数理解析研究所講究録
557巻 1985年 188-202

Adaptive procedure under general conditions

福岡大 理 渡辺正文 (Masafumi Watanabe)

§1. 序

不十分な先験情報のもとでの最適化問題の解を求める方法の一つとして"適応"とか"学習"の考え方がある。確率近似法はその中で確率的な手法の代表的なものとして研究されている。始めは統計的推定の一問題として与えられた (H. Robbins and S. Monro : A stochastic approximation method. Ann. Math. Statist. 22, 1951) がその後逐次的な確率的アルゴリズムとして適応及び学習の面でその有効性が注目された。また、現在はそれらの応用面を考察しながら確率近似法を含むより一般的な逐次的な確率アルゴリズムの研究がなされている。この報告では基本的な Robbins-Monro 型の確率的アルゴリズムの収束について考察する。

確率近似法の収束定理を応用しようとするとき、多くの問題においては観測列 (利用出来る r.v.'s の列) は独立であるこ

とが要求される．従属の場合にも適用出来る収束定理の研究は最近になり研究されている（文献 [1]～[10]）．最初は応用上重要な線形な場合に対し与えられ（[2]），一般的な場合へと拡張されている（[1], [4], [5], [8], [9] etc.）．しかし，独立の場合に比べると条件は限定されたもので，一般的なものとはなっていない．従来の確率近似法の条件に観測列の従属性の条件を加えるだけではうまくいかない．

観測列に対する仮定は以下のタイプが考えられる．

A. 独立性，マルチンゲール性 …… 従来の確率近似法

B. mixing 条件，weak dependent 性 …… [1], [6], [7], [9], [10]

C. summability（和の収束性） ⎫ …… [2], [3], [8], [9]

D. stability，大数の法則，エルゴード性 ⎭

E. A～D より一般的な仮定 …… [4], [6]

従来の確率近似法は A のタイプのもとで考えられている．従属性の仮定として B～E を考える．一般にモーメントに関する適当な条件のもとで，$B \Rightarrow C \Rightarrow D$ が成立する．確率論的な面より考察すると D のタイプが妥当と思われるが，この報告では D より若干一般的な条件（E のタイプと考えられる）のもとでの収束定理を与える．さらに，C のタイプでの収束定理もあわせて考察する．また考える収束のタイプは a.s. 収束（確率 1）である．a.s. 収束の場合は，条件 C～E のタイ

プのもとでは本質的に *deterministic* な議論となり，確率論的な手法は用いられない．確率論的な考察はC～Eを導く時に必要となる．

平均収束に関しては文献[1]，[7]，[9]，[10]の中で考察されている．これらにおいては従属性の条件はいずれもBのタイプである．C以下での条件で考察するのがこれからの問題と思われるが，モーメントの評価を直接必要とするため仲々困難である．

この報告における"general conditions"の意味は従属性の条件C～Eと回帰関数（未知関数）に関する一般性を意味する．

§2. Robbins - Monro stochastic approximation

$(\Omega, \mathcal{A}, P)$：確率空間．以下，考えるr.v.'sは全てこの空間上で定義されているものとする．

H：可分な実Hilbert空間（観測空間）

内積を$\langle \cdot, \cdot \rangle$，ノルムを$\|\cdot\|$で表す．

$M(x)$：$H \to H$ Borel measurable transformation

$Y_n(x)$，$x \in H$，$n = 1, 2, \cdots$：Hの値をとるxをパラメータとするrandom elements（観測）

特に，$Y_n(x) = M(x) + Z_n(x)$，$x \in H$，$n = 1, 2, \cdots$と表せるとする．

ここで，$Z_n(x)$はHの値をとるxをパラメーターとするrandom element（観測誤差）

R－M procedure ：

方程式 $M(x)=0$ の解 $x=\theta$ を求める手法は以下で与える，

$$\begin{cases} X_1 = H\text{の任意の要素} \\ X_{n+1} = X_n - a_n Y_n(X_n) = X_n - a_n\{M(X_n)+Z_n(X_n)\}, \quad n=1,2,\cdots, \end{cases}$$

ここで $\{a_n\}$ は非負実数列.

[1]，[9] においては $H=R^m$ の場合，$Z_n(x)=\sum_{j=1}^{N} m_j(n,x)V_j(n)$ と展開出来る場合を考察している，ただし，$\{m_j(n,x)\}$ は non-random で $\{V_j(n)\}$ の mixing－条件（タイプ B）のもとで procedure の a.s. 及び平均収束を考察している．[9] においては stability（$a_n\delta_n^{-3}\|\sum_{k=1}^{n}V_j(k)\|\longrightarrow 0$ as $n\to\infty$, $j=1,2,\cdots,N$ ；タイプ D）のもとで a.s. 収束を示した．また，$M(x)$ の条件は最も一般的な条件を仮定している．

[2]，[3] では $M(x)=Ax+B$，$Y_n(x)=A_n x+B_n$ の仮定のもとで

$$\lim_{n\to\infty} n^{-1}\sum_{j=1}^{n}A_j = A \quad , \quad \lim_{n\to\infty} n^{-1}\sum_{j=1}^{n}B_j = B$$

の場合（タイプ D）を考察している．

[4]，[8] では $H=R^m$ の時，$\sup_n \|X_n\|<\infty$ (a.s.)，$M(x)$ は連続という仮定のもとで E のタイプ（特に，[8] では D のタイプ）の条件のもとで，a.s. 収束を考察している．また，[6] では

[2], [3] と同じ線型の場合, $\lim_{n\to\infty}\sum_{j=1}^{n} a_j \prod_{k=j+1}^{n}(1-a_k) A_j = A$, $\lim_{n\to\infty}\sum_{j=1}^{n} a_j \prod_{k=j+1}^{n}(1-a_k) B_j = B$ (タイプ E) の仮定のもとで, a.s. 収束を考察している.

この報告では $\sup_n \|X_n\| < \infty$ (a.s.) 及び $M(x)$ の連続性は仮定しない. また, 従属性の条件は C (定理 A), E (定理 B) で与える. またさらに, 収束は a.s. 収束のみを考える.

§3. 基本定理

この章では次章の主結果を導く基本定理を与える.

基本定理. 以下の条件 (i), (ii), (iii) を仮定する.

(i) $a_n > 0\ (n=1,2,\cdots)$, $\lim_{n\to\infty} a_n = 0$, $\sum_{n=1}^{\infty} a_n = \infty$,

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

(ii) $v_n \geqq 0\ (n=1,2,\cdots)$, $\sum_{n=1}^{\infty} v_n < \infty$

(iii) $\{w_n\}$ は実数列で, $\{x_n\}$ は非負実数列で

$$x_{n+1} \leqq (1-a_n+v_n)\, x_n + w_n, \quad n=1,2,\cdots$$

この時, 次の事が成立する.

[I] $\forall T>0$, $\exists\, \omega(T)>0$;

$$\sup_n \max_{n\leq k< m(n,T)} \left|\sum_{j=n}^{k} w_j\right| \leqq \omega(T),$$

ここで, $m(n,T) = \max\{k \mid \sum_{j=n}^{k} a_j \leqq T\}$. ・・・・・(3.1)

$\Rightarrow$ $\exists\, T_0>0$, $\exists\, W_0>0$ (W_0 は T_0 に関係せず) ;

$$\sup_n x_n \leqq \{\omega(T_0)+1\}\, W_0 \quad \cdots\cdots (3.2)$$

〔II〕 $\forall T>0$ に対し

$$\lim_{n\to\infty} \max_{n\le k<m(n,T)} \left|\sum_{j=n}^{k} w_j\right| = 0$$

$$\Rightarrow \quad \lim_{n\to\infty} x_n = 0 \qquad \cdots\cdots \quad (3.3)$$

証明） N を $a_n \le 1/2 \quad (n \ge N)$ となる正整数とする．

$$1-a_j+v_j = (1-a_j)\{1+(1-a_j)^{-1}v_j\} \le (1-a_j)(1+2v_j), \quad j \ge N$$

$$A(m,n) \equiv \begin{cases} \prod_{j=m}^{n}(1-a_j)(1+2v_j) & \text{if} \quad n \ge m \\ 1 & \text{if} \quad n=m-1 \end{cases}$$

仮定 (ii) より $\quad 1 \le \prod_{j=m}^{\infty}(1+2v_j) \le \prod_{j=1}^{\infty}(1+2v_j) \equiv \bar{v} < \infty$ ．

従って

$$A(m,n) \le \bar{v}\exp\left(-\sum_{j=m}^{n} a_j\right), \quad N \le m \le n \qquad \cdots\cdots (3.4)$$

$T_0>0$ を次の様にとる，

$$T_0 \ge 2, \quad \bar{v}/\exp(-T_0/2) \le 1/2$$

$m(n,T_0)$ の定義 (3.1) より $\quad T_0 \ge \sum_{j=n}^{m(n,T_0)-1} a_j > T_0/2 \quad (n \ge N)$

従って $\quad A(n, m(n,T_0)-1) \le \bar{v}\exp(-T_0/2) \le 1/2 \quad (n \ge N)$

次に，以下の補題が成立することに注意する．

[補題]

$$\sup_{n} \max_{n\le k<m(n,T_0)} \left|\sum_{j=n}^{k} w_j\right| \le \omega(T_0)$$

$$\Rightarrow \quad \sup_{n\ge N} \max_{n\le k<m(n,T_0)} \left|\sum_{j=n}^{k} w_j A(j+1,k)\right| \le (1+\bar{v}\bar{A}_N)\,\omega(T_0),$$

ここで，

$$\bar{A}_N = \sup_{n\ge N}\left\{\sum_{j=N}^{n} a_j \prod_{k=j+1}^{n}(1-a_k)\right\} < \infty .$$

次に，$\{x_n\}$ の部分列 $\{x_{n_r}\}_{r=0}^{\infty}$ を次の様に定義する．

$$n_0 \equiv N \ , \qquad n_r = m(n_{r-1}, T_0) \ , \quad r = 1, 2, \cdots$$

このとき，

$$x_{n_0} = x_N$$

$$x_{n_1} \leqq A(n_0, n_1 - 1)\, x_{n_0} + \Big| \sum_{j=n_0}^{n_1-1} w_j A(j+1, n_1-1) \Big| \leqq x_N + 2(1+\bar{v}\bar{A}_N)\,\omega(T_0)$$

$$x_{n_2} \leqq \tfrac{1}{2}\, x_{n_1} + (1+\bar{v}\bar{A}_N)\,\omega(T_0) \qquad \leqq x_N + 2(1+\bar{v}\bar{A}_N)\,\omega(T_0)$$

$$\vdots$$

故に，
$$\sup_r x_{n_r} \leqq x_N + 2(1+\bar{v}\bar{A}_N)\,\omega(T_0) \qquad \cdots\cdots (3.5)$$

任意の $n \geqq N$ に対しては，$\exists r$; $n_r \leqq n < n_{r+1}$ ，従って，$\{x_n\}$ の仮定と (3.5) より

$$x_n \leqq A(n_r, n-1)\, x_{n_r} + \max_{n_r \leq k < n_{r+1}} \Big| \sum_{j=n_r}^{k} w_j A(j+1, k) \Big|$$

$$\leqq \bar{v}\, x_{n_r} + (1+\bar{v}\bar{A}_N)\,\omega(T_0)$$

$$\leqq 3\bar{v}\,(1+\bar{v}\bar{A}_N)(x_N + \omega(T_0))$$

故に，$\sup_n x_n \leqq \{1+\omega(T_0)\} W_0$ ，ここで，$W_0 = \max\{3\bar{v}(1+\bar{v}\bar{A}_N),\ 3\bar{v}(1+\bar{v}\bar{A}_N)\bar{x}_N\}$，$\bar{x}_N = \max\{x_1, x_2, \cdots, x_N\}$ ．従って，[I] が示された．

次に [II] を示す．N を [I] の証明と同じものとする．さらに，$0 < \varepsilon < 1/2$ とし，T_1 を次の様にとる，

$$T_1 \geqq 2 \ , \qquad \bar{v}/\exp(T_1/2) < \varepsilon$$

[I] の証明と同様にして

$$A(n, m(n, T_1) - 1) \leqq \bar{v} \exp[-T_1/2] < \varepsilon < 1/2 \qquad \cdots\cdots (3.6)$$

が示される. [III] の仮定より, $\exists N_1 = N_1(\varepsilon, T_1) \geq N$;

$$\sup_{n \geq N_1} \max_{n \leq k < m(n,T_1)} \left| \sum_{j=n}^{k} w_j \right| < \varepsilon$$

が成立する. 従って, [I] の補題と同様にして

$$\sup_{n \geq N_1} \max_{n \leq k < m(n,T_1)} \left| \sum_{j=n}^{k} w_j A(j+1,k) \right| \leq \varepsilon(1+\bar{v}\bar{A}_N) \quad \cdots\cdots (3.7)$$

が示される.

次に, 部分列 $\{x_{n_r}\}_{r=0}^{\infty}$ を次の様に定義する,

$$n_0 \equiv N_1, \qquad n_r = m(n_{r-1}, T_1), \quad r=1,2,\cdots$$

すると, $N \leq N_1$, $\sup_n x_n \leq W_1 = \{1+\omega(T_0)\}W_0$, (3.6), (3.7) より $0<\varepsilon<1/2$ に注意して

$$x_{n_1} \leq \varepsilon(W_1+1+\bar{v}\bar{A}_N) < 2\varepsilon(W_1+1+\bar{v}\bar{A}_N)$$

$$x_{n_2} \leq 1/2\, x_{n_1} + \varepsilon(W_1+1+\bar{v}\bar{A}_N) < 2\varepsilon(W_1+1+\bar{v}\bar{A}_N)$$

$$\vdots$$

$$x_{n_r} \leq 2\varepsilon(1+W_1+\bar{v}\bar{A}_N), \quad r=1,2,\cdots$$

一方, $n \geq N_1$ なる n に対しては, $\exists r \geq 1$: $n_r \leq n < n_{r+1}$ となることより, (3.7) を用いて

$$\begin{aligned} x_n &\leq A(n_r,n)x_{n_r} + \max_{n_r \leq k < n_{r+1}} \left| \sum_{j=k}^{n} w_j A(j+1,k) \right| \\ &\leq \bar{v}\, x_{n_r} + 2\varepsilon(1+W_1+\bar{v}\bar{A}_N) \\ &\leq 3\varepsilon\bar{v}\,(1+W_1+\bar{v}\bar{A}_N), \qquad n \geq N_1 \end{aligned}$$

故に, $\overline{\lim_{n\to\infty}}\, x_n \leq 3\varepsilon\bar{v}\,(1+W_1+\bar{v}\bar{A}_N)$.

従って, $x_n \geq 0$ より (3.3) が示された. ∎

注意. $a_n \downarrow 0$ とし, $w_n = a_n w_n'$ $(n=1,2,\cdots)$ とする.
以下の条件を考える,

(1) $\sum_{n=1}^{\infty} a_n w_n'$ 収束 (タイプ C)

(2) $\lim_{n\to\infty} a_n \sum_{j=1}^{n} w_j' = 0$ (タイプ D)

(3) $\lim_{n\to\infty} \sum_{j=1}^{n} a_j \prod_{k=j+1}^{n} (1-a_k) w_j' = 0$ (タイプ E, [6])

(4) $\lim_{n\to\infty} \max_{n \le k < m(n,T)} \left| \sum_{j=n}^{k} a_j w_j' \right| = 0$ (タイプ E)

この時, 以下の事が成立する.

イ) クロネッカーの補題より, (1) $\Rightarrow$ (2)

$|a_n^{-1} - a_{n+1}^{-1}| \le K$ より, (2) $\Rightarrow$ (3)

又, (3) $\Rightarrow$ (4) も示される. 従って

(1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4).

ロ) $a_n = n^{-1}$ のときは, $(n-2)e^T \le m(n,T) \le ne^T$ となることより, (4) $\Rightarrow$ (2). また, (3) $\Leftrightarrow$ (2) は明らか. 従って

(2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4).

ハ) $a_n = n^{-1}$ とならぬ場合は必ずしも (4) $\Rightarrow$ (2) は成立しない. 例えば, $a_n = n^{-\alpha}$, $w_n' = n^{-\beta}$, $0<\alpha,\beta$, $\alpha+\beta=1$. このとき, 明らかに (1), (2) は成立しない. しかし,

$$m(n,T) \sim \{n^{1-\alpha} + (1-\alpha)T\}^{1/(1-\alpha)}$$

となることより

$$\sum_{j=n}^{m(n,T)} a_j w_j' = \sum_{j=n}^{m(n,T)} j^{-1} \sim \log m(n,T) - \log n$$
$$= (1-\alpha)^{-1} \log\{1+(1-\alpha)Tn^{-(1-\alpha)}\} \to 0$$

従って，(4) は成立する．

上の事に注意すると，条件 (4) はより一般的な条件，タイプEと考える事が出来る．しかし，Dのタイプ（(2)）と本質的には差がないともいえる．

§4. A.s. convergence of R-M procedure

この章では §2 で与えた手法の収束定理を与える．定理Aは §1で述べたタイプCに対応するもので，$M(x)$ に関する条件は最も弱いものと考えられる．定理Bは $M(x)$ に関する条件は定理Aよりも強くなるが，タイプEに対応するものである．

以下の仮定を与える．

A0 : $^\exists$ positive r.v.'s の列 $\{\delta_n\}$;

(i) $\delta_n \downarrow 0$ (a.s.) , $\sup_n \delta_n \delta_{n+1}^{-1} < \infty$ (a.s.)

(ii) $\sup_n \delta_n \|X_n\| < \infty$ (a.s.)

A1 : $\sum_{n=1}^{\infty} a_n^2 < \infty$, $\sum_{n=1}^{\infty} a_n = \infty$

A2 : $^\exists$ constant $C > 0$;

$$\|M(x)\| \leqq C(\|x\| + 1) \quad , \quad x \in H$$

A3 : $^\exists \theta \in H$, $M(\theta) = 0$;

$$\inf_{\varepsilon < \|x-\theta\| < \varepsilon^{-1}} \langle M(x), x-\theta \rangle > 0 \qquad (0 < \varepsilon < 1)$$

A4 : $^\exists$ nonnegative r.v.'s の列 $\{\alpha_n\}$;

(i) $\|Z_n(x)\| \leqq \alpha_n (\|x\| + 1)$, $x \in H$, $n = 1, 2, \cdots$

(ii) $\displaystyle\sum_{n=1}^{\infty} a_n^2 \alpha_n^2 < \infty$ (a.s.)

A5 : $^{\exists}$ positive r.v. γ_0 ;

(i) $\displaystyle\sup_{n,k} \| \sum_{j=n}^{k} a_j \delta_j^{-3} Z_j(x) \| \leqq \gamma_0 (\|x\| + 1)$, $x \in H$

(ii) $\displaystyle\sup_{n,k} \| \sum_{j=n}^{k} a_j \delta_j^{-3} \{ Z_j(x) - Z_j(y) \} \| \leqq \gamma_0 \|x - y\|$, $x, y \in H$

(iii) $\displaystyle\sum_{n=1}^{\infty} a_n \delta_n \alpha_n < \infty$ (a.s.)

B0 : A0 が成立.

B1 : A1 の他に, $\sup_n |a_n^{-1} - a_{n+1}^{-1}| < \infty$ が成立.

B2 : A2 が成立.

B3 : $^{\exists}\theta \in H$, $M(\theta) = 0$; $^{\exists}\lambda > 0$

$$\langle M(x), x - \theta \rangle \geqq \lambda \|x - \theta\|^2 , \quad x \in H$$

B4 : A4 が成立

B5 : $^{\forall}T > 0$, $^{\exists}$ positive r.v. γ_T :

(i) $\displaystyle\sup_n \max_{n \le k < m(n,T)} \| \sum_{j=n}^{k} a_j \delta_j^{-2} Z_j(x) \| \leqq \gamma_T (\|x\| + 1)$, $x \in H$

(ii) $\displaystyle\sup_n \max_{n \le k < m(n,T)} \| \sum_{j=n}^{k} a_j \delta_j^{-2} \{ Z_j(x) - Z_j(y) \} \| \leqq \gamma_T \|x - y\|$, $x, y \in H$

(iii) $\displaystyle\sup_n \sum_{j=n}^{m(n,T)-1} a_j \alpha_j \leqq \gamma_T$ (a.s.)

<u>注意</u> 仮定 A0 (= B0) は, $^{\exists}$ positive r.v.'s $\{\tau_n\}$;

$$\langle \theta - x, Y_n(x) \rangle \leqq \tau_n (\|x\| + 1)$$

$$\sum_{n=1}^{\infty} a_n \delta_n \tau_n < \infty \quad \text{(a.s.)}$$

が成立すると, [9] の Lemma 1 を用いて得ることが出来る.

仮定Aに対応して次の定理が成立する。

定理A. 仮定 A0〜A5 が成立するならば

$$\lim_{n\to\infty} \|X_n-\theta\| = 0 \quad (a.s.)$$

略証) アルゴリズムより

$$\|X_{n+1}-\theta\|^2 = \|X_n-\theta\|^2 - 2a_n \langle X_n-\theta,\ M(X_n)+Z_n\rangle + a_n^2\|M(X_n)+Z_n\|^2$$

$$\leq (1+Ka_n^2+Ka_n^2\alpha_n^2)\|X_n-\theta\|^2 - 2a_n\langle X_n-\theta,\ M(X_n)+Z_n\rangle$$

$$+ Ka_n^2(\alpha_n^2+1) \quad ,\quad n=1,2,\cdots,$$

ここで，$Z_n \equiv Z_n(X_n)$，K は正定数。このとき，

$$\sum_{n=1}^{\infty} a_n \langle X_n-\theta,\ Z_n\rangle \quad \text{converges } (a.s.)$$

が仮定 A0, A5 を用いて示され，[9] の定理1と同様の方法で結論を得る。∎

仮定Bに対応して次の定理が成立する。

定理B. 仮定 B0〜B5 が成立するならば

$$\lim_{n\to\infty} \|X_n-\theta\| = 0 \quad (a.s.)$$

略証) B3 より，

$$\|X_{n+1}-\theta\|^2 \leq (1-2\lambda a_n+Ka_n^2+Ka_n^2\alpha_n^2)\|X_n-\theta\|^2 - 2a_n\langle X_n-\theta,\ Z_n\rangle$$

$$+ Ka_n^2(\alpha_n^2+1) \quad ,\quad n=1,2,\cdots$$

を得る。また，B0, B1, B4, B5 を用いて，$\forall T>0$, $\exists\, \omega_1(T)$;

$$\sup_n \max_{n\leq k<m(n,T)} \left| \sum_{j=n}^{k} a_j \langle X_j-\theta,\ Z_j\rangle \right| \leq \omega_1(T) \quad (a.s.)$$

を得る。従って，§3の基本定理の[I]より，$\exists$ positive r.v. W_0 ;

$$\sup_n \| X_n - \theta \| \leqq \{ \omega_1(T) + 1 \} W_0 \qquad \cdots\cdots (4.1)$$

を得る．B5 (i), (ii) より次の事が成立する，$\exists$ positive r.v.'s $\{\tau_n\}$;

$$\lim_{n\to\infty} \tau_n = 0 \qquad (a.s.)$$

(i)′ $\displaystyle \max_{n\leq k<m(n,T)} \| \sum_{j=n}^{k} a_j Z_j(x) \| \leqq \tau_n (\|x\| + 1), \quad n=1,2,\cdots$

(ii)′ $\displaystyle \max_{n\leq k<m(n,T)} \| \sum_{j=n}^{k} a_j \{ Z_j(x) - Z_j(y) \} \| \leqq \tau_n \|x-y\|, \quad n=1,2,\cdots$

従って，(4.1) と (i)′, (ii)′, B5 (iii) より，$\forall T>0$ に対して

$$\lim_{n\to\infty} \max_{n\leq k<m(n,T)} \left| \sum_{j=n}^{k} a_j \langle X_j - \theta , Z_j(X_j) \rangle \right| = 0 \qquad (a.s.)$$

が示され，基本定理の[II]を用いて結論を得る．▮

注意．B0の代りに，$\sup_n \|X_n\| < \infty \quad (a.s.)$ が成立する場合は定理Bにおいて，B5は次の仮定B5′に置き代えることが出来る．

B5′: $\forall T>0$, $\exists$ positive r.v.'s $\{\tau_n\}$;

(i) $\displaystyle \max_{n\leq k<m(n,T)} \| \sum_{j=n}^{k} a_j Z_j(x) \| \leqq \tau_n (\|x\| + 1), \quad x\in H, \quad n=1,2,\cdots$

(ii) $\displaystyle \max_{n\leq k<m(n,T)} \| \sum_{j=n}^{k} a_j \{ Z_j(x) - Z_j(y) \} \| \leqq \tau_n \|x-y\|, \quad x,y\in H, \quad n=1,2,\cdots$

(iii) B5の (iii) が成立．

(iv) $\displaystyle \lim_{n\to\infty} \tau_n = 0 \qquad (a.s.)$

証明は定理Bの証明の後半部分と同じ方法で，基本定理の[II]を用いて示される．

References

[1] Borodin, A.N. : A stochastic approximation procedure in the case of weakly dependent observations. Theory Prob. Appl. 24, 34 – 52 (1979).

[2] Fritz, J. : Learning from an ergodic training sequence. In "Limit Theorems of Probability Theory"; ed. P. Révész. North Holland, 79 – 91 (1974).

[3] Györfi, L. : Stochastic approximation from ergodic sample for linear regression. Z. Wahrscheinlichkeitstheorie verw. Geb. 54, 47 – 55 (1980).

[4] Kushner, H.J. and Clark D.S. : Stochastic approximation methods for constrained and unconstrained systems. Springer 1978.

[5] Ljung, L. : Strong convergence of a stochastic approximation algorithms. Ann. Statist. 6, 680 – 696 (1978).

[6] Ljung, L. : Analysis of stochastic gradient algorithms for linear regression problem. IEEE Information Theory IT-30, 151 – 160 (1984).

[7] Eweda, E. and Macchi, O. ; Quadratic mean and almost-sure convergence of unbounded stochastic approximation algorithms with correlated observations. Ann. Institut Henri Poincaré, 19 (1983).

[8] Metivier, M. and Priouret, P. : Application of a Kushner and Clark lemma to general classes of stochastic algorithms. IEEE Information Theory IT-30, 140-151 (1984).

[9] Watanabe, M. : A stochastic approximation from dependent observations. Z. Wahrsheinlichkeitstheorie verw. Geb. 62, 279-292 (1983).

[10] Watanabe, M. : The 2r-th mean convergence of adaptive filters with stationary dependent random variables. IEEE Information Theory IT-30, 134-140 (1984).