

## Convergence of stochastic approximation algorithms with dependent random variables

福岡大 理 渡辺正文 (Masafumi Watanabe)

### §1. 序

統計的構造がほとんど未知な系において最適解を推定する方法の一つとして確率近似法が1950年代より研究されて来ている。そのアルゴリズムは逐次的でかつ簡単であるため応用価値が高い。この報告は前回の研究集会の報告の続きであり特に, Robbins-Monro 型の確率近似アルゴリズムの a.s. 収束と平均収束に関するものである。ここでは, 二次平均収束を中心に議論する。従属確率変数列をもつ確率近似法の研究は, 1970年代半ば頃より研究されている。(文献 [1]~[7], [9]~[11] etc.)。従属の場合に於ける収束はほとんどが a.s. 収束に関しであり, 平均収束に関するものはあまりない, 例えば, 文献 [1], [2], [8]~[11] 位である。独立の場合は両者の収束がほとんど同じ条件下で成立するのに対し, 従属の場合は積のモーメントの評価がうまくいかぬため困難である。この報告に

おいては、平均収束に関しては従属性を3つのタイプの  $\varphi$ -mixing 条件で与える ( $\text{strong-}\varphi\text{-mixing}$ ,  $\varphi\text{-mixing}$ ,  $\text{weak-}\varphi\text{-mixing}$  これらの定義は Stout [8] による)。またここで議論する確率近似法は応用上重要な線型の場合を含む条件の下で考察する。文献 [1] においては弱い独立性 ( $\text{weak-}\varphi\text{-mixing}$ ) の下で a.s. 収束と平均収束が論じられているが、線型の場合を含まぬ。一方、[2] では線型のみを扱っている。しかし、従属性は  $M$ -dependent を仮定している。

## §2. Robbins - Monro 型確率近似法

$R^N$  ;  $N$ -次元 Euclid 空間, 内積  $\langle \cdot, \cdot \rangle$ , ノルム  $\|\cdot\|$

$\mathcal{B}^N$  ;  $R^N$  の Borel field

$(\Omega, \mathcal{A}, P)$  ; 確率空間, 以下考える確率変数は全てこの確率空間上で定義されているとする。

$\{\mathcal{A}_n\}$  ;  $\mathcal{A}$  の sub- $\sigma$ -fields の列,

$$\mathcal{A}_n \stackrel{\text{def.}}{=} \sigma\{\mathcal{A}_m, \mathcal{A}_{m+1}, \dots, \mathcal{A}_n\}$$

$M(x)$  ;  $R^N \rightarrow R^N$ ,  $\mathcal{B}^N$ -可測

$Y_n(x, \omega)$  ;  $R^N \times \Omega \rightarrow R^N$ ,  $\mathcal{B}^N \times \mathcal{A}_n$ -可測,  $n=1, 2, \dots$ ,

$$Z_n(x, \omega) \equiv Y_n(x, \omega) - M(x), \quad n=1, 2, \dots$$

Robbins - Monro stochastic approximation algorithm ;

方程式  $M(x) = 0$  の解 (存在を仮定し, それを  $\theta$  とする)

を観測列  $\{Y_n(\cdot, \omega)\}$  を用いて推定するアルゴリズムを次で与える,

$$\begin{cases} X_1(\omega) = R^d \text{ の定数ベクトル (任意) } \\ X_{n+1}(\omega) = X_n(\omega) - a_n Y_n(X_n(\omega), \omega) \quad , \quad n=1, 2, \dots, \\ \quad \quad \quad (= X_n(\omega) - a_n M(X_n(\omega)) - a_n Z_n(X_n(\omega), \omega) ) \end{cases}$$

ここで,  $\{a_n\}$  は正の単調減少列で

$$(A1) \quad \sum_{n=1}^{\infty} a_n^2 < \infty \quad , \quad \sum_{n=1}^{\infty} a_n = \infty$$

を満たすものとする。このとき, 次の定理は一般的 (M(x) に関する仮定が最も一般的) なものとして知られている。

定理 2.1. (Gladyshev, E. G. : On Stochastic Approximation, Theory of Prob. and its appl. Vol. 10, 1965, 275-278). 条件 (A.1) の他に, 以下を仮定する。

$$(i) \quad \inf_{\varepsilon < \|x - \theta\| < \varepsilon^{-1}} \langle x - \theta, M(x) \rangle > 0 \quad \text{for each } \varepsilon > 0$$

$$(ii) \quad \|M(x)\| \leq K(\|x\| + 1)$$

$$(iii) \quad E\{Y_n(X_n) \mid X_1, \dots, X_n\} = M(X_n) \quad \text{a.s.} \quad n=1, 2, \dots$$

$$(iv) \quad E\{\|Z_n(X_n)\|^2 \mid X_1, \dots, X_n\} \leq K(\|X_n\|^2 + 1), \quad \text{a.s.} \quad n=1, 2, \dots$$

$$\Rightarrow \quad \lim_{n \rightarrow \infty} \|X_n - \theta\| = 0 \quad \text{a.s.} \quad \text{かつ} \quad \lim_{n \rightarrow \infty} E\|X_n - \theta\|^2 = 0$$

$\{a_n\}$  が独立の場合は本質的に (iii) が成立する。 $\{a_n\}$  が独立でない場合, すなわち, (iii) が成立しない場合は上の条件の他にいくつかの条件が必要となる。この報告では  $M(x)$  の条件は

出来る限り変えず,  $M(x)$  以外の条件を強めよ又は新たに加えることにより a.s. 収束と 2 次平均収束について考察する。

### §3. a.s. 収束

定理 3.1. 定理 2.1. の条件 (i), (ii) が成立。さらに,

$$(A2) \quad \sup_n |a_n^{-1} - a_{n+1}^{-1}| < \infty$$

$$(B1) \quad \exists \{\delta_n(\omega)\} : \text{正の確率変数列}$$

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

$$(ii) \quad \sup_n \delta_n(\omega) \|X_n(\omega)\| < \infty$$

$$(B2) \quad \exists \{\alpha_n(\omega)\} : \text{非負確率変数列}$$

$$(i) \quad \|Z_n(x, \omega)\| \leq \alpha_n(\omega) (\|x\| + 1), \quad n = 1, 2, \dots$$

$$(ii) \quad \sum_{n=1}^{\infty} \alpha_n^2 \delta_n^{-1}(\omega) < \infty$$

$$(iii) \quad \sum_{n=1}^{\infty} \alpha_n \delta_n(\omega) \delta_n^{-1}(\omega) < \infty$$

$$(B3) \quad \exists \gamma(\omega) : \text{非負確率変数}$$

$$(i) \quad \sup_n \alpha_n \left\| \sum_{j=1}^n \delta_j^{-1} Z_j(x, \omega) \right\| \leq \gamma(\omega) (\|x\| + 1)$$

$$(ii) \quad \sup_n \alpha_n \left\| \sum_{j=1}^n \delta_j^{-1} \{Z_j(x, \omega) - Z_j(y, \omega)\} \right\| \leq \gamma(\omega) \|x - y\|$$

$$\Rightarrow \quad \lim_{n \rightarrow \infty} \|X_n(\omega) - \theta\| = 0$$

(注) (i) 条件 (B1)(ii) は次の条件で置きかえることができる。

3.

$\exists \{\beta_n(\omega)\}$  ; 非負確率変数列

$$(i) \quad \langle 0-x, Y_n(x, \omega) \rangle \leq \beta_n(\omega) (\|x\| + 1), n=1, 2, \dots$$

$$(ii) \quad \sum_{n=1}^{\infty} a_n \delta_n(\omega) \beta_n(\omega) < \infty$$

この条件を満たす例として次の線型の例がある。

$$M(x) = Ax + b, \quad A \text{ は } N \times N\text{-matrix}, \quad b \in R^N$$

$$Y_n(x, \omega) = A_n(\omega)x + b_n(\omega), \quad A_n(\omega) \text{ は } N \times N\text{-random matrix},$$

$b_n(\omega)$  は random vector, とすると, 条件  $\langle x, A_n(\omega)x \rangle \geq 0$

が成立する場合, (B1) (ii) が満たされる。あるいは, 線型と

有限の場合,  $\sup_{n, x} \|Z_n(x, \omega)\| \leq \delta(\omega)$  が成立する場合も

(B1) (ii) が成立する ([1] はこの場合を扱っている)。

(2) この定理において確率的な議論は必要でない。確率的議論は (B3) を導くために必要となる。

#### S4. 2r 次平均収束

$\mathcal{A}', \mathcal{A}''$  を  $\mathcal{A}$  の sub- $\sigma$ -fields とする。

$$\mathcal{G}_1(\mathcal{A}', \mathcal{A}'') \stackrel{\text{def}}{=} \sup \left\{ \left| \frac{P(A \cap B)}{P(A)P(B)} - 1 \right| ; P(A)P(B) \neq 0, \right.$$

$$A \in \mathcal{A}', B \in \mathcal{A}'' \left. \right\}$$

$$\mathcal{G}_2(\mathcal{A}', \mathcal{A}'') \stackrel{\text{def}}{=} \sup \left\{ |P(B|A) - P(B)| ; P(A) \neq 0, \right.$$

$$A \in \mathcal{A}', B \in \mathcal{A}'' \left. \right\}$$

$$\mathcal{G}_3(\mathcal{A}', \mathcal{A}'') \stackrel{\text{def}}{=} \sup \left\{ |P(A \cap B) - P(A)P(B)| ; A \in \mathcal{A}', \right.$$

$$B \in \mathcal{A}'' \left. \right\}$$

さらに,

$$\phi_1(n) \stackrel{\text{def}}{=} \sup_k \varphi_1(\mathcal{A}_1^k, \mathcal{A}_{k+n})$$

$$\phi_2(n) \stackrel{\text{def}}{=} \sup_k \varphi_2(\mathcal{A}_1^k, \mathcal{A}_{k+n})$$

$$\phi_3(n) \stackrel{\text{def}}{=} \sup_k \varphi_3(\mathcal{A}_1^k, \mathcal{A}_{k+n})$$

(注) (1)  $\phi_i(n)$  は任意の  $n$  に対して有限値に限る。

$$(2) \quad \phi_1(n) \geq \phi_2(n) \geq \phi_3(n).$$

$$(3) \quad \lim_{n \rightarrow \infty} \phi_1(n) = 0 \quad \alpha \in \mathbb{R}, \quad \{\mathcal{A}_n\} \text{ は strong-}\phi\text{-mixing}$$

$$\lim_{n \rightarrow \infty} \phi_2(n) = 0 \quad \alpha \in \mathbb{R}, \quad \{\mathcal{A}_n\} \text{ は } \phi\text{-mixing}$$

$$\lim_{n \rightarrow \infty} \phi_3(n) = 0 \quad \alpha \in \mathbb{R}, \quad \{\mathcal{A}_n\} \text{ は weak-}\phi\text{-mixing}$$

となる。

補助定理 4.1.  $\mathcal{A}', \mathcal{A}'' \in \mathcal{A}$  の sub- $\sigma$ -fields.  $X, Y$  は各々  $\mathcal{A}'$ -

可測,  $\mathcal{A}''$ -可測な確率変数とする。このとき, 以下が成立。

$$(i) \quad E|X| < \infty, \quad E|Y| < \infty$$

$$\Rightarrow |EXY - EXEY| \leq \varphi_1(\mathcal{A}', \mathcal{A}'') E|X| E|Y|$$

$$(ii) \quad 0 < p, q < 1, \quad p^{-1} + q^{-1} = 1, \quad E|X|^p < \infty, \quad E|Y|^q < \infty$$

$$\Rightarrow |EXY - EXEY| \leq 2 \varphi_2^{1/p}(\mathcal{A}', \mathcal{A}'') (E|X|^p)^{1/q} (E|Y|^q)^{1/p}$$

$$(iii) \quad 0 < p, q, s < 1, \quad p^{-1} + q^{-1} + s^{-1} = 1, \quad E|X|^p < \infty,$$

$$E|Y|^s < \infty$$

$$\Rightarrow |EXY - EXEY| \leq 10 \varphi_3^{1/p}(\mathcal{A}', \mathcal{A}'') (E|X|^p)^{1/q} (E|Y|^s)^{1/s}$$

(注) (3)は W. Phillip and W. Stout, "Almost sure invariance principles for partial sums of weakly dependent random variables", Mem. Amer. Math. Soc. 6 (1975), Lemma 7.2.1.

による。(i), (ii), (iii) は  $R^N$  の確率変数に對して同様の結果が成立する。

補助定理 4.2.  $\{X_n\}$  : 確率変数列.  $t \in \text{正数}$  とする.

$$(i) \quad \sup_n E|X_n|^t = c < \infty$$

$$(ii) \quad X_n \rightarrow X \text{ in probability}$$

$$\Rightarrow \lim_{n \rightarrow \infty} E|X_n - X|^{t'} = 0 \quad \text{for } t' < t$$

(注) 補助定理 4.2. は Loève, Probability Theory, p164 にある。

補助定理 4.3.  $\{\xi_n\}, \{u_n\}, \{v_n\}$  : 非負実数列

$$0 < t \leq 1, \quad \sum_{n=1}^{\infty} v_n < \infty,$$

$$\xi_{n+1} \leq (1+v_n)\xi_n + u_n(\xi_n^{1-t} + 1), \quad n=1, 2, \dots$$

$$\Rightarrow \xi_n^t \leq 2^t \left( \prod_{i=1}^n (1+v_i)^t \right) \left( \sum_{i=1}^n u_i \right), \quad n=1, 2, \dots$$

(注) 証明は [11] lemma 1.

(A1), (A2) を §2 及び §3 で与えたものとする。さらに, 以下の仮定を与える。以下,  $0 < u \leq 1$ ,  $p$  は 正整数 とする。

(A3)  $\exists \{\delta_n\}$  : 正実数列

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

$$(ii) \quad \sum_{n=1}^{\infty} a_n \delta_n^u < \infty$$

$$(iii) \quad a_n \delta_n^{-(2p+1)} > a_{n+1} \delta_{n+1}^{-(2p+1)}, \quad n=1, 2, \dots$$

$$(C1) \quad \inf_{\varepsilon < \|x-\theta\|} \langle x-\theta, M(x) \rangle > 0 \quad \text{for each } \varepsilon > 0$$

$$(C2) \quad \|M(x)\| \leq K(\|x\| + 1)$$

(C3)  $\exists F: R^N \rightarrow R^{N_0}$ , Borel-可測,  $F(\theta) = 0$ ,  $\exists \{G_n(\omega)\}$  ;

$R^{N_0}$ -valued r.v.'s a.s., 各  $n$  に於て,  $G_n$  は  $\mathcal{A}_n$ -可測,

$$(i) \quad \langle x-\theta, Z_n(x, \omega) \rangle = \langle F(x), G_n(\omega) \rangle_0, \quad n=1, 2, \dots$$

$$(ii) \quad \|F(x) - F(y)\|_0 \leq K(\|x\| + \|y\| + 1) \|x - y\|$$

(C4)  $\exists \{\alpha_n(\omega)\}$  : 非負随率変数数列, 各  $n$  に於て  $\alpha_n$  は  $\mathcal{A}_n$ -可測,

$$\|Z_n(x, \omega)\| \leq \alpha_n(\omega) (\|x\| + 1), \quad n=1, 2, \dots$$

(C5)  $\exists \{\beta_n(\omega)\}$  ; 非負随率変数数列, 各  $n$  に於て  $\beta_n$  は  $\mathcal{A}_n$ -可測,

$$\langle \theta - x, Y_n(x, \omega) \rangle \leq \beta_n(\omega) (\|x\|^{2-u} + 1), \quad n=1, 2, \dots$$

$$(C6) \quad \sup_n a_n \delta_n^{-(2p+1)} \left\| \sum_{j=1}^n E G_j \right\|_0 < \infty$$

(C7)  $\exists M$  (正整数) ;

$$\sum_{n=M}^{\infty} a_n \delta_n^{-(2p+1)} \phi_1(n) < \infty$$

$$(C8) \quad \sup_n E[\gamma_n^{2p(1+M+\frac{2}{u})}] < \infty,$$

$$\gamma_n(\omega) = \max \{ \alpha_n(\omega), \beta_n(\omega), \|G_n(\omega)\|_0 \}.$$

補助定理 4.4.  $\{V_n\}$  : 非負確率変数列,  $V_n$  は  $\mathcal{A}_n$ -可測。

(C4) : 非負実数列, 以下 E 満たす。

$$(i) \quad \sum_{n=1}^{\infty} C_n < \infty, \quad C_{n+1} \leq C_n \leq 1$$

$$(ii) \quad \exists M \text{ (正整数)} : \phi_1(M) < \infty$$

$$(iii) \quad \sup_n E V_n^{\alpha M} < \infty \quad \text{for } \alpha \geq 1$$

$$\Rightarrow \quad \sup_n E \left[ \prod_{i=1}^n (1 + C_i V_i)^{\alpha} \right] < \infty$$

(注) 証明は [1], Lemma 2.

定理 4.1.  $r$  は正整数とする。(A1)~(A3), (C1)~(C8) が

$p=r+1 \geq 1$  で成立。

$$\Rightarrow \quad \lim_{n \rightarrow \infty} E \|X_n - \theta\|^{2r} = 0$$

略証) 補助定理 4.3, 4.4 を用いて,  $\sup_n \int_n^{\infty} E \|X_n\|^{\alpha} < \infty$  を示す。ここで,  $\alpha = 2p(1 + \frac{u}{pu+2})$ 。この  $\alpha$  を用いて,

$$\text{補助定理 4.1 (i) より,} \quad \sum_{n=1}^{\infty} a_n E \left[ \|X_n\|^{2p-2} < F(X_n), G_n \right]$$

収束, を得る。この結果と, (C1) より,  $\lim_{n \rightarrow \infty} E \|X_n - \theta\|^{2p}$  存在,

$\exists \{X_{n_k}\} \subset \{X_n\}$ ,  $\lim_{k \rightarrow \infty} \|X_{n_k} - \theta\| = 0$  a.s. を得る。従って,

補助定理 4.2 より結論を得る。

次に以下の仮定を与える。

(C4)'  $\exists \{\alpha_n(\omega)\}$  ; 非負確率変数列, 各  $n$  に対して  $\alpha_n(\omega)$  は

$d_n$ -可測,

$$\|\sum_n(x, \omega)\| \leq K(\|x\| + d_n(\omega) + 1), \quad n=1, 2, \dots$$

$$(C7)' \quad \sum_{n=1}^{\infty} a_n \delta_n^{-(2p+1)} \{\phi_2(n)\}^{\frac{2p+1}{2p+2}} < \infty$$

$$(C8)' \quad \sup_n E \delta_n^{\frac{2p+2}{u}} < \infty,$$

$$=: z'', \quad \delta_n(\omega) = \max\{\alpha_n(\omega), \beta_n(\omega), \|\zeta_n(\omega)\|_0\}.$$

定理 4.2. (A1)~(A3), (C1)~(C3), (C4)', (C5), (C6), (C7)', (C8)' が  $p=r+1$  として成立. さらに, (C7)' は  $p=r$  に対して成立するとする.

$$\Rightarrow \lim_{n \rightarrow \infty} E \|X_n - \theta\|^{2r} = 0$$

略証) 補助定理 4.3 を用いて,  $\sup_n \delta_n^{2p} E \|X_n - \theta\|^{2p} < \infty$  が示される.  $\therefore a = z$  を用いて, 補助定理 4.1 (ii) より,

$$\sum_{n=1}^{\infty} a_n E [\|X_n\|^{2p-2} \langle F(X_n), G_n \rangle_0] \quad \text{収束, かつ成り立ち, (C1) を}$$

用いて,  $\lim_{n \rightarrow \infty} E \|X_n - \theta\|^{2p}$  存在 ( $\sup_n E \|X_n - \theta\|^{2p} < \infty$ ),

$\exists \{X_{n_k}\} \subset \{X_n\}, \lim_{k \rightarrow \infty} \|X_{n_k} - \theta\| = 0 \quad \text{a.s.} \quad \exists$  得る. 従って,

補助定理 4.2 より結論を得る.

次に, まえ仮定を与える.

$$(C7)'' \quad \sum_{n=1}^{\infty} a_n \delta_n^{-(2p+1)} \{\phi_2(n)\}^{\frac{1}{2p+2}} < \infty$$

定理 4.3.  $(A1) \sim (A3)$ ,  $(C1) \sim (C3)$ ,  $(C4)'$ ,  $(C5)$ ,  $(C6)$ ,  $(C7)'$ ,  $(C8)'$  が  $p = r+1$  に  $\#1$  で成立.

$$\Rightarrow \lim_{n \rightarrow \infty} E \|X_n - \theta\|^{2r} = 0$$

略証) 定理 4.2. と同様にして, 補助定理 4.1(iii) を用いて示す.

### §5. 平均収束の order

この章では,  $a_n = a n^{-1}$  ( $a > 0$ ),  $n = 1, 2, \dots$  とする。§5 に以下の仮定を与える。

$$(C1)' \quad \exists \lambda > 0; \quad \langle x - \theta, M(x) \rangle \geq \lambda \|x - \theta\|^2$$

$$(C6)' \quad \|E G_n\|_0 = 0, \quad n = 1, 2, \dots$$

定理 5.1.  $(C1)', (C2), (C3) \sim (C5), (C6)', (C8)$  が  $p = 2r$  と  $(\#2)$  成立するとき,

$$\phi_r(n) \leq K n^{-b} \quad (b > 0), \quad n = M, M+1, \dots$$

が成立する。

$$\Rightarrow E \|X_n - \theta\|^{2r} \leq \begin{cases} K n^t (\log n)^2, & 0 < t \leq 1 \\ K n^{-1}, & t > 1 \end{cases}$$

$$t = \min \{2a, b\}.$$

略証) 証明は, [10], [11] と同様の方法を用いて示す。

定理 5.2.  $(C1)', (C2), (C3), (C4)', (C5), (C6)', (C8)'$  が  $p = r+1$  に対して成立, さらに,

$$\phi_2(n) \leq k n^{-b} \quad (b > 0), \quad n = 1, 2, \dots$$

$\Rightarrow$  定理 5.1 の結果が成立する, ただし,  $t = \min\{2\alpha, \frac{b(2r+1)}{2r+2}\}$ .

定理 5.2.  $(C1)', (C2), (C3), (C4)', (C5), (C6)', (C8)'$  が  $p = r+1$  に対して成立, さらに,

$$\phi_2(n) \leq K n^{-b} \quad (b > 0), \quad n = 1, 2, \dots$$

$\Rightarrow$  定理 5.1 の結果が成立する, ただし,  $t = \min\{2\alpha, \frac{b}{2r+2}\}$ .

## References

- [1] Borodin, A. N. : A stochastic approximation procedure in the case of weakly dependent observations. Theory Prob. Appl. 24 (1979), 34-52.
- [2] Eweda, E and Macchi, O. : Quadratic mean and almost-sure convergence of unbounded stochastic approximation algorithms with correlated observations. Ann. Institut. Henri Poincaré, Vol. 19, No. 1 (1983).

- [3] Györfi, L. : Stochastic approximation from ergodic sample for linear regression. Z. Wahrscheinlichkeitstheorie verw. Geb. 54 (1980), 47-55.
- [4] Kushner, H. J. and Clark, D. S. : Stochastic approximation methods for constrained and unconstrained systems. Springer 1978.
- [5] Ljung, L. : Analysis of recursive stochastic algorithms. IEEE Trans. Automat. Contr. Vol. AC-22 (1977), 551 - 575.
- [6] Ljung, L. : Strong convergence of a stochastic approximation algorithms. Ann. Statist. 6 (1978), 680 - 696.
- [7] Métivier, M. and Priouret, P. : Application of a Kushner and Clark lemma to general classes of stochastic algorithms. IEEE Inform. Theory Vol. IT-30 (1984), 140 - 151.
- [8] Stout, W. F. : Almost sure convergence. Academic Press, 1974.
- [9] Watanabe, M. : A stochastic approximation from dependent observations. Z. Wahrscheinlichkeitstheorie verw. Geb. 62 (1983), 279 - 292.

- [10] Watanabe, M. : The  $2r$ -th mean convergence of adaptive filters with stationary dependent random variables . IEEE Inform. Theory Vol. IT-30 (1984) , 134-140 .
- [11] Watanabe, M. : The  $2r$ -th mean convergence of a stochastic approximation algorithm with weakly dependent observations . Fukuoka Univ. Science Reports vol. 16 , No.2 (1986) , 71-80 .