The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 1134–1135

ARTHUR E. ALBERT AND LELAND A. GARDNER, JR., *Stochastic Approximation and Nonlinear Regression*. Research Monograph no. 42, The M.I.T. Press, Cambridge, 1967. xv + 204 pp. $16.50.

Review by VÁCLAV FABIAN

Michigan State University

The book extends the Robbins-Monro stochastic approximation method to the problem of sequential estimating of an unknown parameter θ from a sequence $\{Y_n\}$ of random variables with expectations $F_n(\theta)$, where F_n are known functions. The reason for the use of stochastic approximation is that it leads to computationally feasible procedures.

If all F_n are equal to an increasing function F then θ may be sought as the zero point of the regression function of $Z_n(t) = Y_n - F(t)$ by the Robbins-Monro method

$$t_{n+1} = t_n + a_n(Y_n - F(t_n)). \tag{1}$$

Here $\{a_n\}$ is a sequence of positive numbers (e.g. $a_n = a/n$). Under mild conditions, $\{t_n\}$ is known to converge to θ with probability one, and to be asymptotically normal; the asymptotic normal distribution has minimal variance if $a_n = d^{-1}/n$ where d is the derivative of F at θ.

The extension consists of allowing the F to depend on n. The authors then study in detail conditions sufficient for the convergence (Chapter 2), derive the asymptotic moments (Chapter 3) and the asymptotic distribution (Chapter 4). Based on these results they propose and compare various choices of the sequence $\{a_n\}$. A more specialized case is treated in Chapter 5 in which it is assumed that for a sequence $\{b_n\}$ of constants the sequence $\{b_n^{-1} \dot{F}_n\}$ is uniformly convergent. In this case a version of the approximation process is shown to be asymptotically efficient if and only if the observation errors are normally distributed. While the results mentioned up to now were restricted to the one-dimensional case where F_n are real functions on the real line R, the remaining chapters deal with the multi-dimensional case, concentrating on convergence conditions and various details useful in applications. Examples showing verification of assumptions in special cases and some open problems conclude the book.

The monograph extends in a significant way the field of possible applications. Some of the results (e.g. concerning the efficiency) strengthen the known results even if restricted to the original Robbins-Monro situation. There seems to be a misformulation in Assumption A1 where uniform boundedness of variances of the W_i's is later not sufficient and should be replaced, e.g. by that of second moments. And a final remark: The results on asymptotic distribution can be strengthened by observing that it depends, if $t_n \to \theta$ with probability one, only on local properties of F_n at θ, a fact first used in this connection by Hodges and Lehmann, *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** (1956) 96–104, University of California Press. Strengthening of these results will in turn influence the discussion on the relative merits of various versions of the approximation method, since this discussion is based not only on the properties but also on the extent to which the properties of the various versions were established. [The authors intend to publish a note on this.]