On a modified Robbins-Monro procedure approximating the root from below with errors in setting the x-levels

By
Eiichi ISOGAI

(Received November 30, 1978)

1. Introduction and Summary

In the case of finding the unique root θ of the equation $M(x)=0$, situations may occure where even the precise setting of the x-levels of an experiment is impossible without error. DUPAČ and KRÁL [3] and WATANABE [6] dealt with these situations. On the other hand, there are cases in which it is advantageous to use a process which converges to θ from below. ANBAR [1] gave a modified Robbins-Monro (RM) procedure for guaranteeing that with probability one the procedure overestimates θ only finitely many times. In this paper, it is shown that assuming that the error in x-level can be made small at some rate at each step, the modified RM procedure overestimates θ only finitely many times with probability one.

This paper consists of five sections. In section 2, we shall give some assumptions, notations and a lemma. In section 3, we shall show a convergence theorem. Section 4 will give some lemmas which are used in section 5. In section 5, we shall present two theorems which show that with probability one the modified RM process overestimates θ only finitely many times and give an asymptotic normality of the process.

2. Preliminaries

Let R be the real line. Let $\{U^n(x)\}$ and $\{V^n(x)\}$ be two sequences of random variables which depend on parameter $x\in R$. Suppose that for each n, $U^n(x)$ and $V^n(x)$ are measurable functions of x. Further, suppose $E[U^n(x)]=E[V^n(x)]=0$ for all $x\in R$ and all $n\geqslant 1$.

Let $M(x)$ be a real-valued measurable function on R, let θ be the unique root of $M(x)=\alpha$ where α is an arbitrary given number.

Let us define the mdified RM procedure proposed by ANBAR [1] as follows: Let X_1 be a random variable with $E[X_1^2]<\infty$ and let define X_2, X_3, $\cdots$ by the recursive relation

$$(2.\ 1)\qquad X_{n+1}=X_n-a_n[M(X_n+u_n)-v_n-\alpha+b_n]\quad n=1, 2, \cdots\cdots$$

where $\{a_n\}$ is a sequence of positive numbers satisfying

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

$\{b_n\}$ is a sequence of numbers satisfying

(2. 3) $$\lim_{n\to\infty} b_n=0,$$

u_n, $n\geq 1$, are random variables whose conditional distributions, given $X_1, u_1, \cdots, u_{n-1}, v_1, \cdots, v_{n-1}$, coincide with those of $U^n(X_n)$, and v_n, $n\geq 1$, are random variables whose conditional distributions, given $X_1, u_1, \cdots, u_n, v_1, \cdots, v_{n-1}$, coincide with those of $V^n(X_n+u_n)$.

The following lemma given by WATANABE [5] will be needed to prove Theorem 3. 1.

LEMMA 2. 1. *Let $\{U_n\}_{n=1}^{\infty}$ and $\{V_n\}_{n=1}^{\infty}$ be two sequences of random variables on a probability space $(\Omega, \mathfrak{A}, P)$. Let $\{\mathfrak{A}_n\}_{n=1}^{\infty}$ be a sequence of sub-σ-algebras of $\mathfrak{A}$, $\mathfrak{A}_n\subset\mathfrak{A}_{n+1}\subset\mathfrak{A}$, where U_n and V_n are measurable with respect to $\mathfrak{A}_n$ for each $n\geq 1$. Furthermore, let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive numbers satisfying*

(2. 4) $$\lim_{n\to\infty} a_n=0, \qquad \sum_{n=1}^{\infty} a_n=\infty.$$

Suppose that the following conditions are satisfied:

(2. 5) $U_n\geq 0$ *a. s. for all* $n\geq 1$,

(2. 6) $E[U_1]<\infty$,

(2. 7) $E[U_{n+1}|\mathfrak{A}_n]\leq(1-a_n)U_n+V_n$ *a. s. for all* $n\geq 1$,

(2. 8) $\sum_{n=1}^{\infty} E[|V_n|]<\infty$.

Then, it holds that $\lim_{n\to\infty} U_n=0$ *a. s. and* $\lim_{n\to\infty} E[U_n]=0$.

3. Convergence of the modified RM process

In this section, an almost surely convergence of the modified RM process is proved.

THEOREM 3. 1. *Let $\{\alpha_n\}_{n=1}^{\infty}$ be a sequence of non-negative numbers and $\{\beta_n\}_{n=1}^{\infty}$ be a sequence of positive numbers. Suppose the following conditions are satisfied:*

(3. 1) $K_1\leq(M(x)-\alpha)/(x-\theta)\leq K_2$ *for all* $x\neq\theta$,

where K_1 and K_2 are some positive constants;

(3. 2) $\sup_{-\infty<x<\infty} \mathrm{Var}\,[U^n(x)]\leq\alpha_n$ *for all* $n\geq 1$;

(3. 3) $$\sup_{-\infty<x<\infty} \mathrm{Var}\,[V^n(x)] \leqslant \beta_n \qquad \textit{for all } n \geqslant 1$$

(3. 4) $$\sum_{n=1}^{\infty} a_n \alpha_n < \infty;$$

(3. 5) $$\sum_{n=1}^{\infty} a_n^2 \beta_n < \infty;$$

(3. 6) $$\sum_{n=1}^{\infty} a_n |b_n| < \infty.$$

Then, the modified RM process X_n defined by (2. 1) *converges to θ with probability one as well as in mean-square.*

PROOF. Without loss of generality we may assume $\alpha=0$. From (2. 1) we have

(3. 7) $$X_{n+1}-\theta=(X_n-\theta)-a_n M(X_n+u_n)+a_n v_n-a_n b_n.$$

Squaring both sides of (3. 7) and taking conditional expectations on both sides given $X_1, \cdots, X_n$, we obtain

(3. 8) $$\begin{aligned} &E[(X_{n+1}-\theta)^2 | X_1, \cdots, X_n] \\ &= (X_n-\theta)^2 + a_n^2 E[M^2(X_n+u_n) | X_1, \cdots, X_n] \\ &\quad + a_n^2 E[v_n^2 | X_1, \cdots, X_n] + a_n^2 b_n^2 - 2a_n (X_n-\theta) E[M(X_n+u_n) \\ &\quad | X_1, \cdots, X_n] - 2a_n^2 b_n E[v_n | X_1, \cdots, X_n] \\ &\quad + 2a_n (X_n-\theta) E[v_n | X_1, \cdots, X_n] - 2a_n b_n (X_n-\theta) \\ &\quad - 2a_n^2 E[M(X_n+u_n) v_n | X_1, \cdots, X_n] \\ &\quad + 2a_n^2 b_n E[M(X_n+u_n) | X_1, \cdots, X_n]. \end{aligned}$$

From the property of $V^n(x)$ and (3. 3), it is easily seen that $E[v_n | X_1, \cdots, X_n]=0$ and $E[v_n^2 | X_1, \cdots, X_n] \leqslant \beta_n$.

Let us define $Q(x, \theta)$ as follows:

$$\begin{aligned} Q(x, \theta) &= M(x)/(x-\theta) \qquad && \textit{if } x \neq \theta \\ &= \alpha_1 && \textit{if } x=\theta \end{aligned}$$

where α_1 is an arbitrary fixed constant with $K_1 \leqslant \alpha_1 \leqslant K_2$. By (3. 1) we get

(3. 9) $$M(x)=Q(x, \theta)(x-\theta)$$

where $K_1 \leqslant Q(x, \theta) \leqslant K_2$ for all x.

Since $|M(X_n+u_n)| \leqslant K_2(|X_n-\theta|+|u_n|)$, it follows by (3. 2) and (3. 9) that

(3. 10) $$\begin{aligned} &E[M^2(X_n+u_n) | X_1, \cdots, X_n] \\ &\leqslant 2K_2^2 (X_n-\theta)^2 + 2K_2^2 E[u_n^2 | X_1, \cdots, X_n] \\ &\leqslant 2K_2^2 (X_n-\theta)^2 + 2K_2^2 \alpha_n. \end{aligned}$$

Using (3. 9), we have

$$(X_n-\theta)E[M(X_n+u_n)\,|X_1, \cdots, X_n]$$
$$=(X_n-\theta)^2E[Q(X_n+u_n, \theta)\,|X_1, \cdots, X_n]$$
$$+(X_n-\theta)E[Q(X_n+u_n, \theta)u_n\,|X_1, \cdots, X_n].$$

The relation (3. 2) and Schwarz's inequality imply

$$|E[Q(X_n+u_n, \theta)u_n\,|X_1, \cdots, X_n]\,|\leqslant K_2\alpha_n^{\frac{1}{2}}\,.$$

Therefore, we have

$$(X_n-\theta)E[M(X_n+u_n)\,|X_1, \cdots, X_n] \tag{3. 11}$$
$$\geqslant K_1(X_n-\theta)^2-K_2\alpha_n^{\frac{1}{2}}|X_n-\theta|.$$

It follows by (3. 2) and (3. 9) that

$$|E[M(X_n+u_n)\,|X_1, \cdots, X_n]\,| \tag{3. 12}$$
$$\leqslant K_2|X_n-\theta|+K_2\alpha_n^{\frac{1}{2}}\,.$$

By making use of $E[v_n|X_1, \cdots, X_n, u_n]=0$ and taking conditional expectations given X_1, $\cdots$, X_n, we obtain

$$E[M(X_n+u_n)v_n\,|X_1, \cdots, X_n]=0. \tag{3. 13}$$

The relations (3. 8), (3. 10), (3. 11) and (3. 12) yield that

$$E[(X_{n+1}-\theta)^2|X_1, \cdots, X_n] \tag{3. 14}$$
$$\leqslant (X_n-\theta)^2+2K_2{}^2a_n^2(X_n-\theta)^2+2K_2{}^2a_n^2\alpha_n+a_n^2\beta_n+a_n^2b_n^2$$
$$-2K_1a_n(X_n-\theta)^2+2K_2a_n\alpha_n^{\frac{1}{2}}|X_n-\theta|+2a_n|b_n(X_n-\theta)|$$
$$+2K_2a_n^2|b_n(X_n-\theta)|+2K_2a_n^2|b_n|\alpha_n^{\frac{1}{2}}\,\cdot$$

By making use of the inequality $2ab\leqslant ka^2+k^{-1}b^2$ for any $k>0$, we get the following inequalities:

$$2K_2a_n\alpha_n^{\frac{1}{2}}|X_n-\theta|\leqslant 2^{-1}K_1a_n(X_n-\theta)^2+2K_1{}^{-1}K_2{}^2a_n\alpha_n, \tag{3. 15}$$
$$2a_n|b_n(X_n-\theta)|\leqslant 2^{-1}K_1a_n(X_n-\theta)^2+2K_1{}^{-1}a_nb_n^2\,,$$
$$2K_2a_n^2|b_n(X_n-\theta)|\leqslant K_2a_n^2(X_n-\theta)^2+K_2a_n^2b_n^2\,,$$
$$2K_2a_n^2|b_n|\alpha_n^{\frac{1}{2}}\leqslant K_2a_n^2\alpha_n+K_2a_n^2b_n^2\,.$$

Hence, it follows from (3. 14) and (3. 15) that

$$E[(X_{n+1}-\theta)^2|X_1, \cdots, X_n] \tag{3. 16}$$

$$\leqslant \{1-(K_1-2K_2{}^2a_n-K_2a_n)a_n\}(X_n-\theta)^2$$
$$+(2K_2+1)a_n^2b_n^2+(2K_2{}^2+K_2)a_n^2\alpha_n+2K_2{}^2K_1{}^{-1}a_n\alpha_n$$
$$+a_n^2\beta_n+2K_1{}^{-1}a_nb_n^2\,.$$

By (2. 2) and (2. 3), there exists a positive integer n_0 such that for all $n\geqslant n_0$

$$2K_2{}^2a_n+K_2a_n\leqslant 2^{-1}K_1,\quad a_n\leqslant 1 \text{ and } |b_n|\leqslant 1, \tag{3. 17}$$

so that

$$a_n^2b_n^2\leqslant a_n|b_n|,\quad a_n^2\alpha_n\leqslant a_n\alpha_n,\quad a_nb_n^2\leqslant a_n|b_n| \tag{3. 18}$$

for all $n\geqslant n_0$.

Thus, by (3. 16), (3. 17) and (3. 18), we have

$$E[(X_{n+1}-\theta)^2|X_1,\cdots,X_n] \tag{3. 19}$$
$$\leqslant (1-2^{-1}K_1a_n)(X_n-\theta)^2+(2K_2+2K_1{}^{-1}+1)a_n|b_n|$$
$$+(2K_2{}^2+K_2+2K_2{}^2K_1{}^{-1})a_n\alpha_n+a_n^2\beta_n$$

for all $n\geqslant n_0$.

By (2. 2), (3. 4), (3. 5), (3. 6) and (3. 19), all conditions of Lemma 2. 1 are satisfied. Therefore, we obtain

$$\lim_{n\to\infty}(X_n-\theta)^2=0 \text{ a. s. which implies } \lim_{n\to\infty}X_n=\theta \text{ a. s.,}$$

and $\lim_{n\to\infty}E[(X_n-\theta)^2]=0$.

This completes the proof.

4. Auxiliary lemmas

In this section, some lemmas which are needed in later sections are presented. Throughout this section and section 5, suppose $V^n(x)=V(x)$ for all x and all $n\geqslant 1$ and $\beta_n=\beta$ for all $n\geqslant 1$. It is assumed without loss of generality that $\alpha=\theta=0$.

LEMMA 4. 1. *Suppose the conditions* (3. 1) *to* (3. 3) *are satisfied. Further suppose the following conditions*:

$$a_n=An^{-1} \text{ with } 2AK_1>1; \tag{4. 1}$$

$$\alpha_n=Ln^{-d} \text{ with some } L\geqslant 0 \text{ and some } d>1; \tag{4. 2}$$

$$b_n^2\leqslant C(\log_2 n)/n \text{ for some constant } C>0 \text{ and all } n\geqslant 3, \tag{4. 3}$$

where $\log_2 n$ *means* $\log(\log n)$.

Then, there exists a positive constant C_1 *such that*

$$E[X_n^2]\leqslant C_1(\log_2 n)/n \quad \text{for all } n\geqslant 3.$$

PROOF. Throughout this proof, $C_2, C_3, \cdots$ denote positive constants. From (2. 1) and the property of $V(x)$, we get

$$E[X_{n+1}^2]=E[X_n^2]+a_n^2E[M^2(X_n+u_n)]+a_n^2 E[v_n^2]+a_n^2 b_n^2 -2a_nE[X_nM(X_n+u_n)]-2a_n b_nE[X_n] +2a_n^2 b_nE[M(X_n+u_n)]. \tag{4. 4}$$

We put $Q(x)=Q(x, 0)$, where $Q(x, 0)$ is the same as (3. 9). Inserting $\theta=0$ into (3. 10), (3. 11) and (3. 12) and taking expectations on both sides of each inequality, we have

$$E[M^2(X_n+u_n)]\leqslant 2K_2^2E[X_n^2]+2K_2^2\alpha_n, \tag{4. 5}$$

$$E[X_nM(X_n+u_n)]\geqslant K_1E[X_n^2]-K_2\alpha_n^{\frac{1}{2}}E[|X_n|],$$

$$|E[M(X_n+u_n)]|\leqslant K_2E[|X_n|]+K_2\alpha_n^{\frac{1}{2}} .$$

The relations (4. 4) and (4. 5) imply

$$E[X_{n+1}^2] \leqslant\{1-(2K_1 a_n-2K_2^2 a_n^2)\}E[X_n^2]+2K_2a_n\alpha_n^{\frac{1}{2}}E[|X_n|] +2a_n|b_n|E[|X_n|]+2K_2a_n^2 |b_n|E[|X_n|]+2K_2^2a_n^2\alpha_n +\beta a_n^2+a_n^2b_n^2+2K_2 a_n^2|b_n|\alpha_n^{\frac{1}{2}} . \tag{4. 6}$$

Choose $\varepsilon_1>0$ such that $2AK_1(1-\varepsilon_1)>1$ because of $2AK_1>1$. Then there exists a positive integer n_1 such that

$$AK_2^2K_1^{-1}n^{-1}<\varepsilon_1 \quad \text{for all } n\geqslant n_1, \quad \text{so that}$$

$$2K_1 a_n-2K_2^2a_n^2>2AK_1(1-\varepsilon_1)n^{-1} \quad \text{for all } n\geqslant n_1. \tag{4. 7}$$

Choose $\varepsilon_2>0$ such that $2AK_1(1-\varepsilon_1-\varepsilon_2)>1$.
Then we get

$$2K_2 a_n \alpha_n^{\frac{1}{2}}E[|X_n|] \leqslant 2K_1\varepsilon_2 a_n E[X_n^2]+K_2^2(2K_1\varepsilon_2)^{-1}a_n\alpha_n =2AK_1\varepsilon_2n^{-1}E[X_n^2]+ALK_2^2(2K_1\varepsilon_2)^{-1}n^{-d-1}. \tag{4. 8}$$

Choose $\varepsilon_3>0$ such that $2AK_1(1-\varepsilon_1-\varepsilon_2-\varepsilon_3)>1$. Then it follows by (4. 3) that

$$2a_n|b_n|E[|X_n|] \leqslant 2K_1\varepsilon_3 a_n E[X_n^2]+(2K_1\varepsilon_3)^{-1}a_n b_n^2 \leqslant 2AK_1\varepsilon_3n^{-1}E[X_n^2]+AC(2K_1\varepsilon_3)^{-1}(n^{-1}\log_2 n)n^{-1}. \tag{4. 9}$$

Since $2AK_1(1-\varepsilon_1-\varepsilon_2-\varepsilon_3)>1$, there exists $\varepsilon_4>0$ such that $2AK_1(1-\varepsilon_1-\varepsilon_2-\varepsilon_3-\varepsilon_4)>1$ and $2AK_1(1-\varepsilon_1-\varepsilon_2-\varepsilon_3-\varepsilon_4)\neq d$. By (4. 3) we have

$$2K_2 a_n^2 |b_n| E[|X_n|]$$

$$\leqslant K_2 a_n^2 E[X_n^2] + K_2 a_n^2 b_n^2$$

$$\leqslant A^2 K_2 n^{-2} E[X_n^2] + CA^2 K_2 n^{-2}(n^{-1} \log_2 n).$$

Since there exists a positive integer $n_2 \geqslant n_1$ such that for all $n \geqslant n_2$ $A^2 K_2 n^{-1} < 2AK_1 \varepsilon_4$, we obtain

(4. 10) $$2K_2 a_n^2 |b_n| E[|X_n|]$$

$$\leqslant 2AK_1 \varepsilon_4 n^{-1} E[X_n^2] + CA^2 K_2 n^{-2}(n^{-1} \log_2 n) \qquad \text{for all } n \geqslant n_2.$$

(4. 2) and (4. 3) yield

(4. 11) $$2K_2 a_n^2 |b_n| \alpha_n^{\frac{1}{2}} .$$

$$\leqslant K_2 a_n^2 \alpha_n + K_2 a_n^2 b_n^2$$

$$\leqslant K_2 A^2 L n^{-2-d} + CA^2 K_2 n^{-2}(n^{-1} \log_2 n).$$

From (4. 6), (4. 7), (4. 8), (4. 9), (4. 10) and (4. 11), we obtain

$$E[X_{n+1}^2]$$

$$\leqslant (1 - tn^{-1}) E[X_n^2] + C_2 n^{-2} + C_3 n^{-1-d}$$

$$+ C_4 n^{-1}(n^{-1} \log_2 n) \qquad \text{for all } n \geqslant n_2, \qquad \text{where}$$

$$t \equiv 2AK_1(1 - \varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \varepsilon_4) > 1 \text{ and } t \neq d.$$

Repeating this inequality, we have

(4. 12) $$E[X_{n+1}^2]$$

$$\leqslant \beta_{n_2-1n} E[X_{n_2}{}^2] + C_2 \sum_{m=n_2}^{n} \beta_{mn} m^{-2} + C_3 \sum_{m=n_2}^{n} \beta_{mn} m^{-1-d}$$

$$+ C_4 \sum_{m=n_2}^{n} \beta_{mn} m^{-1}(m^{-1} \log_2 m), \quad \text{where}$$

$$\beta_{mn} = \prod_{j=m+1}^{n} (1 - tj^{-1}) \qquad \text{if } m < n$$

$$= 1 \qquad \text{if } m = n.$$

From $E[X_1{}^2] < \infty$ and (4. 6), it follows by induction that $E[X_n^2] < \infty$ for all $n \geqslant 1$. Since $|\beta_{mn}| \leqslant \gamma n^{-t} m^t$ for some $r > 0$ and all $n \geqslant m$, we get

(4. 13) $$\beta_{n_2-1n} E[X_{n_2}{}^2] \leqslant C_5 n^{-t}.$$

After easy calculations, we have

$$(4.14)\qquad C_2\Big|\sum_{m=n_2}^{n}\beta_{mn}m^{-2}\Big|\leqslant C_6\,n^{-1},$$

$$C_3\Big|\sum_{m=n_2}^{n}\beta_{mn}m^{-1-d}\Big|\leqslant C_7\,n^{-t_0},$$

$$C_4\Big|\sum_{m=n_2}^{n}\beta_{mn}\,m^{-1}(m^{-1}\log_2 m)\Big|\leqslant C_8\,n^{-1}\log_2 n,$$

where $t_0\equiv\min(t, d)>1$.

Inserting (4. 13) and (4. 14) into (4. 12), we obtain

$$E[X_{n+1}^2]\leqslant C_5\,n^{-t}+C_6\,n^{-1}+C_7\,n^{-t_0}+C_8\,n^{-1}\log_2 n$$

for all $n\geqslant n_2$. Taking into account this inequality, $t>1$ and $t_0>1$, Lemma 4. 1 follows.

Thus the proof is completed.

LEMMA 4. 2. *Let $p>1/2$ be a fixed number. Then under the conditions of Lemma* 4. 1,

$$n^{-p+\frac{1}{2}}\sum_{m=1}^{n}m^{p-1}X_m^2\longrightarrow 0\quad a.s.\ \ as\ n\to\infty,$$

$$n^{-p+\frac{1}{2}}\sum_{m=1}^{n}m^{p-1}|u_m|\longrightarrow 0\quad a.s.\ \ as\ n\to\infty.$$

and

$$n^{-p+\frac{1}{2}}\sum_{m=1}^{n}m^{p-1}u_m^2\longrightarrow 0\quad a.s.\ as\ \ n\to\infty.$$

PROOF. First, we shall prove the first assertion. It holds

$$n^{-p+\frac{1}{2}}\sum_{m=1}^{n}m^{p-1}X_m^2=n^{-p+\frac{1}{2}}\sum_{m=1}^{n}m^{p-\frac{1}{2}}\ (X_m^2/m^{\frac{1}{2}}).$$

If $\sum_{n=1}^{\infty}(X_n^2/n^{\frac{1}{2}})<\infty$ a. s., using Kronecker's lemma, we have

$$n^{-p+\frac{1}{2}}\sum_{m=1}^{n}m^{p-1}X_m^2\longrightarrow 0\quad \text{a. s. as}\ \ n\to\infty.$$

Hence, it suffices to prove $\sum_{n=1}^{\infty}(E[X_n^2]/n^{\frac{1}{2}})<\infty$.

According to Lemma 4. 1,

$$\sum_{n=3}^{\infty}(E[X_n^2]/n^{\frac{1}{2}})\leqslant C_1\sum_{n=3}^{\infty}(\log_2 n)n^{-\frac{3}{2}}<\infty.$$

Thus the first assertion is proved.

To prove the second and the third assertions, it is sufficient to prove $\sum_{n=1}^{\infty}(E[|u_n|]/n^{\frac{1}{2}})<\infty$ and $\sum_{n=1}^{\infty}(E[u_n^2]/n^{\frac{1}{2}})<\infty$ respectively. Using Schwarz's inequality and (4. 2,) we get

$$\sum_{n=1}^{\infty} (E[|u_n|]/n^{\frac{1}{2}}) \leqslant L^{\frac{1}{2}} \sum_{n=1}^{\infty} n^{-(d+1)/2} < \infty$$

and

$$\sum_{n=1}^{\infty} (E[u_n^2]/n^{\frac{1}{2}}) \leqslant L \sum_{n=1}^{\infty} n^{-\frac{1}{2}-d} < \infty.$$

This completes the proof.

LEMMA 4. 3. *Let* $\bar{\delta}(x)$ *be a measurable function such that* $\lim_{x\to 0} \bar{\delta}(x)/x^2=0$. *Then under the conditions of Lemma* 4. 2,

$$n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1} \bar{\delta}(X_m+u_m) \longrightarrow 0 \quad a.\ s.\ \ as\ n\to\infty.$$

PROOF. By Theorem 3. 1, it follows that $\lim_{m\to\infty} X_m=0$ a. s.
Let $\varepsilon>0$ be arbitrary. By Chebyshev's inequality and (4. 2), we get $\sum_{m=1}^{\infty} P(|u_m|>\varepsilon)<\infty$. Hence, By Borel-Cantelli lemma, we have $\lim_{m\to\infty} u_m=0$ a.s. Thus it follows that

$$\lim_{m\to\infty} (X_m+u_m)=0 \quad \text{a. s.}$$

From the property of the function $\bar{\delta}(x)$,

$$\bar{\delta}(X_m+u_m)/(X_m+u_m)^2=o(1) \quad \text{a. s.} \quad \text{as } m\to\infty.$$

To prove the lemma, it suffices to prove that

$$n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1}(X_m+u_m)^2 \longrightarrow 0 \quad \text{a. s. as } n\to\infty,$$

for

$$n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1} \bar{\delta}(X_m+u_m)$$

$$=n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1}(X_m+u_m)^2\cdot O(1) \quad \text{a. s.}$$

According to Lemma 4. 2,

$$0\leqslant n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1} (X_m+u_m)^2$$

$$\leqslant 2[n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1} X_m^2 + n^{-p+\frac{1}{2}} \sum_{m=1}^{n} m^{p-1} u_m^2]$$

$$\longrightarrow 0 \quad \text{a. s. as } n\to\infty,$$

which concludes the proof.

5. Main results

In this section, the results of the previous sections are used to show that the modified RM process, due to ANBAR [1], converges to θ a. s. from below.

Assume the following:

(5. 1) $$M(x)=\alpha+\alpha_1(x-\theta)+\delta(x,\theta) \quad \text{where}$$
$$\delta(x,\theta)=\alpha_2(x-\theta)^2+\bar{\delta}(x-\theta),$$
$$\bar{\delta}(x)=o(x^2) \quad \text{as } x\to 0,$$
$\alpha_1>0$, α_2 is finite and $\bar{\delta}(x)$ is a measurable function;

(5. 2) $$\sup_{-\infty<x<\infty} E\{|V(x)|^{2+\eta}\}<\infty \quad \text{for some } \eta>0;$$

(5. 3) $$\lim_{x\to\theta} E\{V^2(x)\}=\sigma^2.$$

Consider the modified RM procedure defined by

(5. 4) $$X_{n+1}=X_n-An^{-1}\{M(X_n+u_n)-v_n-\alpha+b_n\} \quad n\geqslant 1$$

where X_1 is a random variable with $E[X_1^2]<\infty$.

Let D_n, $n\geqslant 1$, be a sequence of real numbers satisfying

(5. 5) $$D_n\geqslant Dn^{-\frac{1}{2}}(2\log_2 n)^{\frac{1}{2}}+o(n^{-\frac{1}{2}}(\log_2 n)^{\frac{1}{2}})$$

for all $n\geqslant$ some n_0 and arbitrary positive constant D where

$$D_n=\sum_{m=1}^{n} m^{-1}\beta_{mn}b_m,$$

$$\beta_{mn}=\prod_{j=m+1}^{n}(1-A\alpha_1 j^{-1}) \quad \text{if } m<n$$
$$=1 \quad \text{if } m=n$$

and

$$D>\sigma(2A\alpha_1-1)^{-\frac{1}{2}}.$$

THEROEM 5. 1. *Let X_1, X_2, $\cdots$ be a modified RM process given by* (5. 4). *If the conditions of Theorem* 3. 1 *together with* (4. 1), (4. 2), (4. 3), (5. 1), (5. 2), (5. 3) *and* (5. 5) *hold with* $2AK_1>1$, *then* $\lim_{n\to\infty} X_n=\theta$ *a. s.,* $\lim_{n\to\infty} E[(X_n-\theta)^2]=0$ *and with probability one* $X_n>\theta$ *only finitely many times.*

PROOF. Without loss of generality, we may assume that $\alpha=\theta=0$. Throughout this proof, C_1, C_2, $\cdots$ denote positive constants. By (5. 1) and (5. 4), we have

$$X_{n+1}=(1-A\alpha_1 n^{-1})X_n-A\alpha_1 n^{-1}u_n-An^{-1}\delta_n(X_n)$$
$$+An^{-1}v_n-An^{-1}b_n$$

where

$$\delta_n(X_n) \equiv \delta(X_n + u_n, 0).$$

Repeating this equality, we obtain

$$X_{n+1} = \beta_{0n} X_1 - A\alpha_1 \sum_{m=1}^{n} m^{-1} \beta_{mn} u_m - A \sum_{m=1}^{n} m^{-1} \beta_{mn} \delta_m(X_m) + A \sum_{m=1}^{n} m^{-1} \beta_{mn} v_m - A \sum_{m=1}^{n} m^{-1} \beta_{mn} b_m. \tag{5. 6}$$

Let $\mathfrak{A}_{n-1}$ denote a σ-algebra generated by $X_1, u_1, \cdots, u_n, v_1, \cdots, v_{n-1}$ for each n. Clearly $E\{v_n | \mathfrak{A}_{n-1}\} = 0$ so that v_n's are martingale differences. By (3. 1), (5. 1) and $2AK_1 > 1$, we get

$$2A\alpha_1 > 1. \tag{5. 7}$$

Since $n^{\frac{1}{2}} |\beta_{0n}| \leqslant C_2 n^{\frac{1}{2} - A\alpha_1}$, it follows by (5. 7)

$$\beta_{0n} X_1 = o(n^{-\frac{1}{2}}) \quad \text{a. s. as } n \to \infty. \tag{5. 8}$$

Since $n^{\frac{1}{2}} | \sum_{m=1}^{n} m^{-1} \beta_{mn} u_m | \leqslant C_3 n^{-A\alpha_1 + \frac{1}{2}} \sum_{m=1}^{n} m^{A\alpha_1 - 1} |u_m|$, according to Lemma 4. 2 with $p = A\alpha_1$,

$$\sum_{m=1}^{n} m^{-1} \beta_{mn} u_m = o(n^{-\frac{1}{2}}) \text{ a. s. } \quad \text{as } n \to \infty. \tag{5. 9}$$

The relation (5. 1) implies

$$\sum_{m=1}^{n} m^{-1} \beta_{mn} \delta_m(X_m) = \alpha_2 \sum_{m=1}^{n} m^{-1} \beta_{mn} (X_m + u_m)^2 + \sum_{m=1}^{n} m^{-1} \beta_{mn} \overline{\delta}(X_m + u_m). \tag{5. 10}$$

By making use of the inequality $(a+b)^2 \leqslant 2(a^2 + b^2)$, we have

$$n^{\frac{1}{2}} \left| \sum_{m=1}^{n} m^{-1} \beta_{mn} (X_m + u_m)^2 \right| \leqslant 2C_4 \{ n^{-A\alpha_1 + \frac{1}{2}} \sum_{m=1}^{n} m^{A\alpha_1 - 1} X_m^2 + n^{-A\alpha_1 + \frac{1}{2}} \sum_{m=1}^{n} m^{A\alpha_1 - 1} u_m^2 \}.$$

Taking into account Lemma 4. 2 and this inequality, we obtain

$$\sum_{m=1}^{n} m^{-1} \beta_{mn} (X_m + u_m)^2 = o(n^{-\frac{1}{2}}) \quad \text{a. s. as } n \longrightarrow \infty \tag{5. 11}$$

Since

$$n^{\frac{1}{2}} \left| \sum_{m=1}^{n} m^{-1} \beta_{mn} \overline{\delta}(X_m + u_m) \right|$$

$$\leqslant C_5 n^{-A\alpha_1+\frac{1}{2}} \sum_{m=1}^{n} m^{A\alpha_1-1} \bar{\delta}(X_m+u_m),$$

it follows, according to Lemma 4. 3, that

$$(5.12) \qquad \sum_{m=1}^{n} m^{-1} \beta_{mn} \bar{\delta}(X_m+u_m)=o(n^{-\frac{1}{2}}) \quad \text{a. s.} \quad \text{as } n \longrightarrow \infty.$$

By (5. 10), (5. 11) and (5. 12), we obtain

$$(5.13) \qquad \sum_{m=1}^{n} m^{-1} \beta_{mn} \delta_m(X_m)=o(n^{-\frac{1}{2}}) \quad \text{a. s.} \quad \text{as } n \longrightarrow \infty.$$

From (5. 6), (5. 8), (5. 9) and (5. 13), we get

$$P\{X_{n+1}>0 \quad \text{i. o.}\}$$

$$=P\{A\sum_{m=1}^{n} m^{-1} \beta_{mn} v_m > AD_n + o(n^{-\frac{1}{2}}) \quad \text{i. o.}\}.$$

In the same way as HEYDE [4], we can show that

$$\limsup_{n\to\infty} \{n^{\frac{1}{2}} (2\log_2 n)^{-\frac{1}{2}} \sum_{m=1}^{n} m^{-1} \beta_{mn} v_m\} = \sigma(2A\alpha_1-1)^{-\frac{1}{2}} \quad \text{a. s.}$$

Hence,

$$P\{X_{n+1}>0 \quad \text{i. o.}\}$$

$$\leqslant P\{A \limsup_{n\to\infty} n^{\frac{1}{2}} (2\log_2 n)^{-\frac{1}{2}} \sum_{m=1}^{n} m^{-1} \beta_{mn} v_m$$

$$\geqslant A \limsup_{n\to\infty} n^{\frac{1}{2}} (2\log_2 n)^{-\frac{1}{2}} D_n\}$$

$$=P\{A\sigma(2A\alpha_1-1)^{-\frac{1}{2}} \geqslant AD\}$$

$$=P\{\sigma(2A\alpha_1-1)^{-\frac{1}{2}} > \sigma(2A\alpha_1-1)^{-\frac{1}{2}}\}=0.$$

Therefore, with probability one $X_n>0$ only finitely many times. Also, from Theorem 3. 1,

$$X_n \longrightarrow 0 \quad \text{a. s.} \quad \text{as } n \longrightarrow \infty$$

and

$$E[X_n^2] \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

Thus, the proof is completed.

EXAMPLE OF $\{d_n\}$

The following example is a special one given by ANBAR [1];

$$b_1=b_2=0$$

$$b_n=D'n^{-\frac{1}{2}} (2\log_2 n)^{\frac{1}{2}} \qquad n\geqslant 3$$

with $D' > 2^{-1}\sigma(2A\alpha_1-1)^{\frac{1}{2}}$.

This sequence $\{b_n\}$ satisfies (5. 5).

The following theorem presents the asymptotic normality of the process (5. 4)

THEOREM 5. 2. *Under the conditions of Theorem* 5. 1,

$n^{\frac{1}{2}}(X_n-\theta+AD_n)$ *converges in law to a normal variable with mean zero and variance* $A^2\sigma^2(2A\alpha_1-1)^{-1}$.

PROOF. Throughout this proof, $C_1, C_2, \cdots$ denote positive constants. We may assume $\alpha=0$. From (5. 6) we get

$$(5.14)\qquad (n+1)^{\frac{1}{2}}(X_{n+1}-\theta+AD_{n+1})$$
$$=(n+1)^{\frac{1}{2}}\beta_{0n}(X_1-\theta)-A\alpha_1(n+1)^{\frac{1}{2}}\sum_{m=1}^{n}m^{-1}\beta_{mn}u_m$$
$$-A(n+1)^{\frac{1}{2}}\sum_{m=1}^{n}m^{-1}\beta_{mn}\delta_m(X_m)$$
$$+A(n+1)^{\frac{1}{2}}\sum_{m=1}^{n}m^{-1}\beta_{mn}v_m-A^2\alpha_1(n+1)^{-\frac{1}{2}}D_n$$
$$+A(n+1)^{-\frac{1}{2}}b_{n+1}\qquad \text{a. s.}$$

It is easily seen that as $n\to\infty$

$$(n+1)^{\frac{1}{2}}\beta_{0n}(X_1-\theta)=o(1)\qquad \text{a. s.,}$$
$$(n+1)^{\frac{1}{2}}\sum_{m=1}^{n}m^{-1}\beta_{mn}u_m=o(1)\qquad \text{a. s.,}$$
$$(n+1)^{\frac{1}{2}}\sum_{m=1}^{n}m^{-1}\beta_{mn}\delta_m(X_m)=o(1)\qquad \text{a. s.,}$$
$$(n+1)^{-\frac{1}{2}}D_n=o(1)\qquad \text{and}\qquad (n+1)^{-\frac{1}{2}}b_{n+1}=o(1).$$

Thus, from (5. 14)

$$(5.15)\qquad (n+1)^{\frac{1}{2}}(X_{n+1}-\theta+AD_{n+1})$$
$$=o(1)+A(n+1)^{\frac{1}{2}}\sum_{m=1}^{n}m^{-1}\beta_{mn}v_m\qquad \text{a. s. as } n\to\infty.$$

Choose a positive integer m_0 such that $1-A\alpha_1 m_0^{-1}>0$.

Putting $\gamma_n=\prod_{j=m_0}^{n}(1-A\alpha_1 j^{-1})$, we have

$$\beta_{mn}=\gamma_n\gamma_m^{-1}\qquad \text{for all } n\geqq m\geqq m_0.$$

Since from (5. 2) it holds that $|v_m|<\infty$ a. s. for all m, we get

$$n^{\frac{1}{2}} \sum_{m=1}^{m_0-1} m^{-1}|\beta_{mn}||v_m|$$

$$\leqslant C_1 n^{-A\alpha_1+\frac{1}{2}} \sum_{m=1}^{m_0-1} m^{-1+A\alpha_1}|v_m| \longrightarrow 0 \quad \text{a. s. as } n\to\infty.$$

Thus,

$$n^{\frac{1}{2}} \sum_{m=1}^{m_0-1} m^{-1}\beta_{mn} v_m = o(1) \qquad \text{a. s. as } n\to\infty. \tag{5.16}$$

Setting $U_m = m^{-1}\gamma_m^{-1} v_m$, we have

$$\sum_{m=m_0}^{n} m^{-1}\beta_{mn} v_m = \gamma_n \sum_{m=m_0}^{n} U_m.$$

Let $\mathfrak{A}_{n-1}$ be the same as defined in Theorem 5. 1. Then, $\{U_n, \mathfrak{A}_n; n\geqslant m_0\}$ is a martigale difference. Put $S_n = \sum_{m=m_0}^{n} U_m$. Since as in HEYDE [4]

$$s_n^2 \equiv E[S_n^2] = \sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2} E[v_m^2]$$

$$\sim \sigma^2 \gamma_n^{-2}(2A\alpha_1-1)^{-1} n^{-1} \quad \text{as } n\to\infty$$

and

$$\gamma_n^{-2} n^{-1} \geqslant C_2 n^{2A\alpha_1-1} \nearrow \infty \qquad \text{as } n\to\infty,$$

we have $s_n^2 \nearrow \infty$ as $n\to\infty$. To prove this theorem, we shall use Theorem 2 in BROWN [2]. Firstly, we shall check the Lindeberg condition, i. e.

(5. 17) $\quad s_n^{-2} \sum_{j=m_0}^{n} E[U_j^2 I(|U_j| \geqslant \varepsilon s_n)] \longrightarrow 0$ as $n\to\infty$ for all $\varepsilon>0$, where $I(A)$ denotes the indicator function of a set A. Since $s_n \nearrow \infty$ as $n\to\infty$, we get

$$I(|U_j| \geqslant \varepsilon s_n) \leqslant I(|U_j| \geqslant \varepsilon s_j) \qquad \text{for all } j\leqslant n.$$

Thus, it suffices to show that

$$s_n^{-2} \sum_{j=m_0}^{n} E[U_j^2 I(|U_j| \geqslant \varepsilon s_j)] \longrightarrow 0 \quad \text{as } n\to\infty. \tag{5.18}$$

If

$$\sum_{n=m_0}^{\infty} s_n^{-2} E[U_n^2 I(|U_n| \geqslant \varepsilon s_n)] < \infty, \tag{5.19}$$

using Kronecker's lemma, we have (5. 18). As in [4], it follows that

$$s_n^{-2} E[U_n^2 I(|U_n| \geqslant \varepsilon s_n)]$$

$$\sim \sigma^{-2}(2A\alpha_1-1) n^{-1} E[v_n^2 I(|v_n| \geqslant \varepsilon' n^{\frac{1}{2}})] \qquad \text{as } n\to\infty,$$

where $\varepsilon' = \varepsilon\sigma(2A\alpha_1 - 1)^{-\frac{1}{2}}$
Therefore, it suffices to prove

$$\sum_{n=m_0}^{\infty} n^{-1} E[v_n^2 I(|v_n| \geqslant \varepsilon n^{\frac{1}{2}})] < \infty \qquad \text{for all } \varepsilon > 0.$$

From (5. 2), we have

$$\sum_{n=m_0}^{\infty} n^{-1} E[v_n^2 I(|v_n| \geqslant \varepsilon n^{\frac{1}{2}})]$$

$$\leqslant \varepsilon^{-\eta} \sum_{n=m_0}^{\infty} n^{-1-\frac{1}{2}\eta} E[|v_n|^{2+\eta}]$$

$$\leqslant \varepsilon^{-\eta} \sum_{n=m_0}^{\infty} n^{-1-\frac{1}{2}\eta} \sup_{-\infty < x < \infty} E[|V(x)|^{2+\eta}] < \infty.$$

Hence, (5. 17) is proved.
Secondly, we shall verify

$$s_n^{-2} \sum_{m=m_0}^{n} E[U_m^2 \mid \mathfrak{A}_{m-1}] \longrightarrow 1 \quad \text{a. s. as } n \to \infty. \tag{5. 20}$$

Since

$$E[v_m^2 \mid \mathfrak{A}_{m-1}] \longrightarrow \sigma^2 \qquad \text{a. s. as } m \to \infty$$

and

$$\sum_{m=m_0}^{n} m^{-2} \gamma_m^{-2} \nearrow \infty \quad \text{as } n \to \infty,$$

it follows, using Toeplitz's lemma, that

$$\left(\sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2}\right)^{-1} \left(\sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2} E[v_m^2 \mid \mathfrak{A}_{m-1}]\right) \longrightarrow \sigma^2 \quad \text{a. s. as } n \to \infty. \tag{5. 21}$$

Also, since $s_n^{-2} \sim \sigma^{-2}(2A\alpha_1 - 1) n \gamma_n^2$ as $n \to \infty$, we get

$$s_n^{-2}\left(\sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2}\right) \tag{5. 22}$$

$$\sim \sigma^{-2}(2A\alpha_1 - 1) n \gamma_n^2 (2A\alpha_1 - 1)^{-1} \gamma_n^{-2} n^{-1}$$

$$= \sigma^{-2} \qquad \text{as } n \to \infty.$$

Thus, from (5. 21) and (5. 22), we have

$$s_n^{-2} \sum_{m=m_0}^{n} E[U_m^2 \mid \mathfrak{A}_{m-1}] \tag{5. 23}$$

$$=s_n^{-2}\left(\sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2}\right)\times\left(\sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2}\right)^{-1}\left(\sum_{m=m_0}^{n} m^{-2}\gamma_m^{-2} E[v_m^2 \mid \mathfrak{A}_{m-1}]\right)$$

$$\longrightarrow 1 \quad \text{a.s. as } n\to\infty.$$

Hence, (5. 20) is proved.

Therefore, by Theorem 2 in [2], we obtain

$$S_n/s_n \longrightarrow N(0, 1) \quad \text{in law as } n\to\infty.$$

Since

$$n^{\frac{1}{2}} \sum_{m=m_0}^{n} m^{-1} \beta_{mn} v_m = n^{\frac{1}{2}} \gamma_n s_n (S_n/s_n)$$

and

$$n^{\frac{1}{2}} \gamma_n s_n \sim \sigma(2A\alpha_1-1)^{-\frac{1}{2}} \quad \text{as } n\to\infty,$$

we get

$$\text{(5. 24)} \qquad n^{\frac{1}{2}} \sum_{m=m_0}^{n} m^{-1} \beta_{mn} v_m \longrightarrow N(0, \sigma^2(2A\alpha_1-1)^{-1}) \quad \text{in law as } n\to\infty.$$

Therefore the relations (5. 15), (5. 16) and (5. 24) yield the conclusion of the theorem, which completes the proof.

References

[1] ANBAR, D. *A modified Robbins-Monro procedure approximating the zero of a regression function from below.* Ann. Statist., Vol. 5, No. 1, (1977), 229–234.

[2] BROWN, B. M. *Martingale central limit theorems.* Ann. Math. Stat., Vol. 42, (1971), 59–66.

[3] DUPAČ, V. and KRÁL, F. *Robbins-Monro procedure with both variables subject to experimental error.* Ann. Math. Stat., Vol. 43, No. 4, (1972), 1089–1095.

[4] HEYDE, C. C. *On martingale limit theory and strong convergence results for stochastic approximation procedures.* Stochastic Procosses Appl. 2, (1974), 359–370.

[5] WATANABE, M. *On convergence of asymptotically optimal discriminant functions for pattern classification problems.* Bull. Math. Stat. 16, 1–2, (1974), 23–34.

[6] WATANABE, M. *On RM stochastic approximation method with errors in setting the x-levels.* Fukuoka Univ. Sci. Reports. Vol. 5, (1975), 13–19.