A DYNAMIC STOCHASTIC APPROXIMATION METHOD

By Václav Dupač

Charles University, Prague

1. Introduction and summary. This investigation has been inspired by a paper of V. Fabian [3], where *inter alia* the applicability of stochastic approximation methods for progressive improvement of production processes is discussed; his non-formal discussion includes the case where the optimum of the production process moves during the optimization process.

In the present paper, the last case is treated in a formal way. A modified approximation scheme is suggested, which turns out to be an adequate tool, when the position of the optimum is a linear (or nearly linear) function of time. The domain of effectiveness of the unmodified approximation scheme is also investigated. In this context, the incorrectness of a theorem of T. Kitagawa is pointed out.

The considerations are performed for the Robbins-Monro case in detail; they can all be repeated for the Kiefer-Wolfowitz case and for the multidimensional case, as indicated in Section 4. Among the properties of the method, only the mean convergence and the order of magnitude of $E[(x_n - \theta_n)^2]$ are investigated; (here x_n denotes the estimated and θ_n the true position of the optimum at time n.)

A lemma, due to K. L. Chung [1] is used repeatedly:

Lemma. *Let b_n, $n = 1, 2, \cdots$, be real numbers such that for $n \geqq n_0$,*

$$b_{n+1} \leqq (1 - c/n^s)b_n + c'/n^t, \tag{1}$$

where $0 < s < 1$, $c > 0$, $c' > 0$, *t real. Then*

$$\limsup_{n\to\infty} n^{t-s} b_n \leqq c'/c. \tag{2}$$

The lemma remains true, if the inequalities (1) and (2) are reversed and, simultaneously, $\limsup_{n\to\infty}$ is changed into $\liminf_{n\to\infty}$. (In Chung's paper, a further assumption $t > s$ is made, but it is easily seen, that both versions of the lemma hold true also when $t \leqq s$; this fact is used in Section 3.)

Throughout the paper, $K_0, K_1, K_2, \cdots$ denote positive constants, numbered in order of appearance.

2. The modified Robbins-Monro method. Let $M(x)$, $-\infty < x < +\infty$, be an (unknown) real function. Let θ_n, $n = 1, 2, \cdots$, be (unknown) real numbers, the first, θ_1, being the, unique root of the equation $M(x) = 0$. Set $M_1(x) = M(x)$; for $n = 1, 2, \cdots$, set $M_n(x) = M(x - \theta_n + \theta_1)$ so that θ_n is the unique root of $M_n(x) = 0$. Let a_n, $n = 1, 2, \cdots$, be positive numbers. Let x_1 be an arbitrary random variable; define for $n = 1, 2, \cdots$,

Received 9 December 1964; revised 11 June 1965.

$$x_{n+1} = x_n^* - a_n y_n^*, \tag{3}$$

where $x_n^* = (1 + n^{-1})x_n$ and y_n^* is a rv such that

$$E[y_n^* \mid x_1, \cdots, x_n] = M_{n+1}(x_n^*), \tag{4}$$

$$\text{Var}\,[y_n^* \mid x_1, \cdots, x_n] \leqq \sigma^2 \tag{5}$$

with some constant σ^2.

The meaning of the scheme (3) is the following: at the time-instant $n + 1$, we try to determine an approximate value for θ_{n+1}. We start from the preceding approximation x_n, make first a correction for trend ($x_n^* = (1 + n^{-1})x_n$), then estimate the value of the instantaneous regression function M_{n+1} at x_n^* by means of the observation y_n^* and, finally, take a further correction, $-a_n y_n^*$. It will be seen from the next theorem and its corollary, that the use of this scheme is justified, when θ_n is a linear (nearly linear) function of n.

THEOREM 1. *Suppose that the following conditions are satisfied:*

$$M(x) < 0 \quad \textit{for} \quad x < \theta_1 \quad \textit{and} \quad M(x) > 0 \quad \textit{for} \quad x > \theta_1. \tag{6}$$

There exist K_0, K_1 *such that*

$$K_0\,|x - \theta_1| \leqq |M(x)| \leqq K_1\,|x - \theta_1| \quad \textit{for} \quad -\infty < x < +\infty. \tag{7}$$

For $n = 1, 2, \cdots,$

$$a_n = a/n^\alpha, \qquad a > 0, \qquad \tfrac{1}{2} < \alpha < 1. \tag{8}$$

θ_n *varies in such a way, that*

$$\theta_{n+1} - (1 + n^{-1})\theta_n = O(n^{-\omega}), \quad \textit{where } \omega > \alpha. \tag{9}$$

Further,

$$E[x_1^2] < +\infty. \tag{10}$$

Then $x_n - \theta_n \to 0$ *in the mean, and*

$$\begin{aligned} E[(x_n - \theta_n)^2] &= O(n^{-\alpha}) \quad \textit{for} \quad \omega \geqq \tfrac{3}{2}\alpha, \\ &= O(n^{-2(\omega-\alpha)}) \quad \textit{for} \quad \omega < \tfrac{3}{2}\alpha. \end{aligned} \tag{11}$$

PROOF. From (6) and (7) it follows that

$$M(x) = (x - \theta_1)\mu \quad \text{for} \quad -\infty < x < +\infty, \tag{12}$$

where μ denotes (here as in the sequel) a quantity, the dependence of which on x is not pointed out and which satisfies

$$K_0 \leqq \mu \leqq K_1. \tag{13}$$

Hence,

$$M_n(x) = M(x - \theta_n + \theta_1) = (x - \theta_n)\mu. \tag{14}$$

Calculate the conditional expectation of y_n^*, using successively (4), (14), (9):

$$(15)\qquad E[y_n^* \mid x_1, \cdots, x_n] = M_{n+1}(x_n^*) = (x_n^* - \theta_{n+1})\mu = (1 + n^{-1})(x_n - \theta_n)\mu + O(n^{-\omega}),$$

and form an upper bound for the conditional expectation of y_n^{*2}, using (5), (15), (13):

$$(16)\qquad E[y_n^{*2} \mid x_1, \cdots, x_n] \leqq \sigma_1^2 + K_2(x_n - \theta_n)^2$$

for some $\sigma_1^2 > \sigma^2$, $K_2 > 2K_1^2$ and for sufficiently large n. (In the sequel, all the inequalities in which the constants K_i occur hold only for sufficiently large n; but we shall not repeat that phrase.) Subtract θ_{n+1} on both sides of (3), substitute according to (9) on the right-hand side and then square:

$$(17)\qquad x_{n+1} - \theta_{n+1} = (1 + n^{-1})(x_n - \theta_n) + O(n^{-\omega}) - (a/n^\alpha)\, y_n^*;$$

$$(18)\qquad (x_{n+1} - \theta_{n+1})^2 = (1 + n^{-1})^2(x_n - \theta_n)^2 + O(n^{-2\omega}) + (a^2/n^{2\alpha})y_n^{*2} + O(n^{-\omega})(x_n - \theta_n) - (2a/n^\alpha)(1 + n^{-1})y_n^*(x_n - \theta_n) + O(n^{-(\alpha+\omega)})y_n^*.$$

Now take conditional expectations on both sides of (18) and use (15), (16) and (13):

$$(19)\qquad E[(x_{n+1} - \theta_{n+1})^2 \mid x_1, \cdots, x_n] \leqq (1 + n^{-1})^2(x_n - \theta_n)^2 + K_3/n^{2\omega} + a^2\sigma_1^2/n^{2\alpha} + (a^2K_2/n^{2\alpha})(x_n - \theta_n)^2 + (K_4/n^\omega)|x_n - \theta_n| - (2aK_0/n^\alpha)(1 + n^{-1})^2(x_n - \theta_n)^2 + (K_5/n^{\alpha+\omega})|x_n - \theta_n| + K_6/n^{\alpha+2\omega}.$$

After enlarging the corresponding coefficients, the terms of lower order of magnitude will include the terms of higher order:

$$(20)\qquad E[(x_{n+1} - \theta_{n+1})^2 \mid x_1, \cdots, x_n] \leqq (1 - K_7/n^\alpha)(x_n - \theta_n)^2 + (K_8/n^\omega)|x_n - \theta_n| + K_9/n^{2\alpha},$$

where $K_7 < 2aK_0$, etc.

Now, we take (unconditional) expectations on both sides of (20); when estimating $E[|x_n - \theta_n|]$, we use the inequality

$$(21)\qquad E[|z|] \leqq \epsilon + \epsilon^{-1}E[z^2]$$

(holding true for every $\epsilon > 0$ and every rv z with finite variance), where we set $\epsilon = 1/(\delta n^{\omega-\alpha})$ for some small $\delta > 0$. We get

$$(22)\qquad E[(x_{n+1} - \theta_{n+1})^2] \leqq (1 - K_7/n^\alpha)E[(x_n - \theta_n)^2] + K_8\delta^{-1}/n^{2\omega-\alpha} + (K_8\delta/n^\alpha)E[(x_n - \theta_n)^2] + K_9/n^{2\alpha},$$

consequently,

$$(23)\quad E[(x_{n+1} - \theta_{n+1})^2] \leqq (1 - K_{10}/n^{\alpha})E[(x_n - \theta_n)^2] + K_{11}/n^{2\omega-\alpha} \text{ for } \omega < \tfrac{3}{2}\alpha,$$
$$\leqq (1 - K_{10}/n^{\alpha})E[(x_n - \theta_n)^2] + K_{12}/n^{2\alpha} \text{ for } \omega \geqq \tfrac{3}{2}\alpha.$$

The application of Chung's lemma completes the proof.

COROLLARY 1. *Under the assumptions of Theorem* 1, *let θ_n be a linear function of n, then*

$$(24)\quad E[(x_n - \theta_n)^2] = O(n^{-\alpha}) \quad \textit{for} \quad \tfrac{1}{2} < \alpha \leqq \tfrac{2}{3},$$
$$= O(n^{-[2(1-\alpha)]}) \quad \textit{for} \quad \tfrac{2}{3} < \alpha < 1;$$

let θ_n be proportional to n, then

$$(25)\quad E[(x_n - \theta_n)^2] = O(n^{-\alpha}) \textit{ for } \tfrac{1}{2} < \alpha < 1.$$

PROOF. In the linear case, $\theta_n = \lambda n + \rho$; hence $\theta_{n+1} - (1 + n^{-1})\theta_n = -\rho/n$, so that (9) holds with $\omega = 1$; in the proportionality case, we have moreover $\rho = 0$, so that (9) holds with ω arbitrarily large. This proves the corollary.

The proportionality case can be achieved, when the variation of θ_n is known to be linear and when the exact value of the root θ is known at some time-instant; we need only to choose this value for zero-value and the corresponding time-instant for zero-time, i.e. $\theta_0 = 0$.

The assumption (9) about the variation of θ_n is satisfied—besides by the linear function—by functions of the type, e.g.,

$$(26)\quad \theta_n = cn^{\rho}, \quad \text{with} \quad -\alpha \leqq \rho < 1 - \alpha \quad \text{and} \quad c \quad \text{real},$$

or by every function

$$(27)\quad \theta_n = O(n^{-\tau}) \quad \text{with} \quad \tau > \alpha,$$

and, of course, by linear combinations of all these functions.

For the case, when (9) is not satisfied, only the following partial result may be given. If, instead of (9), the relation $\lim_{n\to\infty} n^{\rho}(\theta_{n+1} - (1 + n^{-1})\theta_n) = q$ holds for some $-\infty < \rho < \alpha$ and $0 < |q| < +\infty$, then the non-random example $M(x) = x$, $x_1 = 0$, $\sigma^2 = 0$ satisfies all the other conditions of Theorem 1, but $x_n - \theta_n$ diverges to infinity. Indeed, we get in this case

$$\theta_{n+1} - x_{n+1} = (1 - (a + o(1))/n^{\alpha})(\theta_n - x_n) + (q + o(1))/n^{\rho};$$

for q positive, the application of Chung's lemma (cf. Section 1) gives $\liminf_{n\to\infty} n^{\rho-\alpha}(\theta_n - x_n) > 0$, so that $\theta_n - x_n \to +\infty$; similarly, $x_n - \theta_n \to +\infty$ for q negative.

3. The unmodified Robbins-Monro method. We shall now investigate, how the unmodified Robbins-Monro procedure works in the presence of trend. We first change the definition of x_n :

$$(3')\quad x_{n+1} = x_n - a_n y_n, \qquad n = 1, 2, \cdots,$$

where y_n is a rv, such that

$$(4')\quad E[y_n \mid x_1, \cdots, x_n] = M_n(x_n),$$

(5′) $$\text{Var}\,[y_n \mid x_1, \cdots, x_n] \leqq \sigma^2.$$

(Theorem 2 holds true, irrespective of the definition of $E[y_n \mid x_1, \cdots, x_n]$ either as $M_n(x_n)$ or as $M_{n+1}(x_n)$.) With this re-definition of x_n, Theorem 1 is falsified in general; the case $M(x) = x$, $x_1 = 0$, $\sigma^2 = 0$ can again serve as a counter-example, whenever the trend of θ_n is such that $\lim_{n\to\infty} n^{\rho}(\theta_{n+1} - \theta_n) = q$ for some $-\infty < \rho < \alpha$ and $0 < |q| < +\infty$. We shall, therefore, replace also (9) by a stronger condition

(9′) $$\theta_{n+1} - \theta_n = O(n^{-\omega}), \quad \text{with} \quad \omega > \alpha;$$

this is again satisfied by functions (26) and (27), but no more by the linear function.

THEOREM 2. *Under the assumptions* (3′), (4′), (5′), (6), (7), (8), (9′), (10), *we have*

(28) $$E[(x_n - \theta_n)^2] = O(n^{-\alpha}) \quad for \quad \omega \geqq \tfrac{3}{2}\alpha;$$
$$= O(n^{-[2(\omega-\alpha)]}) \quad for \quad \omega < \tfrac{3}{2}\alpha.$$

PROOF. As in the preceding theorem, it holds (12), for it follows from (6) and (7) only. Instead of (15) and (16), we get now

(29) $$E[y_n \mid x_1, \cdots, x_n] = (x_n - \theta_n)\mu,$$

(30) $$E[y_n^2 \mid x_1, \cdots, x_n] \leqq \sigma^2 + K_{13}(x_n - \theta_n)^2.$$

The rest of the proof is quite similar to that of Theorem 1 and will be omitted.

A related problem has been treated by T. Kitagawa ([4], p. 12, Theorem 4.2). Under the assumption $K_{14} \leqq (M(x) - M(x'))/(x - x') \leqq K_{15}$ for real x, x', and under the usual assumptions about $\{a_n\}$, Kitagawa considers a given sequence of real numbers $\{\alpha_n\}$ such that the roots θ_n of the equations $M(x) = \alpha_n$, $n = 1, 2, \cdots$, satisfy

(31) $$\sum_{n=1}^{\infty} (\theta_n - \theta_{n+1})^2 < +\infty;$$

then he defines the scheme $x_{n+1} = x_n + a_n(\alpha_{n+1} - y_n)$, with the standard meaning of y_n ($E[y_n \mid x_1, \cdots, x_n] = M(x_n)$, $\text{Var}\,[y_n \mid x_1, \cdots, x_n] \leqq \sigma^2$) and asserts, that $E[(x_n - \theta_n)^2] \to 0$. But this theorem is false, as can again be shown by the following counter-example: $a_n = n^{-\alpha}$, $\frac{1}{2} < \alpha < 1$; $M(x) = x$, $x_1 = 0$, $\sigma^2 = 0$, $\lim_{n\to\infty} n^{\rho}(\theta_{n+1} - \theta_n) = q$, for some $\frac{1}{2} < \rho < \alpha$ and $0 < |q| < +\infty$, which yields the divergence of $\theta_n - x_n$ to infinity. (The gap in Kitagawa's proof is that the constants $F_n = 2^{\frac{1}{2}}(1 - Aa_n)$ do *not* fulfill the condition $\prod_{n=1}^{\infty} F_n^2 = 0$.)

The Kitagawa's theorem will hold true, together with the same order-estimates as in our Theorem 2, if we choose the constants a_n according to (8) and replace (31) by the condition (9′). The proof is then entirely similar to that of Theorem 2.

4. The Kiefer-Wolfowitz method. The investigations made for the Robbins-Monro case in preceding two sections, can all be repeated for the Kiefer-Wolfowitz case as for the multidimensional case. We shall state only the analogue of Theorem 1.

Let $M(x)$, $-\infty < x < +\infty$, be a real function, let θ_n, $n = 1, 2, \cdots$, be real numbers, θ_1 being the value at which $M(x)$ achieves its unique maximum. Set $M_1(x) = M(x)$, for $n = 1, 2, \cdots$ set $M_n(x) = M(x - \theta_n + \theta_1)$. Let a_n, c_n, $n = 1, 2, \cdots$ be two sequences of positive numbers. Let x_1 be an arbitrary rv; define for $n = 1, 2, \cdots$,

$$x_{n+1} = x_n^* + a_n(y_{2n}^* - y_{2n-1}^*)/c_n, \tag{32}$$

where $x_n^* = (1 + n^{-1})x_n$ and y_{2n}^*, y_{2n-1}^* are rv's such that their conditional expectations, given $x_1, \cdots, x_n$, are $M_{n+1}(x_n^* + c_n)$, $M_{n+1}(x_n^* - c_n)$ respectively, their conditional variances are bounded by a constant σ^2, and they are conditionally independent.

Theorem 3. *Suppose that the following conditions are satisfied: $M(x)$ is increasing for $x < \theta_1$ and decreasing for $x > \theta_1$. There exist K_{16}, K_{17}, K_{18} such that*

$$\begin{aligned} &K_{16}|x - \theta_1| \leqq |M'(x)| \leqq K_{17}|x - \theta_1|, \\ &|M'''(x)| \leqq K_{18} \qquad \text{for} \quad -\infty < x < +\infty. \end{aligned} \tag{33}$$

For $n = 1, 2, \cdots$,

$$a_n = a/n^\alpha, \qquad a > 0, \qquad \tfrac{3}{5} < \alpha < 1, \tag{34}$$

$$c_n = c/n^\gamma, \qquad c > 0, \qquad \alpha/6 \leqq \gamma < \alpha - \tfrac{1}{2}. \tag{35}$$

θ_n varies in such a way that

$$\theta_{n+1} - (1 + n^{-1})\theta_n = O(n^{-\omega}), \quad \text{where} \quad \omega > \alpha. \tag{36}$$

Further, $E[x_1^2] < +\infty$. Then $x_n - \theta_n \to 0$ in the mean, and

$$E[(x_n - \theta_n)^2] = \begin{cases} O(n^{-(\alpha-2\gamma)}) & \text{for} \quad \omega \geqq \tfrac{3}{2}\alpha - \gamma, \\ O(n^{-[2(\omega-\alpha)]}) & \text{for} \quad \omega < \tfrac{3}{2}\alpha - \gamma. \end{cases} \tag{37}$$

Proof. Let us denote $D_cM(x) = (M(x + c) - M(x - c))/c$. We have from (33):

$$\begin{aligned} D_cM(x) &= 2M'(x) + \tfrac{1}{6}c^2\{M'''(x + \vartheta_1 c) + M'''(x - \vartheta_2 c\} \\ &= -(x - \theta_1)\mu + \lambda c^2, \quad \text{where} \quad 2K_{16} \leqq \mu \leqq 2K_{17}, \ |\lambda| \leqq \tfrac{1}{3}K_{18}; \end{aligned} \tag{38}$$

hence,

$$D_cM_n(x) = D_cM(x - \theta_n + \theta_1) = -(x - \theta_n)\mu + \lambda c^2. \tag{39}$$

Further, by (39), (36) and (35),

$$\begin{aligned} E[(y_{2n}^* - y_{2n-1}^*)/c_n \mid x_1, \cdots, x_n] &= D_{c_n}M_{n+1}(x_n^*) \\ &= -(x_n^* - \theta_{n+1})\mu + \lambda c_n^2 \\ &= -\mu(1 + n^{-1})(x_n - \theta_n) + \lambda c^2/n^{2\gamma} + O(n^{-\omega}) \end{aligned} \tag{40}$$

(the last term will be neglected in order estimates, because $\gamma < \alpha - \frac{1}{2}$ together with $\omega > \alpha$ implies $\omega > 2\gamma$);

$$(41)\quad E\{[(y_{2n}^* - y_{2n-1}^*)/c_n]^2 \mid x_1, \cdots, x_n\} \leqq 2\sigma^2/c_n^2 + [D_{c_n}M_{n+1}(x_n^*)]^2 \leqq K_{19}n^{2\gamma} + K_{20}(x_n - \theta_n)^2.$$

Subtract θ_{n+1} from both sides of (32), substitute on the right-hand side according to (36), re-arrange and square; we get

$$(42)\quad \begin{aligned}(x_{n+1} - \theta_{n+1})^2 = {} & (1 + n^{-1})^2(x_n - \theta_n)^2 + O(n^{-2\omega}) \\ & + (a^2/n^{2\alpha})[(y_{2n}^* - y_{2n-1}^*)/c_n]^2 + O(n^{-\omega})(x_n - \theta_n) \\ & + (2a/n^{\alpha})(1 + n^{-1})[(y_{2n}^* - y_{2n-1}^*)/c_n](x_n - \theta_n) \\ & + O(n^{-(\omega+\alpha)})(y_{2n}^* - y_{2n-1}^*)/c_n .\end{aligned}$$

Take conditional expectations, use (40) and (41) and neglect terms of higher order; we get

$$(43)\quad \begin{aligned}E[(x_{n+1} - \theta_{n+1})^2 \mid x_1, \cdots, x_n] \leqq {} & (1 - K_{21}/n^{\alpha})(x_n - \theta_n)^2 \\ & + (K_{22}/n^{\omega})|x_n - \theta_n| + [K_{23}/(n^{\alpha+2\gamma})]|x_n - \theta_n| + K_{24}/n^{2\alpha-2\gamma}.\end{aligned}$$

Take unconditional expectations; the terms with $E[|x_n - \theta_n|]$ estimate with help of (21), setting $\epsilon = 1/(\delta n^{\omega-\alpha})$ and $\epsilon = 1/(\delta' n^{2\gamma})$, respectively. We get finally

$$(44)\quad \begin{aligned}E[(x_{n+1} - \theta_n)^2] \leqq {} & (1 - K_{25}/n^{\alpha})E[(x_n - \theta_n)^2] \\ & + K_{26}/n^{2\omega-\alpha} + K_{27}/n^{\alpha+4\gamma} + K_{24}/n^{2\alpha-2\gamma},\end{aligned}$$

but the term $K_{27}/n^{\alpha+4\gamma}$ can be joined to the last term, for the condition $\gamma \geqq \alpha/6$ implies $\alpha + 4\gamma \geqq 2\alpha - 2\gamma$. Chung's lemma gives the statement of the theorem.

5. The use of Dvoretzky's theorem. The mean-square convergence in all above theorems (as well as convergence with probability one) can also be deduced from Dvoretzky's theorem [2], [5], even under slightly more general conditions on θ_n. We shall indicate this implication only for the method of Section 2. Let us replace the conditions (8) and (9) by the more general

$$(8^*)\qquad \lim_{n\to\infty} na_n = +\infty, \qquad \sum_{n=1}^{\infty} a_n^2 < +\infty,$$

$$(9^*)\qquad \theta_{n+1} - (1 + n^{-1})\theta_n = o(a_n).$$

THEOREM 4. *Under the assumptions* (3), (4), (5), (6), (7), (8*), (9*), (10) *it holds that*

$$\lim_{n\to\infty} E[(x_n - \theta_n)^2] = 0 \quad \textit{and} \quad P(\lim_{n\to\infty} (x_n - \theta_n) = 0) = 1.$$

PROOF. Set $z_n = x_n - \theta_n$, $\epsilon_n = -a_n(y_n^* - M_{n+1}(x_n^*))$, $\omega_n = \theta_{n+1} - (1 + n^{-1})\theta_n$. Then the scheme (3) can be rewritten as $z_{n+1} = T_n(z_1, \cdots, z_n) + \epsilon_n$, where

$$(45)\quad T_n(r_1, \cdots, r_n) = (1 + n^{-1})r_n - \omega_n - a_nM((1 + n^{-1})r_n - \omega_n + \theta_1).$$

Among the conditions of Dvoretsky's theorem, $\sum_{n=1}^{\infty} E[\epsilon_n^2] < +\infty$ and $E(\epsilon_n \mid z_1, \cdots, z_n) = 0$ are evidently satisfied; it remains to prove

$$|T_n(r_1, \cdots, r_n)| \leqq \text{Max}\,(\alpha_n, |r_n| - \gamma_n) \tag{46}$$

for some $\alpha_n \rightarrow 0$ and $\sum_{n=1}^{\infty} \gamma_n = +\infty$.

From (8*) and (9*) it follows, that there exist $\rho_n > 0$, such that $\rho_n \rightarrow 0$, $\sum_{n=1}^{\infty} a_n \rho_n = +\infty$ and $\omega_n = o(a_n \rho_n)$.

Using (7) we get

$$T_n(r_1, \cdots, r_n) = (1 - a_n \mu)\{(1 + n^{-1}) r_n - \omega_n\}, \tag{47}$$

whence

$$|T_n(r_1, \cdots, r_n)| \leqq (1 - K_0 a_n)\{(1 + n^{-1})|r_n| + |\omega_n|\} \tag{48}$$

for $n > n_0$.

Now, if $|r_n| \leqq \rho_n$, then

$$|T_n(r_1, \cdots, r_n)| \leqq (1 + n^{-1})\rho_n + |\omega_n| < 2\rho_n; \tag{49}$$

if $|r_n| > \rho_n$, then

$$\begin{aligned} |T_n(r_1, \cdots, r_n)| &\leqq (1 - K_0 a_n + o(a_n))|r_n| + |\omega_n| \\ &\leqq |r_n| - \tfrac{1}{2} K_0 a_n \rho_n. \end{aligned} \tag{50}$$

Thus the condition (46) is satisfied with $\alpha_n = 2\rho_n$ and $\gamma_n = \frac{1}{2} K_0 a_n \rho_n$.

REFERENCES

[1] CHUNG, K. L. (1954). On a stochastic approximation method. *Ann. Math. Statist.* **25** 463–483.

[2] DVORETZKY, A. (1956). On stochastic approximation. *Proc. Third Berkeley Symp. Statist. Prob.* **1** 33–55. Univ. California Press.

[3] FABIAN, V. (1961). A stochastic approximation method for finding optimal conditions in experimental work and in self-adapting systems (in Czech). *Apl. Mat.* **6** 162–183.

[4] KITAGAWA, T. (1959). Successive processes of statistical controls II. *Mem. Fac. Sci. Kyushu Univ. Ser. A* **13** 1–16.

[5] WOLFOWITZ, J. (1956). On stochastic approximation methods. *Ann. Math. Statist.* **27** 1151–1156.