A STOCHASTIC APPROXIMATION METHOD[1]

By Herbert Robbins and Sutton Monro

University of North Carolina

1. Summary. Let $M(x)$ denote the expected value at level x of the response to a certain experiment. $M(x)$ is assumed to be a monotone function of x but is unknown to the experimenter, and it is desired to find the solution $x = \theta$ of the equation $M(x) = \alpha$, where α is a given constant. We give a method for making successive experiments at levels $x_1, x_2, \cdots$ in such a way that x_n will tend to θ in probability.

2. Introduction. Let $M(x)$ be a given function and α a given constant such that the equation

$$M(x) = \alpha \tag{1}$$

has a unique root $x = \theta$. There are many methods for determining the value of θ by successive approximation. With any such method we begin by choosing one or more values $x_1, \cdots, x_r$ more or less arbitrarily, and then successively obtain new values x_n as certain functions of the previously obtained $x_1, \cdots, x_{n-1}$, the values $M(x_1), \cdots, M(x_{n-1})$, and possibly those of the derivatives $M'(x_1), \cdots, M'(x_{n-1})$, etc. If

$$\lim_{n\to\infty} x_n = \theta, \tag{2}$$

irrespective of the arbitrary initial values $x_1, \cdots, x_r$, then the method is effective for the particular function $M(x)$ and value α. The speed of the convergence in (2) and the ease with which the x_n can be computed determine the practical utility of the method.

We consider a stochastic generalization of the above problem in which the nature of the function $M(x)$ is unknown to the experimenter. Instead, we suppose that to each value x corresponds a random variable $Y = Y(x)$ with distribution function $Pr[Y(x) \le y] = H(y \mid x)$, such that

$$M(x) = \int_{-\infty}^{\infty} y \, dH(y \mid x) \tag{3}$$

is the expected value of Y for the given x. Neither the exact nature of $H(y \mid x)$ nor that of $M(x)$ is known to the experimenter, but it is assumed that equation (1) has a unique root θ, and it is desired to estimate θ by making successive observations on Y at levels $x_1, x_2, \cdots$ determined sequentially in accordance with some definite experimental procedure. If (2) holds *in probability* irrespective of any arbitrary initial values $x_1, \cdots, x_r$, we shall, in conformity with usual statistical terminology, call the procedure *consistent* for the given $H(y \mid x)$ and value α.

[1] This work was supported in part by the Office of Naval Research.

In what follows we shall give a particular procedure for estimating θ which is consistent under certain restrictions on the nature of $H(y \mid x)$. These restrictions are severe, and could no doubt be lightened considerably, but they are often satisfied in practice, as will be seen in Section 4. No claim is made that the procedure to be described has any optimum properties (i.e. that it is "efficient") but the results indicate at least that the subject of stochastic approximation is likely to be useful and is worthy of further study.

3. Convergence theorems. We suppose henceforth that $H(y \mid x)$ is, for every x, a distribution function in y, and that there exists a positive constant C such that

$$Pr[\,|\,Y(x)\,| \leq C] = \int_{-C}^{C} dH(y \mid x) = 1 \qquad \text{for all } x. \tag{4}$$

It follows in particular that for every x the expected value $M(x)$ defined by (3) exists and is finite. We suppose, moreover, that there exist finite constants α, θ such that

$$M(x) \leq \alpha \quad \text{for} \quad x < \theta, \qquad M(x) \geq \alpha \quad \text{for} \quad x > \theta. \tag{5}$$

Whether $M(\theta) = \alpha$ is, for the moment, immaterial.

Let $\{a_n\}$ be a fixed sequence of positive constants such that

$$0 < \sum_1^\infty a_n^2 = A < \infty. \tag{6}$$

We define a (nonstationary) Markov chain $\{x_n\}$ by taking x_1 to be an arbitrary constant and defining

$$x_{n+1} - x_n = a_n(\alpha - y_n), \tag{7}$$

where y_n is a random variable such that

$$Pr[y_n \leq y \mid x_n] = H(y \mid x_n). \tag{8}$$

Let

$$b_n = E(x_n - \theta)^2. \tag{9}$$

We shall find conditions under which

$$\lim_{n \to \infty} b_n = 0 \tag{10}$$

no matter what the initial value x_1. As is well known, (10) implies the convergence in probability of x_n to θ.

From (7) we have

$$\begin{aligned} b_{n+1} &= E(x_{n+1} - \theta)^2 = E[E[(x_{n+1} - \theta)^2 \mid x_n]] \\ (11) \qquad &= E\left[\int_{-\infty}^{\infty} \{(x_n - \theta) - a_n(y - \alpha)\}^2 \, dH(y \mid x_n)\right] \\ &= b_n + a_n^2 E\left[\int_{-\infty}^{\infty} (y - \alpha)^2 \, dH(y \mid x_n)\right] - 2a_n E[(x_n - \theta)(M(x_n) - \alpha)]. \end{aligned}$$

Setting

$$d_n = E[(x_n - \theta)(M(x_n) - \alpha)], \tag{12}$$

$$e_n = E\left[\int_{-\infty}^{\infty} (y - \alpha)^2 \, dH(y \mid x_n)\right], \tag{13}$$

we can write

$$b_{n+1} - b_n = a_n^2 e_n - 2a_n d_n. \tag{14}$$

Note that from (5)

$$d_n \geq 0,$$

while from (4)

$$0 \leq e_n \leq [C + |\alpha|]^2 < \infty.$$

Together with (6) this implies that the positive-term series $\Sigma a_n^2 e_n$ converges.
Summing (14) we obtain

$$b_{n+1} = b_1 + \sum_{j=1}^{n} a_j^2 e_j - 2 \sum_{j=1}^{n} a_j d_j. \tag{15}$$

Since $b_{n+1} \geq 0$ it follows that

$$\sum_{j=1}^{n} a_j d_j \leq \tfrac{1}{2}\left[b_1 + \sum_{1}^{\infty} a_n^2 e_n\right] < \infty. \tag{16}$$

Hence the positive-term series

$$\sum_{1}^{\infty} a_n d_n \tag{17}$$

converges. It follows from (15) that

$$\lim_{n\to\infty} b_n = b_1 + \sum_{1}^{\infty} a_n^2 e_n - 2 \sum_{1}^{\infty} a_n d_n = b \tag{18}$$

exists; $b \geq 0$.

Now suppose that there exists a sequence $\{k_n\}$ of nonnegative constants such that

$$d_n \geq k_n b_n, \qquad \sum_{1}^{\infty} a_n k_n = \infty. \tag{19}$$

From the first part of (19) and the convergence of (17) it follows that

$$\sum_{1}^{\infty} a_n k_n b_n < \infty. \tag{20}$$

From (20) and the second part of (19) it follows that for any $\epsilon > 0$ there must exist infinitely many values n such that $b_n < \epsilon$. Since we already know that $b = \lim_{n\to\infty} b_n$ *exists,* it follows that $b = 0$. Thus we have proved

LEMMA 1. *If a sequence* $\{k_n\}$ *of nonnegative constants exists satisfying* (19) *then* $b = 0$.

Let

$$A_n = |x_1 - \theta| + [C + |\alpha|](a_1 + a_2 + \cdots + a_{n-1}); \tag{21}$$

then from (4) and (7) it follows that

$$Pr[|x_n - \theta| \le A_n] = 1. \tag{22}$$

Now set

$$\bar{k}_n = \inf\left[\frac{M(x) - \alpha}{x - \theta}\right] \qquad \text{for} \qquad 0 < |x - \theta| \le A_n. \tag{23}$$

From (5) it follows that $\bar{k}_n \ge 0$. Moreover, denoting by $P_n(x)$ the probability distribution of x_n, we have

$$\begin{aligned} d_n &= \int_{|x-\theta|\le A_n} (x - \theta)(M(x) - \alpha)\, dP_n(x) \\ &\ge \int_{|x-\theta|\le A_n} \bar{k}_n |x - \theta|^2\, dP_n(x) = \bar{k}_n b_n. \end{aligned} \tag{24}$$

It follows that the particular sequence $\{\bar{k}_n\}$ defined by (23) satisfies the first part of (19).

In order to establish the second part of (19) we shall make the following assumptions:

$$\bar{k}_n \ge \frac{K}{A_n} \tag{25}$$

for some constant $K > 0$ and sufficiently large n, and

$$\sum_{n=2}^{\infty} \frac{a_n}{(a_1 + \cdots + a_{n-1})} = \infty. \tag{26}$$

It follows from (26) that

$$\sum_{1}^{\infty} a_n = \infty, \tag{27}$$

and hence for sufficiently large n

$$2[C + |\alpha|](a_1 + \cdots + a_{n-1}) \ge A_n. \tag{28}$$

This implies by (25) that for sufficiently large n

$$a_n \bar{k}_n \ge a_n \frac{K}{A_n} \ge \frac{a_n K}{2[C + |\alpha|](a_1 + \cdots + a_{n-1})}, \tag{29}$$

and the second part of (19) follows from (29) and (26). This proves

LEMMA 2. *If* (25) *and* (26) *hold then* $b = 0$.

The hypotheses (6) and (26) concerning $\{a_n\}$ are satisfied by the sequence $a_n = 1/n$, since

$$\sum_1^\infty \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=2}^\infty \left[\frac{1}{n\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right)}\right] = \infty$$

More generally, any sequence $\{a_n\}$ such that there exist two positive constants c', c'' for which

$$\frac{c'}{n} \leq a_n \leq \frac{c''}{n} \tag{30}$$

will satisfy (6) and (26). We shall call any sequence $\{a_n\}$ which satisfies (6) and (26), whether or not it is of the form (30), a *sequence of type* $1/n$.

If $\{a_n\}$ is a sequence of type $1/n$ it is easy to find functions $M(x)$ which satisfy (5) and (25). Suppose, for example, that $M(x)$ satisfies the following strengthened form of (5): for some $\delta > 0$,

$$M(x) \leq \alpha - \delta \quad \text{for} \quad x < \theta, \qquad M(x) \geq \alpha + \delta \quad \text{for} \quad x > \theta. \tag{5'}$$

Then for $0 < | x - \theta | \leq A_n$ we have

$$\frac{M(x) - \alpha}{x - \theta} \geq \frac{\delta}{A_n}, \tag{31}$$

so that

$$\bar{k}_n \geq \frac{\delta}{A_n}, \tag{32}$$

which is (25) with $K = \delta$. From Lemma 2 we conclude

THEOREM 1. *If* $\{a_n\}$ *is of type* $1/n$, *if* (4) *holds, and if* $M(x)$ *satisfies* (5′) *then* $b = 0$.

A more interesting case occurs when $M(x)$ satisfies the following conditions:

$$M(x) \text{ is nondecreasing}, \tag{33}$$

$$M(\theta) = \alpha, \tag{34}$$

$$M'(\theta) > 0. \tag{35}$$

We shall prove that (25) holds in this case also. From (34) it follows that

$$M(x) - \alpha = (x - \theta)[M'(\theta) + \epsilon(x - \theta)], \tag{36}$$

where $\epsilon(t)$ is a function such that

$$\lim_{t \to 0} \epsilon(t) = 0. \tag{37}$$

Hence there exists a constant $\delta > 0$ such that

$$\epsilon(t) \geq -\tfrac{1}{2}M'(\theta) \qquad \text{for} \qquad | t | \leq \delta, \tag{38}$$

so that

$$\frac{M(x) - \alpha}{x - \theta} \geq \frac{1}{2} M'(\theta) > 0 \quad \text{for} \quad |x - \theta| \leq \delta. \tag{39}$$

Hence, for $\theta + \delta \leq x \leq \theta + A_n$, since $M(x)$ is nondecreasing,

$$\frac{M(x) - \alpha}{x - \theta} \geq \frac{M(\theta + \delta) - \alpha}{A_n} \geq \frac{\delta M'(\theta)}{2A_n}, \tag{40}$$

while for $\theta - A_n \leq x \leq \theta - \delta$,

$$\frac{M(x) - \alpha}{x - \theta} = \frac{\alpha - M(x)}{\theta - x} \geq \frac{\alpha - M(\theta - \delta)}{A_n} \geq \frac{\delta M'(\theta)}{2A_n}. \tag{41}$$

Thus, since we may assume without loss of generality that $\delta/A_n \leq 1$,

$$\frac{M(x) - \alpha}{x - \theta} \geq \frac{\delta M'(\theta)}{2A_n} \quad \text{for} \quad 0 < |x - \theta| \leq A_n, \tag{42}$$

so that (25) holds with $K = \delta M'(\theta)/2 > 0$. This proves

THEOREM 2. *If $\{a_n\}$ is of type $1/n$, if* (4) *holds, and if $M(x)$ satisfies* (33), (34), *and* (35), *then* $b = 0$.

It is fairly obvious that condition (4) could be considerably weakened without affecting the validity of Theorems 1 and 2. A reasonable substitute for (4) would be the condition

$$|M(x)| \leq C, \qquad \int_{-\infty}^{\infty} (y - M(x))^2 \, dH(y \mid x) \leq \sigma^2 < \infty \qquad \text{for all } x. \tag{4'}$$

We do not know whether Theorems 1 and 2 hold with (4) replaced by (4′). Likewise, the hypotheses (33), (34), and (35) of Theorem 2 could be weakened somewhat, perhaps being replaced by

$$M(x) < \alpha \quad \text{for} \quad x < \theta, \qquad M(x) > \alpha \quad \text{for} \quad x > \theta. \tag{5''}$$

4. Estimation of a quantile using response, nonresponse data. Let $F(x)$ be an unknown distribution function such that

$$F(\theta) = \alpha \ (0 < \alpha < 1), \qquad F'(\theta) > 0, \tag{43}$$

and let $\{z_n\}$ be a sequence of independent random variables each with the distribution function $Pr[z_n \leq x] = F(x)$. On the basis of $\{z_n\}$ we wish to estimate θ. However, as sometimes happens in practice (bioassay, sensitivity data), we are not allowed to know the values of z_n themselves. Instead, we are free to prescribe for each n a value x_n and are then given only the values $\{y_n\}$ where

$$y_n = \begin{cases} 1 & \text{if } z_n \leq x_n \qquad (\text{"response"}), \\ 0 & \text{otherwise} \qquad (\text{"nonresponse"}). \end{cases} \tag{44}$$

How shall we choose the values $\{x_n\}$ and how shall we use the sequence $\{y_n\}$ to estimate θ?

Let us proceed as follows. Choose x_1 as our best guess of the value θ and let $\{a_n\}$ be any sequence of constants of type $1/n$. Then choose values $x_2, x_3, \cdots$ sequentially according to the rule

$$x_{n+1} - x_n = a_n(\alpha - y_n). \tag{45}$$

Since

$$Pr[y_n = 1 \mid x_n] = F(x_n), \qquad Pr[y_n = 0 \mid x_n] = 1 - F(x_n), \tag{46}$$

it follows that (4) holds and that

$$M(x) = F(x). \tag{47}$$

All the hypotheses of Theorem 4 are satisfied, so that

$$\lim_{n\to\infty} x_n = \theta \tag{48}$$

in quadratic mean and hence in probability. In other words, $\{x_n\}$ is a consistent estimator of θ.

The *efficiency* of $\{x_n\}$ will depend on x_1 and on the choice of the sequence $\{a_n\}$, as well as on the nature of $F(x)$. For any given $F(x)$ there doubtless exist more efficient estimators of θ than any of the type $\{x_n\}$ defined by (45), but $\{x_n\}$ has the advantage of being distribution-free.

In some applications it is more convenient to make a group of r observations at the same level before proceeding to the next level. The nth group of observations will then be

$$y_{(n-1)r+1}, \cdots, y_{nr}, \tag{49}$$

using the notation (44). Let $\bar{y}_n$ = arithmetic mean of the values (49). Then setting

$$x_{n+1} - x_n = a_n(\alpha - \bar{y}_n), \tag{50}$$

we have $M(x) = F(x)$ as before, and hence (48) continues to hold.

The possibility of using a convergent sequential process in this problem was first mentioned by T. W. Anderson, P. J. McCarthy, and J. W. Tukey in the Naval Ordnance Report No. 65-46(1946), p. 99.

5. A more general regression problem. It is clear that the problem of Section 4 is a special case of a more general regression problem. In fact, using the notation of Section 2, consider any random variable Y which is associated with an observable value x in such a way that the conditional distribution function of Y for fixed x is $H(y \mid x)$; the function $M(x)$ is then the regression of Y on x.

The usual regression analysis assumes that $M(x)$ is of known form with unknown parameters, say

$$M(x) = \beta_0 + \beta_1 x, \tag{51}$$

and deals with the estimation of one or both of the parameters β_i on the basis of observations $y_1, y_2, \cdots, y_n$ corresponding to observed values $x_1, x_2, \cdots, x_n$. The method of least squares, for example, yields the estimators b_i which minimize the expression

$$\sum_{i=1}^{n} (y_i - [\beta_0 + \beta_1 x_i])^2. \tag{52}$$

Instead of trying to estimate the parameters β_i of $M(x)$ under the assumption that $M(x)$ is a linear function of x, we may try to estimate the value θ such that $M(\theta) = \alpha$, where α is given, without any assumption about the form of $M(x)$. If we assume only that $H(y \mid x)$ satisfies the hypotheses of Theorem 2 then the sequence of estimators $\{x_n\}$ of θ defined by (7) will at least be consistent. This indicates that a distribution-free sequential system of making observations, such as that given by (7), is worth investigating from the practical point of view in regression problems.

One of us is investigating the properties of this and other sequential designs as a graduate student; the senior author is responsible for the convergence proof in Section 3.