ii) and (iii). And $a_n = (n \log n)^{-1}$ and $u_n = (\log n)^{-1}$ satisfy (ii) and (iii) also.

Finally we state without proof two lemmas which are proved in [1].

LEMMA 3. *(The strong law of large numbers). Let* $\{X_n\}$ *be a sequence of random variables. Suppose that there exist a positive integer* n_0 *and a decreasing sequence of positive numbers* $\{a_n\}$ *such that*

(i) $\limsup_{n\to\infty} \phi\Big(\bigvee_{1\leqq i\leqq n} \mathcal{A}_i, \bigvee_{n+n_0\leqq i} \mathcal{A}_i\Big) < 1$,

(ii) $\sum_{n=1}^{\infty} \{\sup_m \phi(\mathcal{A}_m, \mathcal{A}_{m+n})\}^{1/2} < \infty$,

(iii) $\lim_{n\to\infty} a_n = 0$,

(iv) $\sum_{n=1}^{\infty} a_n^2 \mathrm{E}(X_n - \mathrm{E}X_n)^2 < \infty$,

where $\mathcal{A}_i$ *is the* σ*-field generated by* X_i *and* $\phi(\cdot, \cdot)$ *is defined by* (2.1). *Then it holds that*

$$\lim_{n\to\infty} a_n \sum_{i=1}^{n} (X_i - \mathrm{E}X_i) = 0 \quad a.\,s. .$$

REMARK. Let $\{X_n\}$ be a sequence of $\boldsymbol{R}^N$-valued random vectors satisfying (i) and (ii). And (iii) and

(iv)′ $\sum a_n^2 \mathrm{E}|X_n - \mathrm{E}X_n|^2 < \infty$ ("$|\cdot|$" denotes the Euclidian norm) are satisfied. Then it also holds that

$$\lim_{n\to\infty} a_n \Big| \sum_{i=1}^{n} (X_i - \mathrm{E}X_i) \Big| = 0 \quad \text{a. s.} .$$

LEMMA 4. *Let the random variable* X_i *be* $\mathcal{A}_i$*-measurable for* $i=1, 2$. *Suppose that* $\mathrm{E}X_i^2 < \infty$, $i=1, 2$. *Then*

$$|\mathrm{E}X_1X_2 - \mathrm{E}X_1\mathrm{E}X_2| \leqq 2\phi^{1/2}(\mathcal{A}_1, \mathcal{A}_2)(\mathrm{E}X_1^2)^{1/2}(\mathrm{E}X_2^2)^{1/2} .$$

REMARK. Let X_i, $i=1, 2$ be $\boldsymbol{R}^N$-valued random vectors satisfying $\mathrm{E}|X_i|^2<\infty$, $i=1, 2$. And X_i is $\mathcal{A}_i$-measurable for $i=1, 2$. Then it is easily seen that

$$|\mathrm{E}\langle X_1, X_2\rangle-\langle \mathrm{E}X_1, \mathrm{E}X_2\rangle| \leqq 2\phi^{1/2}(\mathcal{A}_1, \mathcal{A}_2)(\mathrm{E}|X_1|^2)^{1/2}(\mathrm{E}|X_2|^2)^{1/2}.$$

3. Assumptions.

Assume that, for each $n\geqq 1$, the equation (1.1) has a solution θ_n. Let $\mathcal{A}_n$ be the σ-field generated by the $\boldsymbol{R}^M$-valued random vector Y_n for each $n\geqq 1$. Let $\{a_n\}$ be the decreasing sequence of positive numbers which is used in the procedure (1.2). And $\{a_n\}$ satisfies the following conditions A1 and A2.

A1: $\lim_{n\to\infty} a_n=0, \qquad \sum a_n=\infty.$

A2: There exists a non-increasing sequence of positive numbers $\{\delta_n\}$ such that

$$\lim_{n\to\infty} \delta_n=0, \qquad \sum a_n\delta_n<\infty.$$

Moreover we shall make the following assumptions about $\{a_n\}$ and $\{\delta_n\}$.

A3: $a_n\delta_n^{-3}>a_{n+1}\delta_{n+1}^{-3}, \qquad \lim_{n\to\infty} a_n\delta_n^{-3}=0.$

A4: $\sup_n |a_n^{-1}-a_{n+1}^{-1}|<\infty.$

A5: $\sum a_n^2\delta_n^{-2}<\infty.$

Next we shall make the following assumptions B1 to B7 which are used in the proof of Theorem 1.

B1: For each $n\geqq 1$, let $f_n(\cdot, \cdot)$ be a real valued Borel function defined on $\boldsymbol{R}^N\times\boldsymbol{R}^M$ and $F_n(\cdot)$ and $G_n(\cdot)$ be $\boldsymbol{R}^{N_0}$-valued Borel functions on $\boldsymbol{R}^N$ and $\boldsymbol{R}^M$, respectively. Suppose that there exist a sequence of non-negative numbers $\{\varepsilon_n\}$ which converges to zero and a positive constant α_0 such that, for all $y\in\boldsymbol{R}^M$ and for all $n\geqq 1$,

$$\langle x-\theta_n, \Phi_n(x, y)\rangle \geqq \max\{f_n(x, y), \alpha_0|x-\theta_n|^2+\langle F_n(x), G_n(y)\rangle_0\}$$

if $|x-\theta_n|^2>\varepsilon_n$.

B2: There exist sequences of non-negative Borel functions $\{g_n(\cdot)\}$ and $\{h_n(\cdot)\}$ defined on $\boldsymbol{R}^M$ such that

(i) $|f_n(x, y)|\leqq g_n(y)|x-\theta_n|^2+h_n(y)$

for all $x\in\boldsymbol{R}^N$, $y\in\boldsymbol{R}^M$ and $n\geqq 1$, where $f_n(\cdot, \cdot)$ is to be given in B1,

(ii) $\sum a_n \mathrm{E} g_n(Y_n)<\infty,$

(iii) $\sum a_n\delta_n^2 \mathrm{E} h_n(Y_n)<\infty.$

B3: Let $\{F_n(\cdot)\}$ be as given in B1. There exist non-negative constants K_i, $1\leqq i\leqq 4$ and three sequences of non-negative numbers $\{b_n\}$, $\{c_n\}$ and $\{d_n\}$ such that

(i) $|F_n(x)|_0 \leqq K_1|x-\theta_n|^2+K_2|x-\theta_n|+K_3$ for all $x\in \boldsymbol{R}^N$ and $n\geqq 1$,

(ii) $|F_n(x)-F_n(x')|_0 \leqq K_4(|x-\theta_n|+|x'-\theta_n|+1)|x-x'|$

for all $x, x'\in \boldsymbol{R}^N$ and $n\geqq 1$.,

(iii) $|F_n(x)-F_{n+1}(x)|_0 \leqq b_n|x-\theta_n|^2+c_n|x-\theta_n|+d_n$ for all $x\in \boldsymbol{R}^N$ and $n\geqq 1$,

(iv) $\sum b_n\delta_n<\infty$,

(v) $\sup_n c_n\delta_n<\infty$, $\sum c_n\delta_n^2<\infty$,

(vi) $\sum d_n\delta_n^3<\infty$.

B4: Let $\{G_n(\cdot)\}$ be as given in B1.

(i) $\lim_{n\to\infty}|\mathrm{E}G_n(Y_n)|_0=0$,

(ii) $\sum a_n^2\delta_n^{-6}\mathrm{E}|G_n(Y_n)-\mathrm{E}G_n(Y_n)|_0^2<\infty$.

B5: There exist sequences of non-negative Borel functions $\{\alpha_n(\cdot)\}$ and $\{\beta_n(\cdot)\}$ defined on $\boldsymbol{R}^M$ such that

(i) $|\Phi_n(x, y)|^2 \leqq \alpha_n(y)|x-\theta_n|^2+\beta_n(y)$ for all $x\in \boldsymbol{R}^N$, $y\in \boldsymbol{R}^M$ and $n\geqq 1$,

(ii) $\sum a_n\delta_n\mathrm{E}\alpha_n(Y_n)<\infty$,

(iii) $\sum a_n\delta_n^3\mathrm{E}\beta_n(Y_n)<\infty$.

B6: Let n_0 be a positive integer.

(i) $\limsup_{n\to\infty}\phi\Big(\bigvee_{1\leq i\leq n}\mathcal{A}_i, \bigvee_{n+n_0\leq i}\mathcal{A}_i\Big)<1$,

(ii) $\sum_{n=1}^{\infty}\sup_m \phi^{1/2}(\mathcal{A}_m, \mathcal{A}_{m+n})<\infty$.

B7: $\lim_{n\to\infty}a_n^{-1}|\theta_n-\theta_{n+1}|=0$.

Finally we shall make the following assumptions C1 to C3 which are used in the proof of Theorem 2.

C1: There exist a positive constant α_0 and a sequence of $\boldsymbol{R}^N$-valued Borel functions $\{\gamma_n(\cdot)\}$ defined on $\boldsymbol{R}^M$ such that

(i) $\langle x-\theta_n, \Phi_n(x, y)\rangle \geqq \alpha_0|x-\theta_n|^2+\langle x-\theta_n, \gamma_n(y)\rangle$

for all $x\in \boldsymbol{R}^N$, $y\in \boldsymbol{R}^N$ and $n\geqq 1$,

(ii) $\lim_{n\to\infty}|\mathrm{E}\gamma_n(Y_n)|=0$,

(iii) $\sup_n \mathrm{E}|\gamma_n(Y_n)-\mathrm{E}\gamma_n(Y_n)|^2<\infty$.

C2: There exist a positive constant α_1 and a sequence of non-negative Borel functions $\{\beta_n(\cdot)\}$ defined on $\boldsymbol{R}^M$ such that

(i) $|\Phi_n(x, y)|^2 \leqq \alpha_1 |x-\theta_n|^2 + \beta_n(y)$ for all $x \in \boldsymbol{R}^N$, $y \in \boldsymbol{R}^M$ and $n \geqq 1$,

(ii) $\sup_n \mathrm{E}\, \beta_n(Y_n) < \infty$.

C3: $\sum_{n=1}^{\infty} \sup_m \phi^{1/2}\Big(\bigvee_{1\leq i\leq m} \mathcal{A}_i, \bigvee_{m+n\leq i} \mathcal{A}_i\Big) < \infty$.

REMARKS. (1) Since $\{\delta_n\}$ is non-increasing and bounded, A3 implies

A3′: $a_n\delta_n^{-1} > a_{n+1}\delta_{n+1}^{-1}$, $\lim_{n\to\infty} a_n\delta_n^{-1} = 0$.

Because $a_n^2 = a_n\delta_n a_n\delta_n^{-1}$, A3′ and B5(ii) imply

(3.1) $$\sum a_n^2 \alpha_n(Y_n) < \infty \quad \text{a.s.}.$$

And because $a_n^2 = a_n\delta_n^3 a_n\delta_n^{-3}$, A3 and B5(iii) imply

(3.2) $$\sum a_n^2 \beta_n(Y_n) < \infty \quad \text{a.s.}.$$

(2) Using the inequality $2ab \leqq ka^2 + k^{-1}b^2$ which holds for and $k>0$, it follows from B3(i) that

$$\begin{aligned}
&\alpha_0|x-\theta_n|^2 + \langle F_n(x), G_n(y)\rangle_0 \\
&\leqq \frac{\delta_n}{2}|x-\theta_n|^2 + \frac{\delta_n^{-1}\alpha_0^2}{2} + K_1|G_n(y)|_0|x-\theta_n|^2 \\
&\quad + \frac{\delta_n}{2}|x-\theta_n|^2 + \frac{\delta_n^{-1}K_2^2}{2}|G_n(y)|_0^2 + K_3|G_n(y)|_0 \\
&= (\delta_n + K_1|G_n(y)|_0)|x-\theta_n|^2 + \frac{\delta_n^{-1}\alpha_0^2}{2} + \frac{\delta_n^{-1}K_2^2}{2}|G_n(y)|_0^2 + K_3|G_n(y)|_0 .
\end{aligned}$$

If A2,

(3.3) $$\sum a_n \mathrm{E}|G_n(Y_n)|_0 < \infty$$

and

(3.4) $$\sum a_n\delta_n \mathrm{E}|G_n(Y_n)|_0^2 < \infty$$

are satisfied then we can chose $f_n(\cdot, \cdot)$ so that

$$f_n(x, y) = \alpha_0|x-\theta_n|^2 + \langle F_n(x), G_n(y)\rangle_0 .$$

Putting $g_n(y) = \delta_n + K_1|G_n(y)|_0$ and $h_n(y) = \frac{\delta_n^{-1}\alpha_0^2}{2} + \frac{\delta_n^{-1}K_2^2}{2}|G_n(y)|_0^2 + K_3|G_n(y)|_0$, A2, (3.3) and (3.4) imply B2. If $a_n\delta_n \leqq a_n^2\delta_n^{-6}$ follows then B4(ii) implies (3.4), implying B2(iii). But B4 does not yield (3.3). Hence B2(ii) may not be satisfied. Thus it is needed to introduce the function $f_n(\cdot, \cdot)$ when $K_1>0$. If we can assume that $K_1=0$ then $f_n(\cdot, \cdot)$ is not needed.

(3) Since $\phi(\mathcal{A}_m, \mathcal{A}_{m+n}) \leqq \phi\Big(\bigvee_{1\leq i\leq m} \mathcal{A}_i, \bigvee_{m+n\leq i} \mathcal{A}_i\Big)$, the mixing condition C3 implies B6.

(4) Examples of two sequences $\{a_n\}$ and $\{\delta_n\}$ satisfying A1 to A4 are easy to

obtain. For example, $a_n=n^{-\alpha}$ $\left(\frac{7}{8}<\alpha\leqq 1\right)$ and $\delta_n=n^{-\beta}$ $\left(1-\alpha<\beta\leqq\frac{\alpha}{7}\right)$ satisfy A1 to A4. And A5 is also satisfied. $a_n=n^{-\alpha'}$ $\left(\frac{3}{4}<\alpha'\leqq 1\right)$ and $\delta_n=n^{-\beta'}$ $\left(1-\alpha'<\beta'\leqq\frac{\alpha'}{3}\right)$ satisfy A1, A2, A3′, A4 and A5, where A3′ is to be given in (1).

4. Results.

Using Lemma 1 we obtain the following lemma.

LEMMA 5. *Let $X_1, X_2, \cdots$ be the Robbins-Monro type process defines by* (1.2). *Suppose that* A1 *to* A3, B1, B2, B5 *and* B7 *are satisfied. Then it holds that there exists a non-negative random variable Z such that, for all $n\geqq 1$,*

$$|X_n-\theta_n|\leqq\delta_n^{-1}Z \quad a.s.. \tag{4.1}$$

PROOF. We reduce the lemma to an a.s. pointwise application of Lemma 1. By virtue of (1.2) we obtain directly

$$\begin{aligned}|X_{n+1}-\theta_{n+1}|^2=&|X_n-\theta_n|^2+a_n^2|\Phi_n(X_n,\ Y_n)|^2+|\theta_n-\theta_{n+1}|^2\\ &+2\langle X_n-\theta_n,\ \theta_n-\theta_{n+1}\rangle-2a_n\langle\Phi_n(X_n,\ Y_n),\ \theta_n-\theta_{n+1}\rangle\\ &-2a_n\langle X_n-\theta_n,\ \Phi_n(X_n,\ Y_n)\rangle .\end{aligned}\tag{4.2}$$

Using the inequality $2ab\leqq ka^2+k^{-1}b^2$ $(k>0)$ we obtain

$$2|\langle X_n-\theta_n,\ \theta_n-\theta_{n+1}\rangle|\leqq a_n\delta_n|X_n-\theta_n|^2+a_n\delta_n^{-1}A_n \tag{4.3}$$

where

$$A_n=(a_n^{-1}|\theta_n-\theta_{n+1}|)^2 . \tag{4.4}$$

From B5(i) we botain

$$\begin{aligned}2a_n|\langle\Phi_n(X_n,\ Y_n),\ \theta_n-\theta_{n+1}\rangle|\leqq& a_n^2\alpha_n(Y_n)|X_n-\theta_n|^2\\ &+a_n^2\beta_n(Y_n)+a_n^2A_n ,\end{aligned}\tag{4.5}$$

where A_n is defined by (4.4). And from B1 and B2(i), if $|X_n-\theta_n|^2>\varepsilon_n$, then it follows that

$$\begin{aligned}-2a_n\langle X_n-\theta_n,\ \Phi_n(X_n,\ Y_n)\rangle\leqq& 2a_n|f_n(X_n,\ Y_n)|\\ \leqq& 2a_ng_n(Y_n)|X_n-\theta_n|^2+2a_nh_n(Y_n)^2 .\end{aligned}\tag{4.6}$$

Substituting (4.3), (4.5), (4.6) and B5(i) into (4.2), if $|X_n-\theta_n|^2>\varepsilon_n$ then it follows that

$$\begin{aligned}|X_{n+1}-\theta_{n+1}|^2\leqq&(1+a_n\delta_n+2a_n^2\alpha_n(Y_n)+2a_ng_n(Y_n))|X_n-\theta_n|^2\\ &+2a_n^2\beta_n(Y_n)+2a_nh_n(Y_n)+(2a_n^2+a_n\delta_n^{-1})A_n .\end{aligned}\tag{4.7}$$

Next we suppose that $|X_n-\theta_n|^2\leqq\varepsilon_n$. Substituting B5(i) into (4.2) we obtain

$$\begin{aligned}|X_{n+1}+\theta_{n+1}|^2\leqq& 3|X_n-\theta_n|^2+3a_n^2|\Phi_n(X_n,\ Y_n)|^2+3|\theta_n-\theta_{n+1}|^2\\ \leqq& 3(1+a_n^2\alpha_n(Y_n))\varepsilon_n+3a_n^2\beta_n(Y_n)+3a_n^2A_n\\ &\qquad\qquad\text{if}\quad |X_n-\theta_n|^2\leqq\varepsilon_n .\end{aligned}\tag{4.8}$$

Since $\{\delta_n\}$ is non-increasing, (4.7) and (4.8) imply

$$\delta_{n+1}^2 |X_{n+1}-\theta_{n+1}|^2 \leqq \max \{\delta_n^2 h_n,\ (1+d_n')\delta_n^2 |X_n-\theta_n|^2+\delta_n^2 v_n'\}$$

where

$$h_n=3(1+a_n^2\alpha_n(Y_n))\varepsilon_n+3a_n^2\beta_n(Y_n)+3a_n^2(a_n^{-1}|\theta_n-\theta_{n+1}|)^2, \tag{4.9}$$

$$d_n'=a_n\delta_n+2a_n^2\alpha_n(Y_n)+2a_n g_n(Y_n),$$

$$v_n'=2a_n^2\beta_n(Y_n)+2a_n h_n(Y_n)+(2a_n^2+a_n\delta_n^{-1})(a_n^{-1}|\theta_n-\theta_{n+1}|)^2.$$

From (3.1), (3.2) and B7, if follows that

$$\lim_{n\to\infty} h_n=0 \quad \text{a. s.}. \tag{4.10}$$

From A2, (3.1) and B2(ii), it follows that

$$\sum d_n' < \infty \quad \text{a. s.}.$$

And from (3.2), B2(iii), A2 and B7, it follows that

$$\sum \delta_n^2 v_n' < \quad \text{a. s.}.$$

Therefore, using Lemma 1, (4.1) follows.

Using Lemma 3 and Lemma 5 we obtain the following lemma.

LEMMA 6. *Suppose that the hypotheses of Lemma* 5 *are satisfied. Moreover, suppose that* B3, B4(ii) *and* B6 *are satisfied. Then it holds that*

$$\lim_{n\to\infty} a_n \sum_{k=1}^{n} \langle F_k(X_k),\ G_k(Y_k)-\mathrm{E}G_k(Y_k)\rangle_0=0 \quad a.\ s.. \tag{4.11}$$

PROOF. Define $\delta_0=a_0=1$, $S_0=\mathbf{0}$ (the zero vector of $\boldsymbol{R}^{N_0}$) and

$$S_n=a_n\delta_n^{-3}\sum_{k=1}^{n}(G_k(Y_k)-\mathrm{E}G_k(Y_k)).$$

Using Lemma 3 (The strong law of large numbers), it follows from A3 and B4(ii) that

$$\lim_{n\to\infty} |S_n|_0=0 \quad \text{a. s.}. \tag{4.12}$$

Let us put

$$a_n \sum_{k=1}^{n} \langle F_k(X_k),\ G_k(Y_k)-\mathrm{E}G_k(Y_k)\rangle_0=W_n.$$

Since $G_k(Y_k)-\mathrm{E}G_k(Y_k)=a_k^{-1}\delta_k^3 S_k-a_{k-1}^{-1}\delta_{k-1}^3 S_{k-1}$, W_n is rewritten

$$W_n=\delta_n^3\langle F_n(X_n),\ S_n\rangle_0+a_n\sum_{j=1}^{n-1} a_j^{-1}\delta_j^3\langle F_j(X_j)-F_{j+1}(X_{j+1}),\ S_j\rangle_0.$$

Then we obtain

$$|W_n| \leqq \delta_n^3 |F_n(X_n)|_0 |S_n|_0+a_n\sum_{j=1}^{n} a_j^{-1}\delta_j^3 |F_j(X_j)-F_j(X_{j+1})|_0 |S_j|_0$$

$$+a_n\sum_{j=1}^{n} a_j^{-1}\delta_j^3 |F_j(X_{j+1})-F_{j+1}(X_{j+1})|_0 |S_j|_0.$$

Hence, in order to show (4.11) we have to prove the following (a) to (c).

(a): $\lim_{n\to\infty} \delta_n^3 |F_n(X_n)|_0 |S_n|_0 = 0$ a. s..

(b): $\lim_{n\to\infty} a_n \sum_{j=1}^{n} a_j^{-1}\delta_j^3 |F_j(X_j)-F_j(X_{j+1})|_0 |S_j|_0 = 0$ a. s..

(c): $\lim_{n\to\infty} a_n \sum_{j=1}^{n} a_j^{-1}\delta_j^3 |F_j(X_{j+1})-F_{j+1}(X_{j+1})|_0 |S_j|_0 = 0$ a. s..

PROOF OF (a). It is easily seen from B3(i), Lemma 5 and (4.12) that (a) holds.

PROOF OF (b). By virture of (1.2) and B5(i), we obtain

$$(4.13)\qquad |X_j - X_{j+1}|^2 \leqq a_j^2 \alpha_j(Y_j)|X_j - \theta_j|^2 + a_j^2 \beta_j(Y_j).$$

From B3(ii), (4.13) and Lemma 5 we obtain

$$\begin{aligned} |F_j(X_j)-F_j(X_{j+1})|_0 &\leqq K_4(2|X_j-\theta_j| + |X_j - X_{j+1}| + 1)|X_j - X_{j+1}| \\ &\leqq 2K_4(a_j\delta_j^{-2}\alpha_j^{1/2}(Y_j)Z^2 + a_j\delta_j^{-1}\beta_j^{1/2}(Y_j)Z) \\ &\quad + K_4(a_j^2\delta_j^{-2}\alpha_j(Y_j)Z^2 + a_j^2\beta_j(Y_j)) \\ &\quad + K_4(a_j\delta_j^{-1}\alpha_j^{1/2}(Y_j)Z + a_j\beta_j^{1/2}(Y_j)) \quad \text{a. s.}. \end{aligned}$$

Define $Z_0 = \max\{Z^2, Z, 1\}$. Noting $\lim_{n\to\infty} a_n = 0$ and from A2 and A3, there exists a positive constant K_5 such that

$$\begin{aligned} |F_j(X_j)-F_j(X_{j+1})|_0 \leqq K_5 Z_0 \{&a_j\delta_j^{-2}(\alpha_j^{1/2}(Y_j)+\alpha_j(Y_j)) \\ &+ a_j\delta_j^{-1}(\beta_j^{1/2}(Y_j)+\beta_j(Y_j))\} \quad \text{a. s.}. \end{aligned}$$

Using Schwarz inequality, it follows from A2 and B5(ii) that

$$(4.14)\qquad \sum a_n\delta_n\alpha_n^{1/2}(Y_n) \leqq (\sum a_n\delta_n)^{1/2}(\sum a_n\delta_n\alpha_n(Y_n))^{1/2} < \infty \quad \text{a. s.}.$$

And it also follows from B5(iii) that

$$(4.15)\qquad \sum a_n\delta_n^2\beta_n^{1/2}(Y_n) \leqq (\sum a_n\delta_n)^{1/2}(\sum a_n\delta_n^3\beta_n(Y_n))^{1/2} < \infty \quad \text{a. s.}.$$

Hence it follows from (4.12), (4.14), (4.15), B5(ii) and B5(iii) that

$$\sum \delta_n^3 |F_n(X_n)-F_n(X_{n+1})|_0 |S_n|_0 < \infty \quad \text{a. s.}.$$

Therefore, from Kronecker lemma (see, e. g., [2], page 238), (b) follows.

PROOF OF (c). From B3(iii), (4.13) and Lemma 5 it follows that

$$\begin{aligned} |F_j(X_{j+1})-F_{j+1}(X_{j+1})|_0 &\leqq 2b_j|X_{j+1}-X_j|^2 + 2b_j|X_j-\theta_j|^2 \\ &\quad + c_j|X_j - X_{j+1}| + c_j|X_j-\theta_j| + d_j \\ &\leqq 2b_j\delta_j^{-2}(a_j^2\alpha_j(Y_j)+1)Z^2 + 2b_j a_j^2\beta_j(Y_j) \\ &\quad + c_j\delta_j^{-1}(a_j\alpha_j^{1/2}(Y_j)+1)Z \\ &\quad + c_j a_j\beta_j^{1/2}(Y_j) + d_j \quad \text{a. s.}. \end{aligned}$$

Since $b_n a_n = b_n \delta_n a_n \delta_n^{-1}$, it follows from B3(iv) that $\sup_n b_n a_n < \infty$.

Hence it follows from B5(ii), (iii) and B3(iv) that

$$\sum \{2b_n\delta_n(a_n^2\alpha_n(Y_n)+1)Z^2+2b_n a_n^2\delta_n^3\beta_n(Y_n)\}<\infty \quad \text{a. s.}. \tag{4.16}$$

And it also follows from B3(v), (4.14) and (4.15) that

$$\sum \{c_n\delta_n^2(a_n\alpha_n^{1/2}(Y_n)+1)Z+c_n a_n\delta_n^3\beta_n^{1/2}(Y_n)\}<\infty \quad \text{a. s.}. \tag{4.17}$$

Hence it follows from (4.12), (4.16), (4.17) and B3(vi) that

$$\sum \delta_n^3 |F_n(X_{n+1})-F_{n+1}(X_{n+1})|_0 |S_n|_0 < \infty \quad \text{a. s.}.$$

Then, using Kronecker lemma we obtain (c). Thus proof of lemma is complete.

Using Lemmas 2, 5 and 6 we obtain the following theorem which is the main result of this paper.

THEOREM 1. *Let* $X_1, X_2, \cdots$ *be the Robbins-Monro type process defined by* (1.2). *Suppose that* A1 *to* A4 *and* B1 *to* B7 *are satisfied. Then it holds that*

$$\lim_{n\to\infty} |X_n-\theta_n|=0 \quad a. s.. \tag{4.18}$$

PROOF. We reduce the theorem to an a. s. pointwise application of Lemma 2 (with $u_n=1$). Using the inequality $2ab \leqq ka^2+k^{-1}b^2$ which holds for any $k>0$, we obtain

$$2|\langle X_n-\theta_n,\ \theta_n-\theta_{n+1}\rangle| \leqq \frac{\alpha_0}{2} a_n |X_n-\theta_n|^2+2\alpha_0^{-1}a_n A_n , \tag{4.19}$$

where A_n is defined by (4.4). Let us write

$$|\mathrm{E}G_n(Y_n)|_0=B_n .$$

Then it follows from B3(i) that

$$\begin{aligned} 2|\langle F_n(X_n),\ \mathrm{E}G_n(Y_n)\rangle_0| &\leqq 2K_1B_n|X_n-\theta_n|^2+2K_2B_n|X_n-\theta_n|+2K_3B_n \\ &\leqq \left(2K_1B_n+\frac{\alpha_0}{2}\right)|X_n-\theta_n|^2+(2\alpha_0^{-1}K_2^2B_n+2K_3)B_n \end{aligned}$$

Hence, if $|X_n-\theta_n|^2>\varepsilon_n$ then it follows from B1 that

$$\begin{aligned} -2a_n\langle X_n-\theta_n,\ \Phi_n(X_n,\ Y_n)\rangle &\leqq -2\alpha_0 a_n|X_n-\theta_n|^2-2a_n\langle F_n(X_n),\ G_n(Y_n) \\ &\qquad -\mathrm{E}G_n(Y_n)\rangle_0+2a_n|\langle F_n(X_n),\ \mathrm{E}G_n(Y_n)\rangle_0| \\ &\leqq \left(-\frac{3}{2}\alpha_0 a_n+2K_1B_n a_n\right)|X_n-\theta_n|^2 \\ &\qquad +2(\alpha_0^{-1}K_2^2B_n+K_3)B_n a_n+w_n a_n , \end{aligned} \tag{4.20}$$

where

$$w_n=-2\langle F_n(X_n),\ G_n(Y_n)-\mathrm{E}G_n(Y_n)\rangle_0 .$$

Substituting B5(i), (4.5), (4.19) and (4.20) into (4.2), we obtain

$$|X_{n+1}-\theta_{n+1}|^2 \leqq (1-\alpha_0 a_n+2K_1B_n a_n+2a_n^2\alpha_n(Y_n))|X_n-\theta_n|^2 \tag{4.21}$$

$$+2a_n^2\beta_n(Y_n)+a_nw_n+2a_n(\alpha_0^{-1}K_2^2B_n+K_3)B_n$$
$$+2a_n(a_n+\alpha_0^{-1})A_n \qquad \text{if} \quad |X_n-\theta_n|^2>\varepsilon_n .$$

Hence (4.21) together with (4.8) we obtain

$$|X_{n+1}-\theta_{n+1}|\leqq \max\{h_n, (1-\alpha_0a_n+2K_1B_na_n+2a_n^2\alpha_n(Y_n))|X_n-\theta_n|^2$$
$$+2a_n^2\beta_n(Y_n)+a_nw_n+2a_n(\alpha_0^{-1}K_2^2B_n+K_3)B_n$$
$$+2a_n(a_n+\alpha_0^{-1})A_n\}$$

where h_n is to be defined by (4.9). Since $\lim_{n\to\infty}B_n=\lim_{n\to\infty}|EG_n(Y_n)|_0=0$, there exists a positive integer n_1 such that, for all $n\geqq n_1$,

$$1-\alpha_0a_n+2K_1B_na_n\leqq 1-\frac{1}{2}\alpha_0a_n .$$

Therefore it follows that, for all $n\geqq n_1$,

$$|X_{n+1}-\theta_{n+1}|^2\leqq \max\left\{h_n, \left(1-\frac{1}{2}\alpha_0a_n+d_n''\right)|X_n-\theta_n|^2+a_nb_n''+a_nw_n+v_n''\right. \tag{4.22}$$

where

$$d_n''=2a_n^2\alpha_n(Y_n), \qquad v_n''=2a_n^2\beta_n(Y_n)$$

and

$$b_n''=2(\alpha_0^{-1}K_2^2B_n+K_3)B_n+2(\alpha_0^{-1}+a_n)A_n .$$

Hence applying Lemma 2 (with $u_n=1$), it follows from (4.10), (3.1), (3.2), B7 and Lemma 6 that (4.18) holds. Thus the proof of the theorem is complete.

THEOREM 2. *Let* $X_1, X_2, \cdots$ *be the Robbins-Monro type process defined by* (1.2). *Suppose that* A1, A2, A3′, B7 *and* C1 *to* C3 *are satisfied. Then it holds that*

$$\lim_{n\to\infty} E|X_n-\theta_n|^2=0 . \tag{4.23}$$

Moreover, suppose that A4 *and* A5 *are satisfied. Then it also hold that*

$$\lim_{n\to\infty}|X_n-\theta_n|=0 \quad a.s.. \tag{4.24}$$

PROOF. Using Lemma 2 (with $h_n=0$ and $u_n=1$) and Lemma 3, we can prove (4.24) by the similar arguments of the proof of Theorem 1. Hence the proof of (4.24) is omitted. We shall prove (4.23). Using the inequality $2ab\leqq ka^2+k^{-1}b^2$ $(k>0)$, we obtain

$$2|\langle X_n-\theta_n, \theta_n-\theta_{n+1}\rangle|\leqq\frac{1}{2}\alpha_0a_n|X_n-\theta_n|^2+2(\alpha_0a_n)^{-1}|\theta_n-\theta_{n+1}|^2, \tag{4.25}$$

$$2|\langle\Phi_n(X_n, Y_n), \theta_n-\theta_{n+1}\rangle|\leqq a_n|\Phi_n(X_n, Y_n)|^2+a_n^{-1}|\theta_n-\theta_{n+1}|^2 \tag{4.26}$$

and

$$2|\langle X_n-\theta_n, E\gamma_n(Y_n)\rangle|\leqq\frac{\alpha_0}{2}|X_n-\theta_n|^2+2\alpha_0^{-1}|E\gamma_n(Y_n)|^2 . \tag{4.27}$$

Substituting (4.25) to (4.27), C1(i) and C2(i) into (4.2) we obtain

$$(4.28)\qquad |X_{n+1}-\theta_{n+1}|^2\leqq(1-\alpha_0 a_n+2\alpha_1 a_n^2)|X_n-\theta_n|^2 +2\{1+(\alpha_0 a_n)^{-1}\}|\theta_n-\theta_{n+1}|^2+2a_n^2\beta_n(Y_n) +2\alpha_0^{-1}a_n|\mathrm{E}\gamma_n(Y_n)|^2-2a_n\langle X_n-\theta_n,\ \gamma_n(Y_n) -\mathrm{E}\gamma_n(Y_n)\rangle\,.$$

It is easily seen from A1 that there exists a positive integer n_0 such that, for all $n\geqq n_0$,

$$(4.29)\qquad 1-\alpha_0 a_n+2\alpha_1 a_n^2\leqq 1-\frac{1}{2}\alpha_0 a_n\,.$$

Taking the expectation on the both side of (4.28) and from (4.29) we obtain

$$(4.30)\qquad \mathrm{E}|X_{n+1}-\theta_{n+1}|^2\leqq\Big(1-\frac{1}{2}\alpha_0 a_n\Big)\mathrm{E}|X_n-\theta_n|^2+a_n b_n+a_n w_n\,,\qquad n\geqq n_0\,,$$

where

$$b_n=2(a_n+\alpha_0^{-1})(a_n^{-1}|\theta_n-\theta_{n+1}|)^2+2\alpha_0^{-1}|\mathrm{E}\gamma_n(Y_n)|^2+2a_n\mathrm{E}\beta_n(Y_n)$$

and

$$w_n=2|\mathrm{E}\langle X_n-\theta_n,\ \gamma_n(Y_n)-\mathrm{E}\gamma_n(Y_n)\rangle|\,.$$

It is easily seen from A1, B7, C1(ii) and C2(ii) that

$$(4.31)\qquad \lim_{n\to\infty} b_n=0\,.$$

Hence, from (4.31) and according to Lemma 2.3 of [6], in order to show (4.23), we have only to prove that

$$(4.32)\qquad \lim_{n\to\infty} w_n=0\,.$$

First we shall prove that there exists a positive constant K such that, for all $n\geqq 1$,

$$(4.33)\qquad \mathrm{E}|X_n-\theta_n|^2\leqq\delta_n^{-1}K\,.$$

Substituting

$$w_n\leqq\frac{1}{4}\alpha_0\mathrm{E}|X_n-\theta_n|^2+4\alpha_0^{-1}\mathrm{E}|\gamma_n(Y_n)-\mathrm{E}\gamma_n(Y_n)|^2$$

into (4.30), it follows that, for all $n\geqq n_0$,

$$\delta_{n+1}\mathrm{E}|X_{n+1}-\theta_{n+1}|^2\leqq\Big(1-\frac{1}{4}\alpha_0 a_n\Big)\delta_n\mathrm{E}|X_n-\theta_n|^2+\delta_n a_n b_n +4\alpha_0^{-1}a_n\delta_n\mathrm{E}|\gamma_n(Y_n)-\mathrm{E}\gamma_n(Y_n)|^2\,.$$

Hence, using Lemma 2.3 of [6], it follows from (4.31), A2 and C1(iii) that

$$\lim_{n\to\infty}\delta_n\mathrm{E}|X_n-\theta_n|^2=0\,.$$

Thus (4.33) follows. Next we shall prove (4.32). Let us write

$$\gamma_n'(Y_n)=\gamma_n(Y_n)-\mathrm{E}\gamma_n(Y_n)\,.$$

Then it follows from (1.2) that

$$\langle X_n-\theta_n,\ \gamma_n'(Y_n)\rangle=\langle X_1-\theta_1,\ \gamma_n'(Y_n)\rangle+\sum_{j=1}^{n-1}\langle\theta_j-\theta_{j+1},\ \gamma_n'(Y_n)\rangle$$

$$-\sum_{j=1}^{n-1}a_j\langle\Phi_j(X_j,\ Y_j),\ \gamma_j'(Y_j)\rangle\,.$$

Since X_1 and θ_n's are constant and $|\mathrm{E}\gamma_n'(Y_n)|=0$, it follows that

$$|\mathrm{E}\langle X_n-\theta_n,\ \gamma_n'(Y_n)\rangle|\leqq\sum_{j=1}^{n-1}a_j|\mathrm{E}\langle\Phi_j(X_j,\ Y_j),\ \gamma_n'(Y_n)\rangle|\,. \tag{4.34}$$

Since $\Phi_j(X_j,\ Y_j)$ and $\gamma_n'(Y_n)$ are measurable with respect to $\bigvee_{1\leqq i\leqq j}\mathcal{A}_i$ and $\bigvee_{n\leqq i}\mathcal{A}_i$ respectively, it follows from Lemma 4 and its remark that, for $1\leqq j\leqq n-1$

$$\begin{aligned}&|\mathrm{E}\langle\Phi_j(X_j,\ Y_j),\ \gamma_n'(Y_n)\rangle|\\ &\leqq 2\phi^{1/2}\Big(\bigvee_{1\leqq i\leqq j}\mathcal{A}_i,\ \bigvee_{n\leqq i}\mathcal{A}_i\Big)(\mathrm{E}|\Phi_j(X_j,\ Y_j)|^2)^{1/2}\times(\mathrm{E}|\gamma_n'(Y_n)|^2)^{1/2}\\ &\leqq(\mathrm{E}|\Phi_j(X_j,\ Y_j)|^2+\mathrm{E}|\gamma_n'(Y_n)|^2)\lambda_{n-j}\end{aligned} \tag{4.35}$$

where

$$\lambda_n=\sup_m\phi^{1/2}\Big(\bigvee_{1\leqq i\leqq m}\mathcal{A}_i,\ \bigvee_{n+m\leqq i}\mathcal{A}_i\Big),\qquad n\geqq 1\,.$$

Substitutieg (4.35), (4.33) and C2(i) into (4.34) we obtain

$$|\mathrm{E}\langle X_n-\theta_n,\ \gamma_n'(Y_n)\rangle|\leqq\sum_{j=1}^{n-1}\alpha_1a_j\delta_j^{-1}\lambda_{n-j}+\sum_{j=1}^{n-1}a_j\lambda_{n-1}\mathrm{E}\,\beta_j(Y_j)$$

$$+\sum_{j=1}^{n-1}a_j\lambda_{n-j}\mathrm{E}|\gamma_n'(Y_n)|^2\,.$$

The mixing condition C3 yields

$$\sum_{j=1}^{n-1}\lambda_{n-j}\leqq\sum_{n=1}^{\infty}\lambda_n<\infty\quad\text{and}\quad\lim_{n\to\infty}\lambda_{n-j}=0 \tag{4.36}$$

for any fixed j.

Hence, using Toeplitz lemma (see, e. g., Loéve [2], page 238), it follows from A3$'$ and (4.36) that

$$\lim_{n\to\infty}\sum_{j=1}^{n-1}\alpha_1a_j\delta_j^{-1}\lambda_{n-j}=0\,.$$

And A1, C2(ii) and (4.36) imply

$$\lim_{n\to\infty}\sum_{j=1}^{n-1}a_j\lambda_{n-j}\mathrm{E}\,\beta_j(Y_j)=0\,.$$

Moreover, A1, C1(iii) and (4.36) imply

$$\lim_{n\to\infty}\sum_{j=1}^{n-1}a_j\lambda_{n-j}\mathrm{E}|\gamma_n'(Y_n)|^2=0\,.$$

Thus (4.32) is proved and the proof of the theorem is complete.

REMARK. C3 is only used to prove (4.32). In the proof of (4.24), C3 can be replaced by B6.

5. Examples.

In this section we shall give two examples of our results. The following Example 1 is the direct application of Theorem 2.

EXAMPLE 1. Let us chose $\{a_n\}$ so that $a_n=n^{-1}$. Let us assume that $\Phi_n(x, y)$ can be expressed in the form of

$$\Phi_n(x, y)=M_n(x)+\Gamma_n(y)$$

where $M_n(\cdot)$ and $\Gamma_n(\cdot)$ are $\boldsymbol{R}^N$-balued Borel functions defined on $\boldsymbol{R}^N$ and $\boldsymbol{R}^M$, respectively. Suppose that there exist positive constants A_0, A_1 and A_2 such that

$$\langle x-\theta_n, M_n(x)\rangle \geqq A_0|x-\theta_n|^2,$$

$$|M_n(x)|\leqq A_1|x-\theta_n|^2+A_2.$$

Moreover, suppose that

(5.1) $$|\mathrm{E}\Gamma_n(Y_n)|=0,$$

(5.2) $$\sup_n \mathrm{E}|\Gamma_n(Y_n)|^2<\infty,$$

(5.3) $$\lim_{n\to\infty} n|\theta_n-\theta_{n+1}|=0.$$

Let δ be a positive number satisfying $0<\delta\leqq\frac{1}{3}$. Define $\delta_n=n^{-\delta}$. Then A1, A2, A3′, A4 and A5 are satisfied. Define $\gamma_n(\cdot)=\Gamma_n(\cdot)$, $\beta_n(\cdot)=2A_2+2|\Gamma_n(\cdot)|^2$, $\alpha_0=A_0$ and $\alpha_1=2A_1$. Then it is easily seen that C1 and C2 are satisfied. Moreover, it C3 is satisfied then it follows from Theooem 2 that (4.23) and (4.24) hold.

REMARK. If we can assume that, for all $n\geqq 2$,

(5.4) $$|\mathrm{E}[\Gamma_n(Y_n)|Y_1, \cdots, Y_{n-1}]|=0 \quad \text{a.s.},$$

so that $\{\Gamma_n(Y_n)\}_{n=2}^{\infty}$ is a sequence of martingale differences, then the condition (1.3) is automatically satisfied. Hence this case is the usual Robbins-Monro stochastic approximation. But it has not been assumed in Example 1 that (5.4) holds.

Finally we shall give the following example which is an application of Theorem 1.

EXAMPLE 2. Let $\{a_n\}$ be chosen so that $a_n=n^{-1}$. Let δ be a positive number satisfying $0<\delta\leqq\frac{1}{7}$. Define $\delta_n=n^{-\delta}$. Then A1 to A4 are satisfied. And suppose that all vectors considered here are to be row vectors. And let us assume that $\Phi_n(x, y)$ can be expressed in the form of

$$\Phi_n(x, y)=(x-\theta_n)A_n(y)+\Gamma_n(y).$$

where $A_n(y)$ is an $N\times N$-matrix whose (i, j)-th element is $a_n^{(i,j)}(y)$ which is a real valued Borel function defined on $\boldsymbol{R}^M$ and $\Gamma_n(y)=(\Gamma_n^1(y), \cdots, \Gamma_n^N(y))$ is a $\boldsymbol{R}^N$-valued

Borel function defined on $\boldsymbol{R}^M$. Suppose that the following conditions are satisfied.

(5.5)
$$\langle x-\theta_n,\ (x-\theta_n)A_n(y)\rangle \geqq 0 .$$

(5.6)
$$\langle x-\theta_n,\ (x-\theta_n)\mathrm{E}A_n(Y_n)\rangle \geqq \alpha_0 |x-\theta_n|^2$$

where α_0 is some positive constant. And

(5.7)
$$\sup_n \mathrm{E}\|A_n(Y_n)\|^2 < \infty$$

where $\|A\|$ denotes the operator norm of an $N\times N$-matrix A, i.e., $\|A\| = \sup_{|x|\leqq 1}|xA|$. Moreover, suppose that (5.1) to (5.3) are satisfied. Let us denote $x=(x^1, \cdots, x^N)$ and $\theta_n=(\theta_n^1, \cdots, \theta_n^N)$. And let us write $(x^i-\theta_n^i)(x^j-\theta_n^j)=x_{n,i,j}$ and $a_n^{(i,j)}(y)-\mathrm{E}a_n^{(i,j)}(Y_n) = A_n^{(i,j)}(y)$. Define $N_0=N+N^2$, $\varepsilon_n=0$, $f_n(x, y)=\langle x-\theta_n, \Gamma_n(y)\rangle$,

$$F_n(x)=(x_{n,1,1},\ x_{n,1,2},\ \cdots,\ x_{n,2,1},\ \cdots,\ x_{n,N,N},\ x^1-\theta_n^1,\ \cdots,\ x^N-\theta_n^N)$$

and

$$G_n(y)=(A_n^{(1,1)}(y),\ A_n^{(1,2)}(y),\ \cdots,\ A_n^{(2,1)}(y),\ \cdots,\ A_n^{(N,N)}(y),\ \Gamma_n^1(y),\ \cdots,\ \Gamma_n^N(y)).$$

Since $\langle x-\theta_n, \Phi_n(x, y)\rangle = \langle x-\theta_n, (x-\theta_n)A_n(y)\rangle + \langle x-\theta_n, \Gamma_n(y)\rangle$ and

$$\langle x-\theta_n,\ \Phi_n(x, y)\rangle = \langle x-\theta_n,\ (x-\theta_n)\mathrm{E}A_n(Y_n)\rangle + \langle x-\theta_n,\ (x-\theta_n) \times (A_n(y)-\mathrm{E}A_n(Y_n)) + \Gamma_n(y)\rangle ,$$

(5.5) and (5.6) imply B1. And the fact $\mathrm{E}\|A_n(Y_n)-\mathrm{E}A_n(Y_n)\|^2 \leqq 4\mathrm{E}\|A_n(Y_n)\|^2$ and (5.1), (5.2), (5.3) imply B4. Note that

$$|\langle x-\theta_n,\ \Gamma_n(y)\rangle| \leqq \frac{n^{-\delta}}{2}|x-\theta_n|^2 + \frac{n^{\delta}}{2}|\Gamma_n(y)|^2$$

$$|x_{n,i,j}| \leqq |x-\theta_n|^2$$

$$\begin{aligned}|x_{n,i,j}-x^0_{n,i,j}| &= |(x^i-x_0^i)(x^j-\theta_n^j)+(x^j-x_0^j)(x_0^i-\theta_n^i)| \\ &\leqq (|x-\theta_n|+|x_0-\theta_n|)|x-x_0|\end{aligned}$$

and

$$\begin{aligned}|x_{n,i,j}-x_{n+1,i,j}| &= |(x^i-\theta_n^i)(\theta_{n+1}^j-\theta_n^j)+(x^j-\theta_n^j)(\theta_{n+1}^i-\theta_n^i) \\ &\quad +(\theta_n^i-\theta_{n+1}^i)(\theta_n^j-\theta_{n+1}^j)| \\ &\leqq 2|\theta_n-\theta_{n+1}|\,|x-\theta_n|+|\theta_n-\theta_{n+1}|^2 .\end{aligned}$$

And define $g_n(y)=\frac{n^{-\delta}}{2}$, $h_n(y)=\frac{n^{\delta}}{2}$, $b_n=0$, $c_n=2|\theta_n-\theta_{n+1}|$ and $d_n=|\theta_n-\theta_{n+1}|^2 + |\theta_n-\theta_{n+1}|$. Then it is easily seen that B2 and B3 are satisfied. Futhermore, define $\alpha_n(y)=2\|A_n(y)\|^2$ and $\beta_n(y)=2|\Gamma_n(y)|^2$. Then (5.2) and (5.7) imply B5. Hence, if B6 is satisfied then (4.18) holds. And (5.1) and (5.6) imply that the equation (1.1) has the unique solution $x=\theta_n$. Futhermore, (5.7) implies

$$\lim_{n\to\infty} |M_n(X_n)|=0 \quad \text{a.s.},$$

where $M_n(x)=\mathrm{E}\Phi_n(x, Y_n)=(x-\theta_n)\mathrm{E}A_n(Y_n)$. Thus our analysis is general enough to include our previous result which was discussed in [7] as special case.

References

[1] IOSIFESCU, M. and THEODORESCU, R., Random Processes and Learning, Springer, New York, 1969.

[2] LOÈVE, M., Probability Theory (3rd edition), Van Nostrand, New York, 1963.

[3] ROBBINS, H. and MONRO, S., *A stochastic approximation method*, Ann. Math. Statist., **22** (1951), 400-407.

[4] VENTER, J. H., *On Dvoretzky stochastic approximation theorem*, Ann. Math. Statist., **37** (1966), 1534-1544.

[5] WASAN, M. T., Stochastic Approximation, University Press, Cambridge, 1969.

[6] WATANABE, M., *On asymptotically optimal algorithms for pattern classification problems*, Bull. Math. Statist., **15** (1973), 31-48.

[7] WATANABE, M., *An almost sure convergence theorem in a stochastic approximation method with dependent random variables*, Bull. Math. Statist., **19** (1979), 95-112.