ON RANDOM OBSERVATION PROCESSES FOR STOCHASTIC APPROXIMATION

SIRO YAMAZOE

(Received April 5, 1971; revised June 18, 1973)

Summary

In [1] a new procedure is given to estimate the root of a regression equation. The purposes of this paper are to extend the Lemma 1 in [1] and to give a process involving a randomly determined sequence of observations for finding themaximum of a regression function. The process is similar to that of [1]. Kiefer and Wolfowitz [2] gave a stochastic approximation procedure for the latter purpose. Their process needs the condition of the unimodality of the regression function which is not required for our case.

1. Notation and definition

Let $\left(\Omega=\prod_{i=1}^{\infty}(\mathcal{X}_i\times\mathcal{Y}_i),\ \mathcal{B}\right)$ be a Cartesian product measurable space, where $\mathcal{X}_i=[0,1)$ and $\mathcal{Y}_i=(-\infty,\infty)$. Let $B_{mi}=[(i-1)/2^m, i/2^m)$ $i=1,2,\cdots,2^m$, $m=0,1,\cdots$ and let $\{k(m);\ m=0,1,\cdots\}$ be a strictly increasing sequence of integers and $k(0)=1$.

For each $x\in[0,1)$, $\omega=(x_1, y_1; x_2, y_2;\cdots)\in\Omega$, and $n=1,2,\cdots$, let us define

$$x_n(\omega)=x_n\ ,\qquad y_n(\omega)=y_n$$

and

$$N_n(x,\omega)=k(m_0)\ ,$$

where $m_0=\max\{m;\ \#[j;\ x_j(\omega)\in B_{mi},\ x\in B_{mi} \text{ and } 0<j\leqq n]\geqq k(m)\}$. The symbol $\#[A]$, denotes the number of the elements of the set A. Let

$$B_n(x,\omega)=B_{m_0i_0}\ ,$$

where $N_n(x,\omega)=k(m_0)$ and $x\in B_{m_0i_0}$. Let $C_n(x,\omega)$ be the set of $x_j(\omega)$ belonging to $B_n(x,\omega)$ with the smallest subscripts $k(m_0)$ and define the random functions by

$$M_n(x, \omega)=\frac{1}{N_n(x, \omega)} \sum_{x_j(\omega) \in C_n(x, \omega)} y_j(\omega) ,$$

$$\alpha_n(x, \omega)=h\{N_n(x, \omega)\} \exp\{-N_n(x, \omega)|M_n(x, \omega)-a|^b\} ,$$

$$\alpha'_n(x, \omega)=h\{N_n(x, \omega)\} \exp\left\{-N_n(x, \omega)\left[\max_{x' \in [0,1)} M_n(x', \omega)-M_n(x, \omega)\right]^b\right\} ,$$

where $h(\cdot)>0$ and b is a positive constant,

$$\beta_n(x, \omega)=\alpha_n(x, \omega)\Big/\int_0^1 \alpha_n(x', \omega)dx'$$

and

$$\beta'_n(x, \omega)=\alpha'_n(x, \omega)\Big/\int_0^1 \alpha'_n(x', \omega)dx' .$$

Let $P_{\beta_n(\omega)}$ and $P_{\beta'_n(\omega)}$ be the probability measures on $[0, 1)$ with probability density functions $\beta_n(x, \omega)$ and $\beta'_n(x, \omega)$, respectively. The above notations are well defined.

Let F_x be a probability measure on $(R^1, \mathscr{B}^1)$ for every $x \in [0, 1)$, where $R^1=(-\infty, \infty)$, and $\mathscr{B}^1$ is its Borel σ-field. If

(1) $F_x(A)$ is measurable for every $A \in \mathscr{B}^1$,

then there exists a probability measure P on $(\Omega, \mathscr{B})$ such that, for every $n=1, 2, \cdots$,

(2) if $B \in \mathscr{B}^1$ and $B'=x_{n+1}^{-1}(B)$, then
$$P\{B' | x_i(\omega)=x_i, y_i(\omega)=y_i \quad i=1, 2, \cdots, n\}=P_{\beta_n(\omega)}(B)$$

and such that,

(3) if $B \in \mathscr{B}^1$ and $B'=y_n^{-1}(B)$, then
$$P\{B' | x_i(\omega)=x_i, y_j(\omega)=y_j \quad i=1, \cdots, n \quad j=1, \cdots, n-1\}$$
$$=F_{x_n}(B) ,$$

and further there exists a probability measure P' on $(\Omega, \mathscr{B})$ such that,

(4) if $B \in \mathscr{B}^1$ and $B'=x_{n+1}^{-1}(B)$, then
$$P'\{B' | x_i(\omega)=x_i, y_i(\omega)=y_i \quad i=1, \cdots, n\}=P_{\beta'_n(\omega)}(B)$$

and such that,

(5) if $B \in \mathscr{B}^1$ and $B'=y_n^{-1}(B)$, then
$$P'\{B' | x_i(\omega)=x_i, y_j(\omega)=y_j, \quad i=1, \cdots, n \quad j=1, \cdots, n-1\}$$
$$=F_{x_n}(B) .$$

Since the condition (1) implies the measurability of $M(x)=\int y dF_x(y)$,

$$z_n(\omega)=y_n(\omega)-M\{x_n(\omega)\}$$

can be defined as a random variable for every $n=1, 2, \cdots$, and let

$$M_n'(x, \omega)=\frac{1}{N_n(x, \omega)}\sum_{x_j(\omega)\in C_n(x,\omega)} z_j(\omega) .$$

2. On $\sup_{x,n} |M_n'(x, \omega)|^b N_n(x, \omega)$

Lemma 1 in [1] is essential and especially its condition decides the $k(m)$. It can be extended as follows.

LEMMA 1. *If, for positive integer r,*

$$\int \{y-M(x)\}^{2r} dF_x(y) \leqq V<\infty \tag{6}$$

and

$$\sum_{m=0}^{\infty} 2^m k(m)^{-r(1-2/b)}<\infty \tag{7}$$

then

$$\begin{aligned} (8)\quad & \mathrm{P}\{\omega;\ \sup [|M_n'(x, \omega)|^b N_n(x, \omega):\ n=1, 2, \cdots, x \in [0, 1)]=\infty\} \\ & =\mathrm{P}'\{\omega;\ \sup [|M_n'(x, \omega)|^b N_n(x, \omega):\ n=1, 2, \cdots, x \in [0, 1)]=\infty\} \\ & =0 . \end{aligned}$$

Note. This lemma is available if only a probability measure on $(\Omega, \mathscr{B})$ satisfies the conditions (2) and (3) for some P_β and F_x, that is, it is valid for any form of β_n.

PROOF OF LEMMA 1. It is sufficient to prove the lemma only in case of P. Another part of this lemma follows in an exactly same manner.

Let

$$P_\infty=\mathrm{P}\{\omega;\ \sup [|M_n'(x, \omega)|^b N_n(x, \omega):\ n=1, 2, \cdots, x \in [0, 1)]=\infty\}$$

then

$$\begin{aligned} (9)\quad P_\infty &= \lim_{K\to\infty} \mathrm{P}\{\omega;\ \text{There exist some } i,\ m,\ n \text{ and } x \text{ such that} \\ &\qquad B_{mi}=B_n(x, \omega) \text{ and that } |M_n'(x, \omega)|\, N_n(x, \omega)^{1/b}>K.\} \\ &\leqq \lim_{K\to\infty} \sum_{m=0}^{\infty} \mathrm{P}\{\omega;\ \text{There exist some } i,\ n \text{ and } x \text{ such that} \\ &\qquad B_{mi}=B_n(x, \omega) \text{ and that } |M_n'(x, \omega)|\, N_n(x, \omega)^{1/b}>K.\} . \end{aligned}$$

Let m fix and define the following notations for $i=1, 2, \cdots, 2^m$.

i) For every $A \in \mathcal{B}^1$

$$F_i(A)=2^m \int_{B_{mi}} F_x\{A+M(x)\}dx ,$$

where $A+M(x)=\{y+M(x);\ y \in A\}$. From the condition (1) it is seen that F_i is well defined as a probability measure on $(R^1, \mathcal{B}^1)$.

ii) For $\boldsymbol{l}=1,\cdots,k(m)$ and $\omega \in \Omega$, $t_i(\boldsymbol{l}, \omega)$ is the $\boldsymbol{l}$th smallest index j that $x_j(\omega) \in B_{mi}$, when such index does not exist it is defined as $t_i(\boldsymbol{l}, \omega)=\infty$.

iii) $z_i'(\boldsymbol{l}, \omega)$ is defined at $z_{t_i(\boldsymbol{l},\omega)}(\omega)$ if $t_i(\boldsymbol{l}, \omega)\neq\infty$. If $t_i(\boldsymbol{l}, \omega)=\infty$, $z_i'(\boldsymbol{l}, \omega)$ is the random selection of numbers distributed as F_i.

Then, since $x_{t_i(\boldsymbol{l},\omega)}(\omega)$ has the uniform distribution on B_{mi} if $t_i(\boldsymbol{l}, \omega)\neq\infty$, the conditional distribution of $z_i'(\boldsymbol{l}, \omega)$ given $t_i(\boldsymbol{l}, \omega)$ is mutually independent and has respectively the distribution F_i for $i=1,\cdots,2^m$ and $\boldsymbol{l}=1,\cdots,k(m)$. Hence $z_i'(\boldsymbol{l}, \omega)$, $i=1,\cdots,2^m$, $\boldsymbol{l}=1,\cdots,k(m)$, are mutually independent and have respectively the distribution F_i.

Since $\int ydF_x(y)=M(x)$,

$$\begin{aligned}\int ydF_i(y)&=\int_{B_{mi}} 2^m\left\{\int yF_x(dy+M(x))\right\}dx\\ &=\int_{B_{mi}} 2^m \int (y-M(x))dF_x(y)dx\\ &=0\end{aligned}$$

and

$$\int y^{2r}dF_i(y)=\int_{B_{mi}} 2^m \int \{y-M(x)\}^{2r}dF_x(y)dx\leqq V .$$

Then, from the fact that the total number of terms in expansion of $(\alpha_1+\cdots+\alpha_k)^{2r}$ which contain no first order factors is less than $\alpha(r)k^r$, where $\alpha(r)$ is not depend on k, it follows that

$$\mathrm{E}\left\{\sum_{\boldsymbol{l}=1}^{k(m)} z_i'(\boldsymbol{l}, \omega)\right\}^{2r}\leqq\alpha(r)k^r(m)V .$$

Hence, from the definition of $z_i'(\boldsymbol{l}, \omega)$, the assumptions (7) and (9), it follows

$$\begin{aligned}(10)\qquad P_\infty&\leqq\lim_{K\to\infty}\sum_{m=0}^{\infty}\mathrm{P}\left\{\omega;\ \left|\sum_{\boldsymbol{l}=1}^{k(m)} z_i'(\boldsymbol{l}, \omega)\right|>Kk(m)^{1-1/b} \text{ for some } i=1,\cdots,2^m\right\}\\ &\leqq\lim_{K\to\infty}\sum_{m=0}^{\infty}\sum_{i=1}^{2^m}\mathrm{P}\left\{\omega;\ \left|\sum_{\boldsymbol{l}=1}^{k(m)} z_i'(\boldsymbol{l}, \omega)\right|>Kk(m)^{1-1/b}\right\}\\ &\leqq\lim_{K\to\infty}\sum_{m=0}^{\infty}\sum_{i=1}^{2^m}\frac{\mathrm{E}\left\{\sum_{\boldsymbol{l}=1}^{k(m)} z_i'(\boldsymbol{l}, \omega)\right\}^{2r}}{K^{2r}k(m)^{2r-2r/b}}\end{aligned}$$

$$\leqq \lim_{K\to\infty} \frac{1}{K^{2r}} \sum_{m=0}^{\infty} \frac{2^m}{k(m)^{r-2r/b}} \alpha(r) V$$
$$= 0 .$$

This lemma can be applied to Theorem 1 [1], i.e., if a condition

$$\int \{y - M(x)\}^4 dF_x(y) \leqq V < \infty$$

is added, $k(m) = 2^m$ may be chosen and it may be thought reasonable.

3. Estimation of maximum point

LEMMA 2. *If the conditions of Lemma* 1 *are satisfied and further if*

$$M(x) < M < \infty \qquad \text{for } x \in [0, 1)$$

and an increasing sequence $h(m)$ satisfies the condition

$$\sum_{m=0}^{\infty} \frac{1}{h(m)} = \infty \tag{11}$$

then, for every $x \in [0, 1)$, either $N_n(x, \omega)$ or $N_n'(x, \omega)$ diverges to the infinity a.e. (P′), *where* $N_n'(x, \omega) = \lim_{\varepsilon\to+0} N(x-\varepsilon, \omega)$.

PROOF. Let x^o be an element of $[0, 1)$ and fix it. Let $E = [i/2^k, i'/2^{k'})$ be a neighbourhood of x^o and define the set

$$F_m = \{\omega;\ \text{Both } N_n(x, \omega) \text{ and } N_n'(x, \omega) \text{ are constants with } n \text{ where } n \geqq m \text{ on } x \in E\} .$$

Let's prove $P'(F_m) = 0$ for $m = 1, 2, \cdots$. Let $B_i = \{\omega;\ x_i(\omega) \in E\}$ for $i = 1, 2, \cdots$. Then it follows from the extended Borel's 0–1 law (see, e.g. [3], p. 398) that

(12) both of events $\left\{\omega;\ \sum_{n=1}^{\infty} I_{B_n}(\omega) < \infty\right\}$ and $\left\{\omega;\ \sum_{n=1}^{\infty} P^{\mathcal{B}_{n-1}} B_n(\omega) < \infty\right\}$ are equivalent, where $\mathcal{B}_n$ is the minimal sub-σ-field generated by $B_1, \cdots, B_n$ and I_{B_n} is an indicator function of B_n.

Fix an $\omega \in F_m$ such that $|M_n'(x, \omega)|^b N_n(x, \omega)$ is bounded by $c = c(\omega) < \infty$. Then, for any $x \in [0, 1)$,

$$M_n(x, \omega) = M_n'(x, \omega) + \frac{1}{N_n(x, \omega)} \sum_{x_j(\omega) \in C_n(x, \omega)} M(x_j(\omega)) \tag{13}$$
$$\leqq \left\{\frac{c}{N_n(x, \omega)}\right\}^{1/b} + M \leqq c^{1/b} + M .$$

It follows from (13) that for any $x \in E$ and $n \geqq m$

$$\alpha'_n(x, \omega) \geqq h(N_m(x, \omega)) \exp -N_m(x, \omega)\{c^{1/b}+M-M_m(x, \omega)\}^b = c'(\omega, m, E) = c' > 0 . \tag{14}$$

From (11), (14) and the fact

$$\alpha'_n(x, \omega) \leqq h(N_n(x, \omega)) \leqq h(n)$$

it follows that

$$\sum_{n=1}^{\infty} P'^{\mathscr{B}_{n-1}} B_n = \sum_{n=1}^{\infty} \mathrm{E}\left\{\int_E \beta'_{n-1}(x, \omega)dx\right\} = \sum_{n=1}^{\infty} \mathrm{E}\left\{\frac{\int_E \alpha'_n(x, \omega)dx}{\int_0^1 \alpha'_n(x, \omega)dx}\right\}$$

$$\geqq \sum_{n=m}^{\infty} \frac{\int_E c'dx}{h(n)} = \infty .$$

Then $\omega \in F_m$ implies $\sum_{n=1}^{\infty} I_{B_n}(\omega) = \infty$ a.e. from (12). On the other hand from the definitions of F_m and $N_n(x, \omega)$, if ω belongs to F_m, then $\{n ; x_n(\omega) \in E\}$ is finite. Then $\mathrm{P}'(F_m)=0$ and $\mathrm{P}'\left(\bigcup_{m=1}^{\infty} F_m\right)=0$. Since k, k', i and i' are arbitrary, either $N_n(x, \omega)$ or $N'_n(x, \omega)$ diverges to the infinity a.e..

THEOREM. *If the conditions* (6), (7) *and* (11) *are satisfied, and further if*

(15) $M(x_0) = \max_{0 \leqq x < 1} M(x)$ *and* $\sup_{x \notin G} M(x) < M(x_0)$ *for every neighbourhood* G *of* x_0,

(16) $M(x)$ *is continuous at* x_0,

(17) *for some positive constant* c, $(1/2^m)h(k(m)) > c$ $m = 1, 2, \cdots$

and

(18) $\{h(m)e^{-\alpha m} ; m = 1, 2, \cdots\}$ *is bounded for every* $\alpha > 0$

then, for every open interval E *with* x_0,

$$\lim_{n\to\infty} \int_E \beta'_n(x, \omega)dx = 1 \qquad \textit{a.e.} \ (\mathrm{P}') \tag{19}$$

and

$$\lim_{n\to\infty} \bar{x}_n(\omega) = x_0 \qquad \textit{a.e.} \ (\mathrm{P}') , \tag{20}$$

where $\max_{x \in [0,1)} M_n(x, \omega) = M_n(\bar{x}_n(\omega), \omega)$.

PROOF. Since, from Lemma 2,

$P'\{\omega;\ N_n(r,\omega)$ or $N_n'(r,\omega)\to\infty$ as $n\to\infty$ for every rational number $r\in[0,1)\}=1$

it follows from the definition of $N_n(x,\omega)$ that $N_n(x,\omega)\to\infty$ uniformly as $n\to\infty$ a.e.. Then

i) for any open neighbourhood E of x_0, there exist real numbers $\varepsilon>0$ and $N=N(\varepsilon,\omega)$ such that

$$M(x_0)-\max_{x\in A_n} M(x)>\varepsilon \qquad \text{for any } n\geqq N, \tag{21}$$

where $A_n=\bigcup_{x\notin E} B_n(x,\omega)$

ii) for any positive number ε, there exists an integer $N'=N'(\varepsilon,\omega)$ such that $n\geqq N'$ and $x\in B_n(x_0,\omega)$ imply $|M(x_0)-M(x)|<\varepsilon$

and

iii) for any real number $\varepsilon>0$, there exists $N''=N''(\varepsilon,\omega)$ such that

$$|M_n'(x,\omega)|\leqq\varepsilon \qquad \text{for all } n\geqq N''.$$

In fact, i) is shown from the uniform convergence to 0 of the width of $B_n(x,\omega)$ a.e., which is the consequence of uniform divergence of $N_n(x,\omega)$; ii) is easily proved from (16); iii) is shown in consequence of Lemma 1 and Lemma 2.

Let $N_0=N_0(\varepsilon,\omega)=\max\{N(\varepsilon,\omega),\ N'(\varepsilon/4,\omega),\ N''(\varepsilon/8,\omega)\}$. Then, from i),

$$\begin{aligned} M_n(x,\omega)&=M_n'(x,\omega)+\frac{1}{N_n(x,\omega)}\sum_{x_j(\omega)\in C_n(x,\omega)} M(x_j(\omega))\\ &\leqq\varepsilon/8+M(x_0)-\varepsilon \qquad \text{for } x\notin E \text{ and } n\geqq N_0 \end{aligned} \tag{22}$$

and, from ii) and iii),

$$\begin{aligned} \max_{x\in[0,1)} M_n(x,\omega)&\geqq M_n(x_0,\omega)\\ &=M_n'(x_0,\omega)+\frac{1}{N_n(x_0,\omega)}\sum_{x_j(\omega)\in C_n(x_0,\omega)} M(x_j(\omega))\\ &\geqq-\varepsilon/4+M(x_0)-\varepsilon/8=-3\varepsilon/8+M(x_0) \qquad \text{for } n\geqq N_0. \end{aligned} \tag{23}$$

Now, from (22) and (23), we obtain

$$\max_{x'\in[0,1)} M_n(x',\omega)-M_n(x,\omega)\geqq\varepsilon/2 \qquad \text{for } x\notin E \text{ and } n\geqq N_0 \tag{24}$$

and, for the complement E^c of E,

$$\int_{E^c}\beta_n'(x,\omega)dx\leqq\frac{\int_{E^c} h\{N_n(x,\omega)\}\exp -N_n(x,\omega)(\varepsilon/2)^b dx}{\int_{H_n}\alpha_n'(x,\omega)dx}$$

$$\leqq \frac{\int_{E^c} h\{N_n(x, \omega)\} \exp -N_n(x, \omega)(\varepsilon/2)^b dx}{h\{k(m)\}/2^{m+1}},$$

where H_n is a set $B_n(x, \omega)$ such that $\max_{x' \in [0,1)} M_n(x', \omega) = M_n(x'', \omega)$ for some $x'' \in B_n(x, \omega)$ and $N_n(x'', \omega) = k(m)$.

Then, from (17) and (18),

$$\int_E \beta_n'(x, \omega) dx \to 1 \qquad \text{as } n \to \infty \text{ a.s. (P')}.$$

The second part of the theorem can be proved from i) and ii) in the proof.

Remark 1. In the proof of the theorem, the property of divergence of $N_n(x, \omega)$ to the infinity at every $x \in [0, 1)$ is essential; This property is a necessary condition if $M(x)$ is not unimodal.

2. Under the same conditions of the theorem, it is easily seen that maximum points of $\beta_n'(x, \omega)$ converge to x_0 a.e.. For the proof the property of the remark 1 and the condition (11) are used. Thus when going back to the case of estimation of the root of $M(x)=a$, the condition (11) permits that maximum points of both $M_n(x, \omega)$ and $\beta_n(x, \omega)$ converge to x_0 a.e.. Though this characteristic is favourable, the divergence of $N_n(x, \omega)$ at any $x \in [0, 1)$ implies that the distribution of $\beta_n(x, \omega)$ converges loosely to x_0.

3. There exist sequences $k(m)$ and $h(m)$ which satisfy the conditions in Theorem. We may take, for instance, as follows.

$$r=1,\ K(m)=\text{the integer part of } (2^m m^2)^{b/b-2},\ b>2,\ h(m)=m$$

or

$$r=2, 3, \cdots,\ k(m)=2^m,\ b>2,\ (b-2)/b>1/r,\ h(m)=m.$$

NARA MEDICAL COLLEGE

REFERENCES

[1] Yamazoe, S. (1972). A random observation process for stochastic approximation, *Ann. Inst. Statist. Math.*, **24**, 309-317.

[2] Kiefer, J. and Wolfowitz, J. (1952). Stochastic approximation of the maximum of a regression function, *Ann. Math. Statist.*, **23**, 462-466.

[3] Loève, M. (1963). *Probability Theory*, Third edition, Van Nostrand, Princeton.