数理解析研究所講究録
48巻 1968年 147-160

Stochastic Approximation の応用に関する基礎的研究

阪大 工 草野シナ子
阪大 工 田坂誠男
阪大 工 杉山 博

序

Robbins-Monroは回帰方程式$M(x)=\alpha$がunique root $x=\theta$を有するとき$x_{n+1}=x_n+a_n(\alpha-y_n)$によって定義される系列$\{x_n\}$が若干の条件のもとに$\theta$に平均収束することを明らかにし、又Kiefer-Wolfowitzは回帰関数$M(x)$が唯一つの最大値θを有するとき、$x_{n+1}=x_n+(a_n/c_n)(y_{2n}-y_{2n-1})$によって定義される$\{x_n\}$が$\theta$に確率収束することを証明した。その後、この問題に関する多くの研究が発表されているが我々は次のような諸問題について若干の結果を得た。ここでは、これらの諸問題を論ずるとともに、それらの応用の可能性をexampleとともに示すことにする。

(1) R-M過程，K-W過程におけるx_nの分布に関する基本方程式を導いて、これを解くことにより$M(x)$がlinearの場合，山一つの場合にfinite stageに於けるx_nの分布を求めること

(2)　回帰方程式 $M(x)=\alpha$ が *multiple roots* を有する場合の R-M 過程の収束について。

(3)　回帰関数 $M(x)$ が *multi-peaks* を有する場合の K-W 過程の収束について。

§1　R-M および K-W 過程における基本方程式とその応用

一次元 R-M *process* X_n の分布関数 $G_n(x|x_1)$ は次の方程式を満足する。

$$G_{n+1}(x|x_1)=P_r\{x_n+a_n(\alpha-y_n)<x\}=1-P_r\{y_n<\alpha-\frac{x-x_n}{a_n}\}$$

$$=1-\int_{-\infty}^{\infty}H(\alpha-\frac{x-\xi}{a_n}|\xi)\,dG_n(\xi|x_1)$$

$$G_2(x|x_1)=1-H(\alpha-\frac{x-x_1}{a_1}|x_1)$$

これを R-M 過程の基本方程式ということにする。さて，

[仮定]　　$M(x)=\lambda x+\mu$

$$H(y|\xi)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{y}e^{-\frac{\{x-M(\xi)\}^2}{2\sigma^2}}dx$$

のもとで基本方程式の解はつぎのようになる。なお一般性を失うことなく $\alpha=0$ とする。

$$g_n(x|x_1)=\frac{1}{(\prod_{i=1}^{n-1}a_i)\sqrt{A_n}\sqrt{2\pi\sigma^2}}\,e^{-\frac{[\prod_{i=1}^{n-1}\frac{1}{a_i}x-B_n]^2}{2A_n\sigma^2}}$$

$$A_n=(\frac{1}{a_{n-1}}-\lambda)^2A_{n-1}+\{\prod_{i=1}^{n-2}\frac{1}{a_i}\}^2$$

$$B_n=(\frac{1}{a_{n-1}}-\lambda)B_{n-1}-\{\prod_{i=1}^{n-2}\frac{1}{a_i}\}\mu$$

さらにつぎの仮定のもとでx_{n+1}の特性関数を求め、x_{n+1}の期待値，分散を求めるとつぎのようになる。

［仮定］ $$H\left(-\frac{x-\xi}{a_n}\middle|\xi\right)=\Phi(t_n(x,\xi))=\frac{1}{\sigma}\int_{-\infty}^{-\frac{x-\xi}{a_n}}\phi\left(\frac{z-M(\xi)}{\sigma}\right)dz$$

ここに $$\Phi(x)=\int_{-\infty}^{x}\phi(t)dt\ ,\quad \phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

$$t_n(x,\xi)=-\frac{x-\xi+a_nM(\xi)}{a_n\sigma}$$

$$\sigma_x^2=\int_{-\infty}^{\infty}\{y-M(x)\}^2dH(y|x)=\sigma^2<\infty$$

このときx_{n+1}の特性関数は，

$$\Phi_{n+1}(t)=\int_{-\infty}^{\infty}e^{itx}dG_{n+1}(x|x_1)=1+itI_n^{(1)}-\frac{t^2}{2!}(a_n^2\sigma^2+I_n^{(2)})+O(t^3)$$

ここで $$I_n^{(\nu)}=\int_{-\infty}^{\infty}(x-a_nM(x))^{\nu}dG_n(x|x_1)$$

$$I_1^{(\nu)}=(x_1-a_1M(x_1))^{\nu}$$

したがってx_{n+1}の期待値および分散は

$$E(x_{n+1}|x_1)=I_n^{(1)}$$

$$V(x_{n+1}|x_1)=a_n^2\sigma^2+I_n^{(2)}-(I_n^{(1)})^2$$

によって与えられる。このことを用いて

(i) *linear case* $M(x)=\beta x$ とすると

$$I_1^{(1)}=(x_1-a_1M(x_1))=(1-a_1\beta)x_1$$

$$I_n^{(1)}=\int_{-\infty}^{\infty}(x-a_nM(x))dG_n(x|x_1)=(1-a_n\beta)I_{n-1}^{(1)}$$

更に，$I_1^{(2)}=(1-a_1\beta)^2x_1^2$

$$I_n^{(2)}=(1-a_n\beta)^2\int_{-\infty}^{\infty}x^2dG_n(x|x_1)=(1-a_n\beta)^2(a_{n-1}^2\sigma^2+I_{n-1}^{(2)})$$

したがってx_1が与えられたとき，x_{n+1}は期待値，分散がそれぞれ

$$E(x_{n+1}|x_1)=\prod_{i=1}^{n}(1-a_i\beta)x_1$$

$$V(x_{n+1}|x_1) = \left[a_n^2 + \sum_{j=2}^{n}\prod_{i=j}^{n}(1-a_i\beta)^2 a_{j-1}^2\right]\sigma^2$$

なる正規分布に従う。

(ii) *linearly bounded case*

[仮定]　$0<\delta_1<\delta_2<\infty$ なる δ_1，δ_2 に対し $\inf_{\delta_1\leq|x|\leq\delta_2}|M(x)|>0$ が成立するものとする。且つ

$$\sup_x |M(x)-\beta x| = \gamma$$

を満足する β，γ に対して

$$-\gamma+\beta x \leqq M(x) \leqq \gamma+\beta x$$

なる場合，$M(x)=\beta x+\delta(x)$ とおくと $-\gamma\leqq\delta(x)\leqq\gamma$ となる。このとき、

$$\prod_{i=1}^{n}(1-a_i\beta)x_1 - \sum_{j=2}^{n}\prod_{i=j}^{n}(1-a_i\beta)a_{j-1}\gamma - a_n\gamma$$

$$\leqq E(x_{n+1}|x_1) \leqq \prod_{i=1}^{n}(1-a_i\beta)x_1 + \sum_{j=2}^{n}\prod_{i=j}^{n}(1-a_i\beta)a_{j-1}\gamma + a_n\gamma$$

更に

$$V(x_{n+1}|x_1) \leqq a_n^2\sigma^2 + \left(|1-a_n\beta|\sqrt{a_{n-1}^2\sigma^2+I_{n-1}^{(2)}} + a_n\gamma\right)^2 - \left((1-a_n\beta)I_{n-1}^{(1)} - a_n\gamma\right)^2$$

このことから我々は *finite stage* における x_n の平均，分散の評価を行うことができる。

次に一次元 K-W process x_n の分布関数 $G_n(x|x_1)$ は方程式

$$G_{n+1}(x|x_1) = \int_{-\infty}^{\infty} H^*\left(\frac{x-\xi}{a_n c_n^{-1}}\,\middle|\,\xi+c_n,\ \xi-c_n\right)dG_n(\xi|x_1)$$

$$G_2(x|x_1) = H^*\left(\frac{x-x_1}{a_1 c_1^{-1}}\,\middle|\,x_1\right)$$

を満足する。$H^*\left(\frac{x-\xi}{a_n c_n^{-1}} \middle| \xi+c_n, \xi-c_n\right) = \frac{1}{\sigma}\int_{-\infty}^{\frac{x-\xi}{a_n c_n^{-1}}} \phi\left(\frac{z-(M(\xi+c_n)-M(\xi-c_n))}{\sigma}\right) dz$

を仮定し R-M 過程の場合と同様に x_{n+1} の特性関数を求めることができて，

$$\Psi_{n+1}(t) = \int_{-\infty}^{\infty} e^{itx} dG_n(x|x_1) = 1 + itI_n^{(1)} - \frac{t^2}{2!}\left[\sigma^2 a_n^2 c_n^{-2} + I_n^{(2)}\right] + O(t^3)$$

ここで
$$I_n^{(\nu)} = \int_{-\infty}^{\infty} \left(x + a_n c_n^{-1}(M(x+c_n) - M(x-c_n))\right)^{\nu} dG_n(x|x_1)$$
$$I_1^{(\nu)} = \left(x_1 + a_1 c_1^{-1}(M(x_1+c_1) - M(x_1-c_1))\right)^{\nu}$$

したがって x_{n+1} の期待値および分散は

$$E(x_{n+1}|x_1) = I_n^{(1)}$$
$$V(x_{n+1}|x_1) = a_n^2 c_n^{-2}\sigma^2 + I_n^{(2)} - (I_n^{(1)})^2$$

によって与えられる。

§2　$M(x)=\alpha$ が *multiple roots* を有する場合．

[定理 1]

$M(x)=\alpha$ が *multiple roots* $\theta_1, \theta_2, \cdots, \theta_\nu$ を有するとき条件 I ～ V のもとで R-M process $x_{n+1} = x_n + a_n(\alpha - y_n)$ によって定義される系列 $\{x_n\}$ について

$$\sum_{i=1}^{\frac{\nu+1}{2}} \Pr\left\{\lim_{n\to\infty} x_n = \theta_{2i-1}\right\} = 1$$

が成立する。

[条件]

I　$|M(x)| \leqq d|x-A| + d|x-B|$, $(d>0,\ A<B,\ -\infty<x<\infty)$

II　$x<\theta_1$ のとき $M(x)<\alpha$

$x > \theta_\lambda$ のとき $M(x) > \alpha$

III $\quad \sigma_x^2 = \int_{-\infty}^{\infty} (y - M(x))^2 \, dH(y \mid x) \leqq \sigma^2 < \infty$

IV $\quad 0 < \delta_1 < \delta_2 < \infty$ となるような δ_1, δ_2 に対して

$$\inf_{\bigcap_{i=1}^{\nu} \{\delta_1 \leqq |x - \theta_i| \leqq \delta_2\}} |M(x) - \alpha| > 0$$

V $\quad$ 正数列 $\{a_n\}$ について

① $\sum a_n = \infty$, ② $\sum a_n^2 < \infty$

この定理を証明するために次の補助定理を用いる。

[補題 1]

条件IIIおよびVの②が成立つとき確率変数列

$$\left\{ x_{n+1} - \sum_{j=1}^{n} a_j (\alpha - M(x_j)) \right\}, \qquad n = 1, 2, \cdots$$

は確率1である確率変数に収束する。

[補題 2]

条件I〜IIIのもとで $\{x_n\}$ はある確率変数に確率1で収束する。

[補題 2 の証明]

x_n を $\lim_{n\to\infty} x_n = +\infty$ となるような標本系列とすると $x_n \leqq \theta_\nu$ となる n は高々有限個でそれ以外のものについては $x_n > \theta_\nu$ となるからII, Vによって、$a_n(\alpha - M(x_n))$ が負となる。それゆえ、

$$\lim_{n\to\infty} \left\{ x_{n+1} - \sum_{j=1}^{n} a_j (\alpha - M(x_j)) \right\} = +\infty$$

となる系列が存在することになる。これは補題1に矛盾する。

よって $P_r\{\lim_{n\to\infty} x_n = +\infty\} = 0$ 。同様にして

$P_r\{\lim_{n\to\infty} x_n = -\infty\} = 0$ が証明できる。

したがって，$\{x_n\}$ がある確率変数に確率1で収束しないとすれば，次の2つの性質をもつ標本系列が存在する確率が正である。

(α) $\quad x_{m+1} - \sum_{j=1}^{m} a_j(\alpha - M(x_j))$ はある有限な数に収束する。

(β) $\quad \liminf x_n < \limsup x_n$

いまこのような標本系列の1つをとって，$\limsup x_n > \theta_\nu$ としよう。（$\theta_i < \limsup x_n \leqq \theta_{i+1}$，$i = 1, 2, \cdots \nu-1$，$\limsup x_n \leqq \theta_1$ の場合も同様に証明できる。）以上のことから

(1) $\quad a > \theta_\nu$，$\quad \liminf x_n < a < b < \limsup x_n$

を満足する a, b をとることができる。又 N を十分大きくとって $N \leqq n < m$ であるすべての m, n について，

(2) $\quad$ ① $\quad a_n \leqq \min\left\{\frac{1}{6d},\ \frac{b-a}{3[|\alpha| + d(|A|+|B|) + 2d|\theta_\lambda|]}\right\}$

$\quad$ ② $\quad \left| x_m - x_n - \sum_{j=n}^{m-1} a_j(\alpha - M(x_j))\right| < \frac{b-a}{3}$

適当な m, n をとって次の条件を満足させるようにできる。

(3) $\quad$ ① $\quad N \leqq n < m$

$\quad$ ② $\quad x_n < a$，$\quad x_m > b$

$\quad$ ③ $\quad n < j < m$ のとき $\quad a \leqq x_j \leqq b$

しかるに (1)～(3) はこれから示すように矛盾を生ずる。まず (1)～(3)，と条件Ⅱによって

$$x_m - x_n \leqq \frac{b-a}{3} + \sum_{j=n}^{m-1} a_j (\alpha - M(x_j))$$
$$\leqq \frac{b-a}{3} + a_n (\alpha - M(x_n))$$

ここで $\theta_\nu < x_n$ とすると、$\alpha - M(x_n) < 0$ であるから

$$x_m - x_n < \frac{b-a}{3}$$

となって(3)の②と矛盾する。

$\theta_\nu \geqq x_n$ とすると，$B < \theta_\nu$ に対して、

$$|M(x_n)| \leqq 2d|x_n| + d(|A|+|B|)$$
$$\leqq d(|A|+|B|) + 2d|\theta_\nu| + 2d|x_m - x_n|$$
$$x_m - x_n \leqq \frac{b-a}{3} + a_n[|\alpha| + d(|A|+|B|) + 2d|\theta_\nu| + 2d|x_m - x_n|]$$
$$x_m - x_n \leqq \frac{2(b-a)}{3(1-2a_n d)} \leqq b-a$$

$B > \theta_\nu$ に対して

$$x_m - x_n \leqq \frac{b-a}{3} + a_n[|\alpha| + d|B-A|]$$
$$< \frac{b-a}{3} + a_n[|\alpha| + d(|B|+|A|) + 2d|\theta_\nu|]$$
$$< b-a$$

となって(3)の②と矛盾する。よって補題2は証明された。

［定理1の証明］

まず $P_r\{\lim_{n\to\infty} x_n = x \neq \theta_i\} > 0$ を仮定すると矛盾することを云う。$\theta_\nu < \varepsilon_1 < \varepsilon_2 < \infty$ であるような ε_1，ε_2 をとり、これについて $P_r\{\varepsilon_1 < x < \varepsilon_2\} > 0$ とすることができる。（$\theta_i < \varepsilon_1 < \varepsilon_2 < \theta_{i+1}$，$i = 1, 2, \cdots, \nu-1$ 。 $-\infty < \varepsilon_1 < \varepsilon_2 < \theta_1$ の場合も同様であ

る。) このとき $\varepsilon_1 < x < \varepsilon_2$ なる x に確率1で収束するすべての標本系列 $\{x_n\}$ については、十分大きなすべての n について $\varepsilon_1 \leqq x_n \leqq \varepsilon_2$ が成立つ。したがって $\theta_\nu < \varepsilon_1 \leqq x_n \leqq \varepsilon_2$ が成立つ系列が正の確率で存在する。このような系列では補題1, 2によって、つねに $\sum_{j=1}^{n} a_j(\alpha - M(x_j))$, $n=1, 2, \cdots$ は収束する。他方条件IVおよびVの①によると、この級数は発散する。これは矛盾である。よって、

$$P_r\{\lim_{n\to\infty} x_n = x \neq \theta_i\} = 0$$

次に $P_r\{\lim_{n\to\infty} x_n = \theta_{2i}\} > 0$ が成立するものと仮定すると $\theta_{2i} < a < b < \delta < \theta_{2i+1}$ に対して十分大きいNをとることにより $N \leqq n < m$ なるすべての m, n に対して

$$0 \leqq |x_m - \theta_{2i}| < a - \theta_{2i} < b - \theta_{2i} < x_n - \theta_{2i} < \delta - \theta_{2i}$$

かつ、$m > j > n$ に対して

$$a - \theta_{2i} < |x_j - \theta_{2i}| < b - \theta_{2i}$$

を満足する標本系列 $\{x_n\}$ が存在する。一方 Blum の補題1を適用すると $\left|x_m - x_n - \sum_{j=n}^{m-1} a_j(\alpha - M(x_j))\right| < b - a$ とすることができる。故に

$$|x_m - x_n| < b - a$$

よって矛盾を生ずる。

$$P_r\{\lim_{n\to\infty} x_n = \theta_{2i}\} = 0$$

[系]

定理1の条件IIを $x<\theta_1$ のとき $M(x)>\alpha$，$x>\theta_\nu$ のとき $M(x)<\alpha$ とすると

$$P_r\{\lim x_n=-\infty\}+\sum P_r\{\lim x_n=\theta_{2i}\}+P_r\{\lim x_n=\infty\}=1$$

［定 理 2］

［条 件］　$0<K_1<K_2<\infty$ に対し

(i)　$K_1|x-\theta_i|<|M(x)-\alpha|<K_2|x-\theta_i|$

(ii)　$|y-M(x)|<C$

のもとで

$$|x_j-\theta_i|<\rho,\quad \min(|\underline{\rho}|\ |\tilde{\rho}|)>\rho>\frac{C}{K_1}\quad,\quad a_j<\frac{1}{K_2}$$

ならば $|x_N-\theta_i|<\rho$ （整数 $N\geqq j+1$）

この条件を満足する $\rho>0$ をとると，x_j が区間 $[\theta_i-\rho,\ \theta_i+\rho]$ に入ったとき，$j+1$ 以後のすべての N について x_N はこの区間内に留まる。さて，区間 $[\theta_i-\rho,\ \theta_i+\rho]$ 内では，

$$-\infty<R_1<\alpha-y_i<R_2<\infty\quad,\quad \lim_{n\to\infty}a_n=0 \text{ より}$$

$$\lim_{n\to\infty}\{x_{n+\mu}-x_n\}=\lim_{n\to\infty}\sum_{i=n}^{n+\mu-1}a_i(\alpha-y_i)=0\quad(\text{a. e.})$$

よって，$N\geqq j+1$ については x_N は確率1で θ_i に収束する。

ゆえに x_n が θ_i に収束する確率に対して

$$P_r\{\lim_{n\to\infty}x_n=\theta_i\}\geqq G_j(\theta-\rho|x_1)-G_j(\theta+\rho|x_1)$$

のような lower bound が定まる。ここで $G_n(x|x_1)$ は，

$$G_{n+1}(x|x_1)=1-\int_{-\infty}^{\infty}H\left(\alpha-\frac{x-\xi}{a_n}\middle|\xi\right)dG_n(\xi|x_1)$$

［Example 1］

$M(x)=x^2-4x-5$, $\alpha=0$, ($\theta_1=-1$, $\theta_2=5$)

$Y(x)=M(x)+\varepsilon$, ここにεは$N(0,1)$に従う確率変数である。定理1の系によるとこのcaseは

$$P_r\{\lim x_n=\theta_2\}+P_r\{\lim x_n=-\infty\}=1$$

となる。モンテカルロ実験の結果は次の表の通りである。

Starting points	n=3		n=10		n=30	
	r_1	r_2	r_1	r_2	r_1	r_2
−3	0	1.00	0	1.00	0	1.00
−1	0.65	0.35	0.478	0.522	0.478	0.522
0	0.991	0.009	0.971	0.029	0.971	0.029
2	0	1.00	0	1.00	0	1.00
5	0.99	0.01	0.97	0.03	0.97	0.03

$r_1\cong P_r\{\lim_{n\to\infty}x_n=\theta_2\}$, $r_2\cong P_r\{\lim_{n\to\infty}x_n=-\infty\}$

[Example 2]

① $M(x)=-x^5+2x^4+16x^3-2x^2-15x$, $\alpha=0$

($\theta_1=-3$, $\theta_2=-1$, $\theta_3=0$, $\theta_4=1$, $\theta_5=5$)

$Y(x)=M(x)+\varepsilon$ 　　$\varepsilon\in N(0,1)$ とする。

② $x=\pm10$に吸収壁を設けた。区間$[-10, 10]$で任意にx_1を与え壁に達したらそこで打切りa_nの係数をかえx_1を変えることにした。

③ 壁によって発散する標本系列を取除いたので結果はθ_2およびθ_4へ収束

④ $M(x)=x^5-2x^4-16x^3+2x^2+15x$ とすると，θ_1，θ_3およびθ_5へ収束

[Example 3]　Chebycheff lower bound の一例

$M(x) = x^2 - 4x - 5$

$K_1 = 3$, $K_2 = 8$ とすると

$\min(|\underset{\sim}{\rho}|\ |\bar{\rho}|) = 2$

$$\left.\begin{array}{l} \frac{1}{3} \leqq \rho \leqq 2 \\ a_1 < \frac{1}{8} \\ |x_2 - \theta_2| < \rho \end{array}\right\} \rightarrow |x_N - \theta_2| < \rho \quad (\forall, N \geqq 3)$$

x_1	$G_2(7\mid x_1) - G_2(3\mid x_1)$
0	0
2	0.16
3	0.99
6	0.99
7	0.99
10	0.99
12	0.16

$$P_r\{\lim x_n = \theta_2\} \geqq G_2(7|x_1) - G_2(3|x_1)$$

$$g_2(x|x_1) = \frac{1}{a_1\sigma\sqrt{2\pi}} \exp\left[-\frac{\{x-(x_1-a_1M(x_1))\}^2}{2a_1^2\sigma^2}\right]$$

§3　multi-peaks を有する一次元回帰関数の極値探索

[定理 3]

回帰関数 $M(x)$ が multi-peaks θ_1, θ_2 $\cdots$ θ_ν を有するとき

つぎの条件 I 〜 V のもとで K-W process

$$x_{n+1} = x_n + (a_n/c_n)(y_{2n} - y_{2n-1})$$

によって定義される系列 $\{x_n\}$ について

$$\sum_{i=1}^{\nu} P_r\{\lim x_n = \theta_i\} = 1$$

が成立する。

[条件]

I　$M(x)$ は　$x < \theta_1$ で単調増加

$\theta_i < x < \lambda_i$ で単調減少　　$i = 1, 2, \cdots, \nu-1$

$\lambda_i < x < \theta_{i+1}$ で単調増加

$x > \theta_\nu$ で単調減少

ここで $\theta_1 < \theta_2 < \cdots < \theta_\nu$ $\lambda_1 < \lambda_2 < \cdots < \lambda_{\nu-1}$

Ⅱ $\sigma_x^2 = \int_{-\infty}^{\infty} (y - M(x))^2 dH(y|x) \leqq \sigma^2 < \infty$

Ⅲ 正数 ρ, R が存在して

$|x' - x''| < \rho$ ならば $|M(x') - M(x'')| < R$

Ⅳ すべての $\delta > 0$ に対して正数 $\pi(\delta)$ が存在し

$\bigcap_{i=1}^{\nu}\{|x - \theta_i| > \delta\}$ かつ $\bigcap_{i=1}^{\nu-1}\{|x - \lambda_i| > \delta\}$ において

$$\inf_{\xi > \varepsilon > 0} \frac{|M(x+\varepsilon) - M(x-\varepsilon)|}{\varepsilon} > \pi(\delta)$$

Ⅴ 正数列 $\{a_n\}$, $\{c_n\}$ について

① $\lim_{n\to\infty} c_n = 0$, ② $\sum_{n=1}^{\infty} a_n = \infty$, ③ $\sum_{n=1}^{\infty} (a_n/c_n)^2 < \infty$

この定理の証明には定理1と同様に次の補助定理を用いる。
具体的な証明は定理1と形式的に同じであるので省略する。

[補題 3]

条件Ⅱ, Ⅴが成立つとき

$x_{n+1} - \sum_{j=1}^{n} (a_j/c_j)\{M(x_{2j}) - M(x_{2j-1})\}$

は確率1である確率変数に収束する。

[補題 4]

条件Ⅰ,Ⅱ,Ⅲ,Ⅴのもとで K-W process $\{x_n\}$ は確率1で, ある確率変数に収束する。

おわりに

多次元回帰関数がmulti-peaksを有するとき多次元Kiefer-Wolfowiz processの収束については次の機会に論ずる。

[References]

(1) H. Robbins and S. Monro, "A Stochastic Approximation Method", Annals of Math. Stat., Vol. 22, pp. 400-407.

(2) J. R. Blum, "Approximation Methods which converge with Probability One", Ann. Math. Stat., Vol. 25, pp. 382-386.

(3) J. Wolfowitz, "On the Stochastic Approximation Method of Robbins and Monro", Ann. Math. Stat., Vol. 23, pp. 457-461

(4) M. Loève, "On almost sure Convergence", Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, pp. 279-303.

(5) M. Loève, "Probability Theory", 1955, D. Van Nostrand.

(6) J. Kiefer and J. Wolfowitz, "Stochastic Estimation of the Maximum of a Regression Function", Ann. Math. Stat., Vol. 23, pp. 462-466.

(7) J. H. Venter, "On Convergence of the K-W Approximation Procedure", Ann. Math. Stat., Vol. 38, pp. 1031-1036.