数理解析研究所講究録
第 299 巻 1977 年 160-174

従属確率変数列が与えられた場合の確率近似法

福岡大 理　渡辺正文

§1. 序

R-M stochastic approximation method ([5], [7], [8]) ;

$\{Y_n(x)\}_{n=1}^{\infty}$ を x をパラメータとする確率変数列とし，分布は未知とする。　$E[Y_n(x)] = M_n(x)$ とおくとき，十分大なる n に対して方程式

(1.1) $$M_n(x) = 0$$

の根 $x=\theta_n$（存在するとする）を求める R-M procedure は次の形で与えられる．

(1.2) $$X_0 = 0$$

$$X_{n+1} = X_n - a_{n+1} Y_{n+1}(X_n) \quad , \quad n = 0, 1, 2, \cdots\cdots$$

ここで，$\{a_n\}_{n=1}^{\infty}$ は正の実数列で

(1.3) $$\sum_{n=1}^{\infty} a_n = \infty \quad , \quad \sum_{n=1}^{\infty} a_n^2 < \infty \quad .$$

このとき，

(1.4) $$E[Y_{n+1}(X_n) \mid X_1, X_2, \cdots, X_n] = M_{n+1}(X_n) \quad a.s. \quad , n=0,1,\cdots$$

なる仮定の下で

(1.5) $\lim_{n\to\infty} |X_n - \theta_n| = 0$ a.s. , $\lim_{n\to\infty} E|X_n - \theta_n|^2 = 0$

が成り立つための条件が色々と研究されてきている。また，確率近似法（S.A.）は統計的構造があまり良く知られていない系において，なんらかの意味で最適な解を逐次的に推定出来るという点から，学習制御の面に応用されている。S.A.を応用しようとする場合，例えば，パターン分類問題における，識別関数の学習，与えられた input と output の列により，未知の関数を学習する問題等においては，仮定（1.4）の成立の下でS.A.を用いる事により，その収束が示される。すなわち，（1.4）が満たされる自然な仮定として，観測列（学習列）は独立であることが要求される。独立でない場合は，かならずしも（1.4）が成り立つとはかぎらない。

本報告は（1.4）がかならずしも満たされぬ場合のS.A.を与える。すなわち，観測 r.v.'s の列が従属の場合を仮定する。一般に，任意の従属 r.v.'s の列に対して，（1.5）が成り立つための条件を求めることは困難であろう。しかし，ある種のエルゴート性（mixing condition）が満たされる場合は可能であることを示す。本報告では，従属確率変数列の「大数の法則」が成り立つ条件の下でのS.A.を考える。また，収束は a.s. 収束について示す。

§.2 準 備

・R^N, R^M を各々 N 次元 Euclid 空間, M 次元 Euclid 空間とする.

・$\{Y_n\}_{n=1}^{\infty}$ を R^M の値をとる random vectors の列（観測列）とする.

・ $\Phi_n(x, y) = (\Phi_n^{(1)}(x, y), \cdots\cdots, \Phi_n^{(N)}(x, y))$, $n = 1, 2, \cdots$ を $R^N \times R^M \to R^N$ の可測変換とする。

〈問題〉

十分大なる n に対して，方程式

$$(2.1) \qquad E[\Phi_n(x, Y_n)] = 0$$

の根 $x = \theta_n$ を求めよ。但し, $E[\Phi_n(x, Y_n)]$ は未知とし，その代りに，各 n と (x, Y_n) に対し，$\Phi_n(x, Y_n)$ の観測値（雑音混り）

$$(2.2) \qquad Y_n(x) = \Phi_n(x, Y_n) + Z_n$$

を利用出来る。ここで，Z_n は <u>x に無関係な分布</u>をする R^N の値をとる random vector とする。

上の問題に対し，根 θ_n（存在するとする）を求める procedure は (1.2) の R-M procedure を用いる。すなわち，

〈Procedure〉

$$(2.3) \qquad X_0 \equiv 0$$

$$X_{n+1} = X_n - a_{n+1} Y_{n+1}(X_n), \quad n = 0, 1, 2, \cdots\cdots$$

ここで，

$$(2.4) \qquad Y_{n+1}(X_n) = \Phi_{n+1}(X_n, Y_{n+1}) + Z_{n+1},$$

$\{a_n\}_{n=1}^{\infty}$ は正の実数列とする。

記号

- $\langle \cdot , \cdot \rangle$ ；R^Nの内積
- $\| \cdot \|$ ；R^Nのノルム，内積より induced される Euclid norm.
- $\| \cdot \|_N$ ；N×N-matrix のノルム，i.e. A : $N \times N$ - matrix

$$\|A\|_N = \sup_{\|x\|=1} \langle Ax , Ax \rangle^{\frac{1}{2}}$$

- I ；単位行列（恒等変換）
- $\mathbb{O}$ ；零ベクトルまたは，零変換.

さらに，ベクトルは「たて」ベクトルとする。

Lemma 1.

$\| \cdot \|_0$ を $\|A\|_0 = \max_{1 \leq i,j \leq N} |a_{ij}|$, $A = (a_{ij})$; $N \times N$ - matrix で定義されるノルムとすると，2つのノルム $\| \cdot \|_N$, $\| \cdot \|_0$ は同値となる。すなわち，$\exists \alpha , \beta > 0$; $\alpha \|A\|_0 \leqq \|A\|_N \leqq \beta \|A\|_0$

Lemma 2. ([7], [8])

$\{x_n\}_{n=1}^{\infty}$, $\{\alpha_n\}_{n=1}^{\infty}$, $\{\beta_n\}_{n=1}^{\infty}$, $\{K_n\}_{n=1}^{\infty}$ を非負の実数列とし，以下条件が成り立つとする，

(i) $$x_n = \sum_{k=1}^{n} \alpha_k \beta_k \prod_{t=k+1}^{n} (1-\alpha_t) + \sum_{k=1}^{n} K_k \prod_{t=k+1}^{n} (1-\alpha_t) , \quad n=1,2,\cdots ,$$

(ii) $$\lim_{n\to\infty} \alpha_n = 0 , \quad \sum_{n=1}^{\infty} \alpha_n = \infty ,$$

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

(iv) $$\sum_{n=1}^{\infty} K_n < \infty$$

このとき， $\lim_{n\to\infty} x_n = 0$ が成り立つ，さらに，$\lim_{n\to\infty} \beta_n = {}^{\exists}\beta < \infty$ が成り立つときは，$\{x_n\}_{n=1}^{\infty}$ は有界列となる。

Lemma 3. (Kronecker の lemma [2])

$\{x_n\}_{n=1}^{\infty}$ を実数列とし，$\left|\sum_{n=1}^{\infty} x_n\right| < \infty$ とする。このとき，任意の 0 に収束する単調減少数列 $\{k_n\}_{n=1}^{\infty}$ に対して，

$$\lim_{n\to\infty} k_n \sum_{i=1}^{n} x_i/k_i = 0 \quad .$$

Lemma 4.

X ; 線形ノルム空間

$\{x_n\}_{n=0}^{\infty}$, $\{u_n^{(1)}\}_{n=1}^{\infty}$, $\{u_n^{(2)}\}_{n=1}^{\infty}$, $\{v_n\}_{n=1}^{\infty}$; X の 4 つの列

$\{T_n\}_{n=1}^{\infty}$; $X \to X$ の有界線形作用素の列で次の事を満たす，

$$x_{n+1} = T_{n+1} x_n + u_{n+1}^{(1)} + u_{n+1}^{(2)} + v_{n+1} \,, \quad n = 0, 1, 2, \cdots \quad .$$

さらに，以下の条件を満たす，正の実数列 $\{\alpha_n\}_{n=1}^{\infty}$, $\{\beta_n\}_{n=1}^{\infty}$，非負の実数列 $\{M_n\}_{n=1}^{\infty}$，正の定数 K_i $(i=1,2,3,4,5)$ および，正整数 n_0 が存在するとする，

(i) $\quad \lim_{n\to\infty} \alpha_n = 0 \,, \quad \sum_{n=1}^{\infty} \alpha_n = \infty \,,$

(ii) $\quad \sup_{n\geq 1} \left\{ \left| 1/\alpha_{n+1} - 1/\alpha_n \right| \beta_n \right\} \leq K_1 \,,$

(iii) $\quad \sup_{n\geq 1} \beta_n \leq K_2 \,, \quad \sup_{n\geq 1} \beta_n/\beta_{n+1} \leq K_3 \,,$

(iv) $\quad \| T_n T_{n-1} \cdots\cdots T_m T_{m+1} \| \leq K_4 \prod_{k=m+1}^{n} (1-\alpha_k) \,, \quad n > m \geq n_0$

(v) $\quad \| I - T_n \| \leq \alpha_n M_n \,, \quad n \geq 1$

(vi) $\quad \sup_{n\geq 1} \beta_n M_n \leq K_5$ または $\sum_{n=1}^{\infty} \alpha_n \beta_n M_n < \infty$

(vii) $\quad \lim_{n\to\infty} \| u_n^{(1)} \| / \alpha_n = 0 \,, \quad \sum_{n=1}^{\infty} \| u_n^{(2)} \| < \infty$

(viii) $\quad \lim_{n\to\infty} \alpha_n \beta_n^{-1} \left\| \sum_{m=1}^{n} \alpha_m^{-1} v_m \right\| = 0$

このとき，$\lim_{n\to\infty} \| x_n \| = 0$ となる。

<u>proof.</u> $n_0 = 1$ として示す。さらに、$1 > \alpha_n \ (n \geq 1)$ とする。

$$T_m^{(n)} = T_n T_{n-1} \cdots T_{m+1}, \qquad T_n^{(n)} = I \quad \text{とおくと、}$$

$$x_n = \sum_{k=1}^{n} T_k^{(n)} u_k^{(1)} + \sum_{k=1}^{n} T_k^{(n)} u_k^{(2)} + \sum_{k=1}^{n} T_k^{(n)} v_k + T_0^{(n)} x_0$$

ここで、(i)と(iv)より

$$\| T_0^{(n)} x_0 \| \leq K_4 \prod_{k=1}^{n} (1-\alpha_k) \| x_0 \| \longrightarrow 0 \quad \text{as} \quad n \to \infty .$$

(i), (iv), (vii)と Lemma 2. より

$$\sum_{k=1}^{n} \| T_k^{(n)} u_k^{(1)} \| + \| \sum_{k=1}^{n} T_k^{(n)} u_k^{(2)} \|$$

$$\leq K_4 \sum_{k=1}^{n} \prod_{t=k+1}^{n} (1-\alpha_t)(u_k^{(1)} + u_k^{(2)}) \longrightarrow 0 \quad \text{as} \quad n \to \infty .$$

次に、$\lim_{n\to\infty} \| \sum_{k=1}^{n} T_k^{(n)} v_k \| = 0$ を示す。

$$S_n = \alpha_n \sum_{k=1}^{n} \alpha_k^{-1} v_k \quad , \quad n = 1, 2, \cdots$$

$$S_0 \equiv 0 \quad , \quad \alpha_0 = 1$$

とおくと、

$$y_n \equiv \sum_{k=1}^{n} T_k^{(n)} v_k = S_n + \sum_{k=1}^{n-1} (\alpha_k T_k^{(n)} - \alpha_{k+1} T_{k+1}^{(n)}) \alpha_k^{-1} S_k$$

$$= S_n + \sum_{k=1}^{n-1} T_{k+1}^{(n)} (T_{k+1} - I) S_k + \sum_{k=1}^{n-1} \alpha_{k+1} (\alpha_{k+1}^{-1} - \alpha_k^{-1}) T_{k+1}^{(n)} S_k$$

したがって、(ii), (iv), (V) により

$$\| y_n \| \leq \beta_n (\beta_n^{-1} \| S_n \|) + K_4 \sum_{k=1}^{n-1} \alpha_{k+1} \beta_{k+1} M_{k+1} \prod_{t=k+1}^{n} (1-\alpha_t) \beta_{k+1}^{-1} \| S_k \|$$

$$+ K_1 \sum_{k=1}^{n-1} \alpha_{k+1} \prod_{t=k+1}^{n} (1-\alpha_t) \beta_k^{-1} \| S_k \|$$

ここで、(iii), (viii) より

$$\beta_{k+1}^{-1} \| S_k \| = \beta_k \beta_{k+1}^{-1} (\beta_k \| S_k \|) \to 0 \quad \text{as} \quad n \to \infty$$

したがって、Lemma 2 より $\lim_{n\to\infty} \| y_n \| = 0$.

よって、証明された。

注. $\alpha_n = n^{-1}$, $\beta_n = n^{-\beta}$ $(0<\beta<1)$ とすれば条件 (i),(ii),(iii) は満たされる。また，$\{\alpha_n\}_{n=1}^{\infty}$ を (i) を満たす正実数列とし，$\beta_n = \alpha_n^{\delta}$ $(0<\delta<1)$ として，(ii) を仮定すると，条件 (iii) は満たされる。

次の定理は Iosifescu and Theodorescu [3] による。

Lemma 5. (Strong law of large numbers)

$(\Omega, \mathcal{A}, P)$; Pr. space

$\{\mathcal{B}_n\}_{n=1}^{\infty}$; $\mathcal{A}$ の sub-σ-fields の列

$\{X_n\}_{n=1}^{\infty}$; r.v.'s の列で，各 n に対し，X_n は $\mathcal{B}_n$-可測.

(i) ∃正整数 n_1 ; $\overline{\lim_{n\to\infty}}\ \phi\left(\bigvee_{i=1}^{n}\mathcal{B}_i, \bigvee_{i=n+n_1}^{\infty}\mathcal{B}_i\right) < 1$,

ここで，$\phi(\mathcal{A}_1, \mathcal{A}_2) = \sup_{B\in\mathcal{A}_2}\left\{\operatorname{ess\,sup}_{\omega\in\Omega} |P(B|\mathcal{A}_1)(\omega) - P(B)|\right\}$

で $\mathcal{A}_1, \mathcal{A}_2$ は $\mathcal{A}$ の sub-σ-fields とする. (dependent coefficient of $\mathcal{A}_1$ and $\mathcal{A}_2$)

(ii) $\sum_{n=1}^{\infty} \rho_n^{1/2} < \infty$,

ここで，$\rho_n = \sup_{m\geq 1} \phi(\mathcal{B}_m, \mathcal{B}_{m+n})$

(iii) $\{k_n\}_{n=1}^{\infty}$ を $k_n \downarrow 0$ なる正の実数列とし，

$$\sum_{n=1}^{\infty} k_n^2\, E|X_n - EX_n|^2 < \infty$$

ならば，

$$\lim_{n\to\infty} k_n \sum_{i=1}^{n} (X_i - EX_i) = 0 \quad \text{a.s.}$$

Proof.

[3]を参照.

§3. 定理

<u>仮定</u> 以下において，δ, α_0, K_0は正の定数とし，n_0, n_1は正整数，$\{\theta_n\}_{n=0}^{\infty}$は$R^N$の列とする。特に，$\theta_0 = 0$，<u>$0 < \delta < 1/7$</u>とする。

(A 0) $a_n \downarrow 0$, $\sum_{n=1}^{\infty} a_n = \infty$

(A 1) $\sum_{n=1}^{\infty} a_n^{1+\delta} < \infty$

(A 2) $\sup_{n \geq 1} (a_{n+1}^{-1} - a_n^{-1})\, a_n^{\delta} \leqq K_0$

(H 0) $\Phi_n(x, y) = A_n(y)(x - \theta_n) + B_n(y)$, $n = 1, 2, \cdots$

ここで，各 $n = 1, 2, \cdots$ に対して

$$A_n(y) = (A_n^{(i,j)}(y)) \quad : N \times N\text{- matrix}$$

$$B_n(y) = (B_n^{(1)}(y), \cdots, B_n^{(N)}(y))$$

各，$A_n^{(i,j)}(\cdot)$，$B_n^{(i)}(\cdot)$ は Borel measurable function.

(H 1) $\sum_{n=1}^{\infty} a_n^{1+\delta} E\|A_n(Y_n)\|_N^2 < \infty$

(H 2) $\inf_{\|x\|=1} \langle A_n(Y_n)x, x \rangle \geqq 0$ a.s. for $n \geqq n_0$

(H 3) $\inf_{\|x\|=1} \langle E[A_n(Y_n)]x, x \rangle \geqq \alpha_0$ for $n \geqq n_0$

(H 4) $\| E[B_n(Y_n)] \| = 0$ for $n \geqq n_0$

(H 5) $\sum_{n=1}^{\infty} a_n^{1+\delta} E\|B_n(Y_n)\|^2 < \infty$

(H 6) $\lim_{n \to \infty} \|E[Z_n]\| = 0$

(H 7) $\sum_{n=1}^{\infty} a_n^{1+\delta} E\|Z_n - EZ_n\|^2 < \infty$

(H 8) $\|\theta_{n-1} - \theta_n\| = o(a_n)$ または $\sum_{n=1}^{\infty} \|\theta_{n-1} - \theta_n\| < \infty$

$\mathcal{B}_n = \sigma(Y_n, Z_n)$, $n = 1, 2, \cdots$ とする。

(H 9) $\overline{\lim_{n \to \infty}}\ \phi\left(\bigvee_{i=1}^{n} \mathcal{B}_i, \bigvee_{i=n+n_1}^{\infty} \mathcal{B}_i\right) < 1$

(H10) $\sum_{n=1}^{\infty} \rho_n^{1/2} < \infty$

ここで，$\rho_n = \sup_{m \geq 1} \phi(\mathcal{B}_m, \mathcal{B}_{m+n})$

注. 仮定(H0)，(H4)より θ_n は方程式

$$E[\Phi_n(x, Y_n)] = 0 \quad , \quad n \geq n_0$$

の根となる。

Theorem.

仮定(A0)〜(A2)，(H0)〜(H10)が成り立つならば，(2.3)のprocedureの下で，

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

が成り立つ。

Proof.

(2.3)より，

(3.1) $$X_{n+1} - \theta_{n+1} = (I - a_{n+1} A_{n+1})(X_n - \theta_n) + a_{n+1} \left\{ (I - a_{n+1} A_{n+1}) \frac{(\theta_n - \theta_{n+1})}{a_{n+1}} - B_{n+1} - Z_{n+1} \right\}$$

ここで，$A_n = A_n(Y_n)$，$B_n = B_n(Y_n)$ とする．

さらに，

(3.2) $X_n - \theta_n = x_n$，$EA_n = \hat{A}_n$，$EZ_n = \hat{Z}_n$

(3.3) $T_n = I - a_n \hat{A}_n$

(3.4) $U_n = \left\{ (I - a_n \hat{A}_n) a_n^{-1} (\theta_{n-1} - \theta_n) + \hat{Z}_n \right\}$

(3.5) $V_n = \left\{ (\hat{A}_n - A_n) x_{n-1} + (\hat{A}_n - A_n)(\theta_{n-1} - \theta_n) + (\hat{Z}_n - Z_n) - B_n \right\}$

とおくと，

(3.6) $x_{n+1} = T_{n+1} x_n + a_{n+1} U_{n+1} + a_{n+1} V_{n+1}$, $n = 0, 1, 2, \cdots\cdots$

したがって，Lemma 4 において，$\alpha_n = \alpha_0 a_n$，$\beta_n = a_n^{\delta}$ として，Lemma 4 の条件を満たすことを示すとよい。すなわち，以下の事を証明するとよい。

(1) $\lim_{n\to\infty} \|U_n^{(1)}\| = 0$ ，$\sum_{n=1}^{\infty} a_n \|U_n^{(2)}\| < \infty$

ここで，$U_n^{(1)} = a_n^{-1}(\theta_{n-1} - \theta_n) + \hat{Z}_n$

$U_n^{(2)} = \hat{A}_n(\theta_{n-1} - \theta_n)$

(2) $\lim_{n\to\infty} a_n^{1-\delta} \|\sum_{k=1}^{n} V_k\| = 0$ a.s.

(3) $^{\exists}$正定数 C_0，正整数 N_0；

$\|T_m^{(n)}\|_N \leqq C_0 \prod_{t=m+1}^{n} (1 - \alpha_0 a_t)$ ，$n > m \geqq N_0$

ここで，$T_m^{(n)} = T_n T_{n-1} \cdots T_{m+1}$，$T_n^{(n)} = I$ とする．

(4) $\|I - T_n\|_N \leqq \alpha_0 a_n M_n$ ，$\sum_{n=1}^{\infty} a_n^{1+\delta} M_n < \infty$ となる，正の実数列 $\{M_n\}_{n=1}^{\infty}$ が存在する．ここで，$T_n = I - a_n \hat{A}_n$ ．

(1)の証：(H1)，(H6)，(H8)より明らか．

(3)の証：$\{a_n\}_{n=1}^{\infty}$ の単調性，(H1)，(H3)より成り立つ。

(4)の証：$I - T_n = a_n \hat{A}_n$ なることと，(H1)より成り立つ。

(2)の証：$V_k = V_k^{(1)} + V_k^{(2)} + V_k^{(3)} + V_k^{(4)}$ とする，ここで，

(3.7) $V_k^{(1)} = (\hat{A}_k - A_k) x_{k-1}$ ，$V_k^{(2)} = (\hat{A}_k - A_k)(\theta_{k-1} - \theta_k)$

$V_k^{(3)} = Z_k - \hat{Z}_k$ ，$V_k^{(4)} = -B_k$

$i = 1, 2, 3, 4$ に対して，$\lim_{n\to\infty} a_n^{1-\delta} \|\sum_{k=1}^{n} V_k^{(i)}\| = 0$ を示すと(2)は示される。

$i = 3, 4$ に対しては，lemma 5 と(H4)，(H5)，(H7)，(H9)， (H10) および，$0<\delta<1/7$ なることより成り立つ。

$i=2$ のときは，

(3.8) $a_n^{1-\delta}\|\sum_{k=1}^{n}V_k^{(2)}\| \leqq a_n^{1-\delta}\sum_{k=1}^{n}a_k\|\hat{A}_k\|_N a_k^{-1}\|\theta_k-\theta_{k-1}\| + a_n^{1-\delta}\sum_{k=1}^{n}a_k\|A_k\|_N a_k^{-1}\|\theta_k-\theta_{k-1}\|$

より，Lemma 3 (Kronecker) と (H1)，(H8) により，成り立つことがわかる。

$i=1$ のとき，$n_0=1$ として示す。(3.1) と (H1)，(H2) により，次が成り立つ．

(3.9) $$\|x_n\| \leqq L_0\sum_{k=1}^{n}a_k\|R_k\| \quad \text{a.s.}$$

ここで，$R_k=\{(I-a_kA_k)(\theta_{k-1}-\theta_k)-B_k-Z_k\}$，

L_0 は a.s. finite な r.v. である。さらに，次も成り立つ。

(3.10) $$\|x_n-x_{n-1}\| \leqq a_n\|A_n\|_N\|x_{n-1}\| + a_n\|R_n\| \quad \text{a.s.}$$

次に，

(3.11) $$S_n = a_n^{1-\hat{\delta}}\sum_{k=1}^{n}(\hat{A}_k-A_k)\ , \quad n=1,2,\cdots$$

$$S_0=\mathbb{O}\ , \quad a_0=1$$

ここで，$\hat{\delta}$ は $(1-\delta)/2 \geqq \hat{\delta} \geqq 3\delta$ を満たすものとする。

このとき，Lemma 5より (Lemma 1に注意して)

(3.12) $$\lim_{n\to\infty}\|S_n\|_N = 0 \quad \text{a.s.}$$

が成り立つ。さらに，y_n を次の様におく，

(3.14) $$y_n = a_n^{1-\delta}\sum_{k=1}^{n}(\hat{A}_k-A_k)x_{k-1} \quad \left(= a_n^{1-\delta}\sum_{k=1}^{n}V_k^{(1)}\right).$$

このとき，$\hat{A}_k-A_k=(a_k^{1-\hat{\delta}})^{-1}S_k-(a_{k-1}^{1-\hat{\delta}})^{-1}S_{k-1}$ となることより，

(3.15) $$\|y_n\| \leqq a_n^{\hat{\delta}-\delta}\|S_n\|_N\|x_{n-1}\| + a_n^{1-\delta}\sum_{k=1}^{n-1}(a_k^{1-\hat{\delta}})^{-1}\|S_k\|_N\|x_{k-1}-x_k\|$$

となる。よって，(3.9)，(3.10)，(3.12) および，仮定 (H1)，(H5)，(H6)，(H7)，(H8) により，Lemma 3 を用い（$\frac{1-\delta}{2} \geq \delta \geq 3\delta$ に注意して）$\lim_{n\to\infty} \|y_n\| = 0$　a.s. が成り立つ。したがって，$i=1$ のときも成り立ち，定理は証明された。

§4. 応用例

The construction of an unknown function ([4] etc.):

$$y\ (\text{input}) \longrightarrow \boxed{\ f\ } \longrightarrow w = f(y)\ \ (\text{output}),\quad y, w \in R.$$

input y に対し，output $w=f(y)$ が得られる system を考える。ここで，$f(\cdot)$ は未知の実数値 Borel 可測関数とする。

我々は，未知の関数 f を input r.v.'s の列 $\{Y_n\}_{n=1}^{\infty}$ と output r.v.'s の列 $\{W_n\}_{n=1}^{\infty}$，$W_n = f(Y_n)$，により学習（推定）する問題を考える。ここで，$\{Y_n\}_{n=1}^{\infty}$ は独立とは限らない。

$\{\varphi_i(\cdot)\}_{i=1}^{N}$ を実数値 Borel 関数の N ケの組とする。このとき，我々の問題は，

(4.1)　　$$J_n(\mathbb{C}) = E\left[f(Y_n) - \sum_{i=1}^{N} c_i \varphi_i(Y_n) \right]^2$$

を最小にする R^N の vector $\mathbb{C}$ を求めることとなる。ここで，

$$\mathbb{C} = (c_1, c_2, \cdots, c_N)^t,\ t \text{ は転置を表す。}$$

(4.1) を最小にする vector は次の方程式の根となる．

(4.2)　　$$E\left[\Phi(Y_n)\Phi(Y_n)^t \mathbb{C} - f(Y_n)\Phi(Y_n) \right] = \mathbb{O}$$

ここで，$\Phi(\cdot) = (\varphi_1(\cdot), \cdots, \varphi_N(\cdot))^t$ とする。

上の式において

$$\mathscr{G}(y)\mathscr{G}(y)^t = A(y)$$

とおくと，

(4.3) $\langle A(y)\mathbb{C}, \mathbb{C} \rangle \geqq 0 \qquad \forall y \in R, \ \forall \mathbb{C} \in R^N$

となる。さらに，以下の仮定をする：

(i) $f(\cdot)$ ；有界実数値 Borel 関数

(ii) $\mathscr{G}_i(\cdot) \ (i=1,2,\cdots N)$ ；有界実数値 Borel 関数

(iii) $E[A(Y_n)]$ は正値定符号行列

(iv) $\{Y_n\}_{n=1}^{\infty}$ は r.v.'s の列で以下の条件を満足する；

A) 定常過程である，

B) $\overline{\lim_{n\to\infty}} \phi(\bigvee_{i=1}^{n} \mathcal{B}_i, \bigvee_{i=n+n_1}^{\infty} \mathcal{B}_i) < 1$,

C) $\sum_{n=1}^{\infty} \rho_n^{1/2} < \infty$,

ここで， $\mathcal{B}_n = \sigma(Y_n) \quad (n=1,2,\cdots)$.

以上の条件の下では，方程式(4.2)は一意根 $\mathbb{C}^*$を持つ。すなわち， $E[A(Y_n)]\mathbb{C}^* = E[f(Y_n)\mathscr{G}(Y_n)]$.

したがって，定理において

(4.4) $\Phi_n(\mathbb{C}, y) = \mathscr{G}(y)\mathscr{G}(y)^t(\mathbb{C}-\mathbb{C}^*) + \mathscr{G}(y)\{\mathscr{G}(y)^t\mathbb{C}^* - f(y)\}$

すなわち，

$A(y) = \mathscr{G}(y)\mathscr{G}(y)^t, \quad B(y) = \mathscr{G}(y)\{\mathscr{G}(y)^t\mathbb{C}^* - f(y)\}$

$\theta_n = \mathbb{C}^*$ （nに関係しない；定常性より）， $Z_n \equiv \mathbb{O}$,

として，定理の条件を満足することがわかる。

したがって，$\{a_n\}_{n=1}^{\infty}$ を（A 0）～（A 2）を満足する正の実数列として，

$$(4.5) \qquad \mathbb{C}_0 = 0$$

$$\mathbb{C}_{n+1} = \mathbb{C}_n - a_{n+1} Y_{n+1}(\mathbb{C}_n), \quad n = 0, 1, 2, \cdots\cdots,$$

ここで，
$$Y_{n+1}(\mathbb{C}_n) = \mathcal{G}(Y_{n+1})\mathcal{G}(Y_{n+1})^t \mathbb{C}_n - W_{n+1}\mathcal{G}(Y_{n+1})$$
$$= \mathcal{G}(Y_{n+1})\mathcal{G}(Y_{n+1})^t(\mathbb{C}_n - \mathbb{C}^*) + \mathcal{G}(Y_{n+1})\{\mathcal{G}(Y_{n+1})^t(\mathbb{C}_n - \mathbb{C}^*) - f(Y_{n+1})\}, \quad n = 0, 1, 2, \cdots,$$

なる R-M procedure の下で $\lim_{n\to\infty} \|\mathbb{C}_n - \mathbb{C}^*\| = 0$ a.s. が定理の直接の結果として成り立つ。したがって，input 列が従属確率変数列であっても，ある種のエルゴート性が仮定されるならば，独立のときと同じ procedure で同様の結果が得られることがわかる。

参考文献

[1] Doob, J. ; Stochastic processes, Wiley, New York.

[2] Loéve, M.; Probability Theory, Van Nostrand.

[3] Iosifescu, M. and Theodorescu, R.; Random processes and learning, Springer, (1969).

[4] Mendel, J.M. and Fu, K.S. ; Addaptive, learning, and pattern recognition system, Theory and applications, Academic Press, (1970).

[5] Robbins, H. and Monro, S. ; A stochastic approximation method, Ann. Math. Stat., 22,(1951), 400~407.

[6] Stout, W. F. ; Almost sure convergence, Academic Press, (1974).

[7] Wasan, M. T. ; Stochastic approximation method, Cambridge Univ. Press,(1969).

[8] Watanabe, M. ; On Robbins-Monro stochastic approximation method with time varying observations, Bull. Math. Stat., 16, No.3~4, (1975), 73~91.

[9] ———— ; A stochastic approximation method with a sequence of dependent random variables, (to appear).

[10] ———— ; An application of a stochastic approximation method with a sequence of dependent random variables, Fukuoka Univ. Sci. Reports , Vol.6, 43~51,(1976).