STOCHASTIC APPROXIMATION

by

Tze Leung Lai

Technical Report No. 2002-31
September 2002

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065

STOCHASTIC APPROXIMATION

by

Tze Leung Lai
Department of Statistics
Stanford University

Technical Report No. 2002-31
September 2002

This research was supported in part by National Science Foundation
grant DMS 0072523

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu

STOCHASTIC APPROXIMATION[1]

TZE LEUNG LAI
Stanford University

Abstract

Stochastic approximation, introduced by Robbins and Monro in 1951, has become an important and vibrant subject in optimization, control and signal processing. This paper reviews Robbins' contributions to stochastic approximation and gives an overview of several related developments.

[1]Research supported by the National Science Foundation and National Security Agency.
AMS 2000 subject classifications. Primary 62L20. Secondary 60F15, 93E23, 93E24.

Key words and phrases. Robbins-Monro and Kiefer-Wolfowitz schemes, stochastic Lyapunov functions, adaptive stochastic approximation, recursive stochastic algorithms, adaptive control.

1. Introduction. In 1951, Herbert Robbins and his student, Sutton Monro, founded the subject of stochastic approximation (SA). Not only did their celebrated paper attract immediate attention in the statistics literature, resulting in many important developments and advances in the 1950's and 1960's, but SA also eventually developed into an important area in systems control and optimization, to which Robbins made further seminal contributions in the late 1970's and early 1980's. Sections 2-5 review Robbins' work in SA and related adaptive control problems. They also provide a road map for certain portions of the extensive literature on these topics. SA has served as a prototype for the development and analysis of recursive algorithms for on-line estimation and control of stochastic systems, a brief account of which is given in Section 6 together with a discussion of its connections to Robbins' work. After five decades of continual development, SA still remains vibrant and Section 7 gives examples of new developments and some concluding remarks.

2. The Robbins-Monro and Kiefer-Wolfowitz schemes. The seminal paper of Robbins and Monro (1951) considers the problem of using successive approximations x_n to find the unique root θ of a regression function M in the regression model

$$y_n = M(x_n) + \epsilon_n \quad (n = 1, 2, \dots), \tag{1}$$

where the ϵ_n represent unobservable random errors. If one uses Newton's scheme $x_{n+1} = x_n - y_n/M'(x_n)$ as in the deterministic case $\epsilon_n \equiv 0$ (assuming M to be smooth with $M'(\theta) \neq 0$), then the presence of random errors ϵ_n in (1) implies that

$$x_{n+1} = x_n - M(x_n)/M'(x_n) - \epsilon_n/M'(x_n). \tag{2}$$

Hence, if x_n should converge to θ, so that $M(x_n) \to 0$ and $M'(x_n) \to M'(\theta)$, then (2) entails that $\epsilon_n \to 0$, which is not possible for many kinds of random errors (e.g., i.i.d. ϵ_n with positive variance). To average out the errors ϵ_n, Robbins and Monro (1951) proposed to use a recursive scheme of the form

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

where a_n are positive constants such that $\Sigma a_n^2 < \infty$ and $\Sigma a_n = \infty$, assuming that $M(x) > 0$ if $x > \theta$ and $M(x) < 0$ if $x < \theta$. The assumption $\Sigma a_n^2 < \infty$ ensures that $\Sigma a_n \epsilon_n$ converges in L_2 and a.s. for many stochastic models of random noise. Under certain regularity conditions, this in turn implies that $x_n - \theta$ converges in L_2 and a.s., and the assumption $\Sigma a_n = \infty$ then ensures that the limit of $x_n - \theta$ is 0; see Section 3 for a more detailed discussion.

A specific application of (3) mentioned by Robbins and Monro (1951) is estimation of the pth quantile θ_p of a distribution function F, for which $M(x) = F(x) - p$ and $y_n + p$ is a Bernoulli random variable with mean $F(x_n)$ in (1). Their interest in such applications might have been inspired by the "up-and-down" method of Dixon and Mood (1948), which requires parametric assumptions on F to estimate θ_p and for which the design levels x_n do not converge to θ_p (although the Dixon-Mood method aims at choosing design levels concentrated around the unknown θ_p).

There was substantial interest during the 1940's in the problem of finding the maximum (or minimum) of a regression function M and choosing design levels around the unknown optimum, beginning with Hotelling's (1941) paper on polynomial regression models. The basic ideas of the Robbins-Monro scheme can be readily modified to provide successive approximations for the minimum (or maximum) of a unimodal regression function, as was shown by Kiefer and Wolfowitz (1952) who introduced a recursive scheme of the form

$$x_{n+1} = x_n - a_n \Delta(x_n) \tag{4}$$

to find the minimum θ of M (or equivalently, the solution of $dM/dx = 0$). During the nth stage of the Kiefer-Wolfowitz scheme, observations y_n'' and y_n' are taken at the design levels $x_n'' = x_n + c_n$ and $x_n' = x_n - c_n$, respectively. In (4), $\Delta(x_n) = (y_n'' - y_n')/2c_n$, a_n and c_n are positive constants such that $c_n \to 0$, $\Sigma(a_n/c_n)^2 < \infty$ and $\Sigma a_n = \infty$.

3. Convergence analysis and stochastic Lyapunov functions. The approach used by Robbins and Monro (1951) in the convergence analysis of their scheme (3) is to convert (3) to a corresponding recursion for $E(x_{n+1} - \theta)^2$. This approach, which proves L_2 convergence of the Robbins-Monro scheme, was also used by Kiefer and Wolfowitz (1952) for the scheme (4), first showing that $E(x_{n+1} - \theta)^2$ converges to some limit and then using this to arrive at a contradiction if x_n does not converge to θ in probability. Their argument also requires that $\Sigma a_n c_n < \infty$.

Blum (1954) made use of the a.s. convergence of $\Sigma a_n \epsilon_n$ to prove first that $x_{n+1} + \Sigma_{j=1}^n a_j M(x_j)$ converges a.s. and then that x_n converges a.s. to θ for the Robbins-Monro scheme (3), under the following conditions on M that are weaker than those of Robbins and Monro:

$$|M(x)| \le c(|x - \theta| + 1) \text{ for all } x \text{ and some } c > 0, \tag{5}$$

$$\inf_{\epsilon \le |x-\theta| \le \epsilon^{-1}} \{M(x)(x - \theta)\} > 0 \text{ for all } 0 < \epsilon < 1. \tag{6}$$

He also established by a similar argument the a.s. convergence of (4) to θ under weaker assumptions than those of Kiefer and Wolfowitz (1952), in particular removing the assumption $\Sigma a_n c_n < \infty$. Dvoretzky (1956) subsequently proved the a.s. and L_2 convergence of a general class of recursive stochastic algorithms which include the Robbins-Monro and Kiefer-Wolfowitz schemes as special cases. This result is commonly called "Dvoretzky's approximation theorem".

In 1971, generalizing Gladyshev's (1965) approach to prove a.s. convergence of the Robbins-Monro scheme (3), Robbins and Siegmund proved a general convergence theorem which yields the a.s. part of Dvoretzky's approximation theorem as a corollary. Instead of using (3) or (4) to find a recursion for $E(x_{n+1} - \theta)^2$ as in Robbins and Monro (1951), this convergence theory starts with a corresponding recursive relation for $V_n := (x_{n+1} - \theta)^2$ and then takes conditional expectation with respect to $\mathcal{F}_{n-1}$, where $\mathcal{F}_t$ is the σ-field generated by $\{(x_i, y_i) : i \le t\}$. Note in this connection that x_{n+1} is $\mathcal{F}_n$-measurable, and it will be assumed that the ϵ_n in the regression model (1) form

a martingale difference sequence with respect to $\{\mathcal{F}_n\}$ such that $\sup_n E(\epsilon_n^2|\mathcal{F}_{n-1}) < \infty$ a.s. In particular, the recursion (3) in conjunction with (5) implies that

$$E(V_n|\mathcal{F}_{n-1}) \leq (1+\alpha_{n-1})V_{n-1} + \beta_{n-1} - \gamma_{n-1}, \tag{7}$$

where $\alpha_{n-1} = 2c^2a_n^2$, $\beta_{n-1} = a_n^2\{2c^2 + E(\epsilon_n^2|\mathcal{F}_{n-1})\}$ and $\gamma_{n-1} = 2a_nM(x_n)(x_n-\theta) \geq 0$.

Calling a sequence V_n that satisfies (7), in which V_i, α_i, β_i and γ_i are nonnegative $\mathcal{F}_i$-measurable random variables such that $\Sigma\alpha_i + \Sigma\beta_i < \infty$ a.s., a *nonnegative almost supermartingale*, Robbins and Siegmund (1971) proved the following convergence theorem:

$$V_n \text{ converges a.s. and } \sum \gamma_n < \infty \text{ a.s.} \tag{8}$$

Applying this result to the Robbins-Monro scheme (3) yields the a.s. convergence of $|x_n-\theta|(=V_{n-1}^{1/2})$ and of $\Sigma a_nM(x_n)(x_n-\theta)(=\Sigma\gamma_{n-1}/2)$. Since $\Sigma a_n = \infty$ and (6) holds, it then follows that $x_n \to \theta$ a.s. This approach is particularly useful for multidimensional SA schemes for which V_n takes the form $\|x_{n+1}-\theta\|^2$; see Robbins and Siegmund (1971) for an application of (8) to the infinite dimensional case where the x_n belong to a Hilbert space.

Without taking conditional expectation with respect to $\mathcal{F}_{n-1}$, Lai (1989) modified the Robbins-Siegmund theorem for the case where $\Sigma\beta_n$ need not be finite. Let $\{\epsilon_n, \mathcal{F}_n, n \geq 1\}$ be a martingale difference sequence such that $\sup_n E(\epsilon_n^2|\mathcal{F}_{n-1}) < \infty$ a.s. Let $\alpha_n, \beta_n, \gamma_n, V_n$ be nonnegative $\mathcal{F}_n$-measurable random variables such that $\Sigma\alpha_n < \infty$ a.s., and let w_n be $\mathcal{F}_n$-measurable. Suppose that for $n \geq 2$,

$$V_n \leq (1+\alpha_{n-1})V_{n-1} + \beta_n - \gamma_n + w_{n-1}\epsilon_n \quad \text{a.s.} \tag{9}$$

Then for every $\delta > 0$,

$$\max\left(V_n, \sum_{i=1}^n \gamma_i\right) = O\left(\sum_{i=1}^n \beta_i + \left(\sum_{i=1}^{n-1} w_i^2\right)^{\frac{1}{2}+\delta}\right) \text{ a.s.} \tag{10}$$

Moreover, V_n converges a.s. and $\Sigma E(\gamma_n|\mathcal{F}_{n-1}) < \infty$ a.s. on $\{\Sigma E(\beta_n|\mathcal{F}_{n-1}) < \infty\}$.

The convergence of V_n on $\{\Sigma E(\beta_n|\mathcal{F}_{n-1}) < \infty\}$ basically follows by the same argument as that leading to (8) upon taking conditional expectation with respect to $\mathcal{F}_{n-1}$ in (9). On the other hand, on $\{\Sigma E(\beta_n|\mathcal{F}_{n-1}) = \infty\}$, V_n need not converge and (10) provides a bound on the order of magnitude of V_n and $\Sigma_{i=1}^n\gamma_i$. The function $V_n = \|x_{n+1}-\theta\|^2$ used by Robbins and Siegmund (1971) in closely related to Lyapunov functions in the stability theory of ordinary differential equations (ODE); see Lai (1989). Analogous to Lyapunov functions in ODEs, the *stochastic Lyapunov function* V_n inherits an almost supermartingale structure from the dynamics of the original stochastic system. The idea behind the Lai's (1989) *extended stochastic Lyapunov functions* (9), whose applications will be discussed in Section 5, is to achieve greater flexibility by not insisting on the almost supermartingale property that guarantees convergence.

4. Asymptotic normality and adaptive stochastic approximation. Whereas the Lyapunov function approach (also called Lyapunov's second method) studies stability properties of an equilibrium of an ODE via a Lyapunov function, Lyapunov's first method uses the explicit solution of an ODE to study stability of an equilibrium. Beginning with Sacks (1958), a commonly used method to prove asymptotic normality of SA schemes is to first show its convergence to θ and then to replace the regression function (or its gradient vector) with a linear approximation in a small neighborhood of θ to develop an explicit non-recursive representation of x_n. This is, therefore, similar in spirit to Lyapunov's first method. In particular, Sacks (1958) showed that under certain regularity conditions an asymptotically optimal choice of a_n in the Robbins-Monro scheme (3) is $a_n \sim (n\beta)^{-1}$, for which $\sqrt{n}(x_n - \theta)$ has a limiting $N(0, \sigma^2/\beta^2)$ distribution, where $\beta = M'(\theta)$ (which also appears in Newton's scheme (2)) and $\sigma^2 = \lim_{n\to\infty} E(\epsilon_n^2|\mathcal{F}_{n-1})$. This result was obtained earlier under stronger assumptions by Chung (1954), Hodges and Lehmann (1956) and Burkholder (1956) who used recursions for $E(x_n - \theta)^k$ (similar to those of Robbins and Monro (1951) for $k = 2$) and the method of moments. By the late 1970's when Robbins returned to work on SA, there was a relatively complete theory on convergence and asymptotic normality of the Robbins-Monro and Kiefer-Wolfowitz schemes and their multivariate extensions; see e.g. the review papers by Schmetterer (1968) and Fabian (1971) and the monographs by Albert and Gardner (1967), Tsypkin (1971) and Nevel'son and Has'minskii (1973).

In 1976, Anderson and Taylor proposed a *certainty equivalence rule* for the "multiperiod control problem" in econometrics, which is concerned with choosing the inputs $x_1, \ldots, x_N$ sequentially in the linear regression model $y_i = \alpha + \beta x_i + \epsilon_i$ (with unknown parameters $\beta \neq 0$ and α, and i.i.d. random errors ϵ_i having mean 0 and variance σ^2) so that the outputs are as close as possible to a target value y^*. Assuming prior knowledge of bounds K_1 and K_2 such that $K_1 < \theta := (y^* - \alpha)/\beta < K_2$, the Anderson-Taylor rule is defined recursively by

$$x_{n+1} = K_1 \vee \{\widehat{\beta}_n^{-1}(y^* - \widehat{\alpha}_n) \wedge K_2\}, \quad n \geq 2, \tag{11}$$

where $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ are the least squares estimates of α and β at stage n. Based on the results of simulation studies, Anderson and Taylor (1976) conjectured that this certainty equivalence rule converges to θ a.s. and that $\sqrt{n}(x_n - \theta)$ has a limiting $N(0, \sigma^2/\beta^2)$ distribution. They also raised the question whether $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ are strongly consistent. Clearly, if the x_i should cluster around θ, then there would not be much information for estimating the slope β. There is, therefore, an apparent dilemma between the control objective of setting the design levels as close as possible to θ and the need for an informative design with sufficient dispersion to estimate β.

To resolve this dilemma, Lai and Robbins (1979) began by considering the case of known β. Replacing y_i by $y_i - y^*$, it can be assumed without loss of generality that $y^* = 0$ so that $y_i = \beta(x_i - \theta) + \epsilon_i$. Let $\bar{x}_n = n^{-1}\Sigma_{i=1}^n x_i$, $\bar{y}_n = n^{-1}\Sigma_{i=1}^n y_i$. With known β, the least squares certainty equivalence rule becomes $x_{n+1} = \bar{x}_n - \bar{y}_n/\beta$, which turns out to be equivalent to the SA scheme (3) with $a_n = (n\beta)^{-1}$. Since $\bar{x}_n - \bar{y}_n/\beta = \theta - \bar{\epsilon}_n/\beta$, $E(x_{n+1} - \theta)^2 = \sigma^2/(n\beta^2)$ for $n \geq 1$ and

therefore

$$E\left(\sum_{n=1}^{N} y_n^2\right) = \sum_{n=1}^{N} E\{\beta^2(x_n-\theta)^2+\epsilon_n^2\} = N\sigma^2 + \beta^2 \sum_{n=0}^{N-1} E(x_{n+1}-\theta)^2 = \sigma^2(N+\log N + O(1)).$$

Moreover,

$$\sum_{n=1}^{N}(x_n-\theta)^2 \sim (\sigma^2/\beta^2)\log N \text{ a.s. and } \sqrt{N}(x_N-\theta) \Rightarrow N(0,\sigma^2/\beta^2). \tag{12}$$

As shown by Lai and Robbins (1982a) who used dynamic programming after putting a prior distribution on θ, the optimal control rule when the ϵ_i are normal also has expected regret $\sigma^2 \log N + O(1)$, where $\beta^2\Sigma_{n=1}^N(x_n-\theta)^2$ $(= \Sigma_{n=1}^N(y_n-\epsilon_n)^2)$ is called the *regret* (due to ignorance of θ) of the design. Thus, for normally distributed errors, (12) shows that the least squares certainty equivalence rule when β is known yields both asymptotically minimal regret and efficient final estimate. The next step, therefore, is to try also to achieve (12) even when β is unknown.

An obvious way to modify the preceding rule for the case of unknown β is to use an estimate b_n to substitute for β either in the recursion $x_{n+1} = \bar{x}_n - \bar{y}_n/\beta$ or in the equivalent SA scheme $x_{n+1} = x_n - y_n/(n\beta)$. The equivalence between the two recursive schemes, however, no longer holds when β is replaced by b_n. The first recursion (with β replaced by b_n) is called *iterated least squares* and treated in Lai and Robbins (1982a, b). The second recursion, called *adaptive stochastic approximation*, is treated in Lai and Robbins (1979, 1981).

Under the assumptions (5) and (6) on the regression function M and also assuming that $M'(\theta) = \beta > 0$, Lai and Robbins (1979) consider SA schemes of the type

$$x_{n+1} = x_n - y_n/(nb_n), \tag{13}$$

where b_n is $\mathcal{F}_{n-1}$-measurable and $\lim_{n\to\infty} b_n = b > 0$ a.s. By representing x_n as a weighted sum of the i.i.d. random variables ϵ_i, they prove limit theorems on $x_N - \theta$ and $\Sigma_{n=1}^N(x_n-\theta)^2$. In particular, for $0 < b < 2\beta$, they generalize (12) to

$$\sum_{n=1}^{N}(x_n-\theta)^2 \sim (\sigma^2/\beta^2)f(b/\beta)\log N \text{ a.s. and } \sqrt{N}(x_N-\theta) \Rightarrow N(0,(\sigma^2/\beta^2)f(b/\beta)), \tag{14}$$

where $f(t) = 1/\{t(2-t)\}$ for $0 < t < 2$ and has a minimum value of 1 at $t = 1$. Lai and Robbins (1981) showed that by choosing b_n to be a truncated version of the least squares estimate in the adaptive SA scheme (13), one indeed has $b_n \to \beta$ a.s. and therefore (12) still holds for (13). Another way to achieve (12) for the adaptive SA scheme (13), used by Lai and Robbins (1979), is to take two points $x_n^{(1)} = x_n - c_n$ and $x_n^{(2)} = x_n + c_n$ at stage n, following an earlier idea of Venter (1967). Whereas Venter estimated β by $b_n = \zeta \vee \{\xi \wedge n\Sigma_{j=1}^{n-1}(y_j^{(2)} - y_j^{(1)})/(2c_j)\}$, with $0 < \zeta < \beta < \xi$ and $c_n \sim cn^{-\gamma}$ for some constants $c > 0$ and $\frac{1}{4} < \gamma < \frac{1}{2}$, Lai and Robbins (1979) proposed to use

$$b_n = \zeta_n \vee \left\{\xi_n \wedge \sum_{j=1}^{n-1} c_j\left(y_j^{(2)} - y_j^{(1)}\right) \Big/ \left(2\sum_{j=1}^{n-1} c_j^2\right)\right\}, \tag{15}$$

with $\zeta_n \to 0$ and $\xi_n \to \infty$ such that $\Sigma_1^\infty (i\zeta_i)^{-2} < \infty$ and $\Sigma_1^\infty (i\xi_i)^{-1} = \infty$, and $0 < c_j = O(j^{-\gamma})$ for some $\gamma > \frac{1}{4}$ but $\Sigma_1^\infty c_j^2 = \infty$, $\Sigma_1^n c_j^2 = o(\log n)$. The y_n in the SA scheme (13) is given by $y_n = (y_n^{(1)} + y_n^{(2)})/2$. Both (15) and Venter's estimate of β converge to β a.s., but the regret $\beta^2\{\Sigma_{n=1}^m (x_n^{(1)} - \theta)^2 + \Sigma_{n=1}^m (x_n^{(2)} - \theta)^2\}$ of Venter's design is of the order $2\beta^2 c^2 m^{1-2\gamma}/(1-2\gamma)$, whereas that of its modification by Lai and Robbins (1979) using (15) instead has order $\sigma^2 \log m$. Generalizations of these results to adaptive multivariate SA schemes have been provided by Wei (1987).

5. Strong consistency of least squares estimates and associated adaptive control rules. A key step in the analysis of adaptive SA and iterated least squares procedures in Lai and Robbins (1981, 1982a) is to establish strong consistency of the least squares estimate

$$\widehat{\beta}_n = \left\{ \sum_{i=1}^n (x_i - \bar{x}_n) y_i \right\} \Big/ \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

when the x_i are sequentially determined random variables. Also encountering such consistency issues in the analysis of their certainty equivalence control rule (11), Anderson and Taylor (1979) considered the more general multiple regression model $y_n = \theta^T \psi_n + \epsilon_n$, and showed for the least squares estimate $\widehat{\theta}_n = (\Sigma_{i=1}^n \psi_i \psi_i^T)^{-1} \Sigma_{i=1}^n \psi_i y_i$ that

$$(16)\quad \widehat{\theta}_n \to \theta \text{ a.s. if } \lambda_{\min}\left(\sum_{i=1}^n \psi_i \psi_i^T \right) \to \infty \quad \text{and} \quad \lambda_{\max}\left(\sum_{i=1}^n \psi_i \psi_i^T \right) = O\left(\lambda_{\min}\left(\sum_{i=1}^n \psi_i \psi_i^T \right) \right) \text{a.s.}$$

When specialized to the simple linear model $y_n = \alpha + \beta x_n + \epsilon_n$, (16) yields the strong consistency of $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ if $\liminf \Sigma_{i=1}^n (x_i - \bar{x}_n)^2/n > 0$ with probability 1. Lai and Robbins (1981) proved strong consistency of $\widehat{\beta}_n$ under the substantially weaker condition $\Sigma_{i=1}^n (x_i - \bar{x}_n)^2 / \log n \to \infty$ a.s. and gave a counter-example in which x_i is $\mathcal{F}_{i-1}$-measurable and $\Sigma_{i=1}^n (x_i - \bar{x}_n)^2 / \log n$ converges a.s. to some constant $c > 0$ but in which $\widehat{\beta}_n \to \beta - c^{-1}$ a.s. Since $\Sigma_{i=1}^n (x_i - \theta)^2$ is desired to have the order $(\sigma^2/\beta^2) \log n$ in view of (12), this counter-example also shows that proving strong consistency of $\widehat{\beta}_n$ in an adaptive SA environment requires delicate arguments that involve monitoring how $\Sigma_1^n (x_i - \theta)^2$ and $\bar{x}_n - \theta$ evolve over time. Theorem 8 of Lai and Robbins (1981) provides a set of sufficient conditions on $\Sigma_1^n (x_i - \theta)^2$ and $\bar{x}_n - \theta$ that can be verified for adaptive SA schemes, thereby proving the desired strong consistency of $\widehat{\beta}_n$ and establishing (12) for these schemes.

Lai and Wei (1982) subsequently generalized the Lai-Robbins condition $\Sigma_{i=1}^n (x_i - \bar{x}_n)^2 / \log n \to \infty$ a.s. to multiple regression models, giving the following improvement of (16):

$$(17)\qquad \widehat{\theta}_n \to \theta \text{ a.s. on } \left\{ \lambda_{\min}\left(\sum_{i=1}^n \psi_i \psi_i^T \right) \Big/ \log \lambda_{\max}\left(\sum_{i=1}^n \psi_i \psi_i^T \right) \to \infty \right\}.$$

Whereas Lai and Robbins (1981) assumed the ϵ_n to be i.i.d. with mean 0 and finite variance, Lai and Wei (1982) established (17) under the assumption that $\{\epsilon_n, \mathcal{F}_n, n \geq 1\}$ is a martingale difference sequence such that $\sup_n E(|\epsilon_n|^{2+\delta} | \mathcal{F}_{n-1}) < \infty$ a.s. for some $\delta > 0$. When the ψ_i

are nonrandom, Lai, Robbins and Wei (1979) proved strong consistency of $\widehat{\theta}_n$ under the minimal condition $\lambda_{\min}(\Sigma_{i=1}^n \psi_i \psi_i^T) \to \infty$, assuming only that $\{\epsilon_n\}$ is a "convergence system" (i.e., $\Sigma_1^\infty c_i \epsilon_i$ converges a.s. for all nonrandom sequences $\{c_n\}$ such that $\Sigma_1^\infty c_i^2 < \infty$.) For nonrandom ψ_i, the jth component $\widetilde{\theta}_{nj}$ of $\widehat{\theta}_n - \theta$ can be expressed as $(\Sigma_{i=1}^n a_{ni}\epsilon_i)/\Sigma_{i=1}^n a_{ni}^2$, where the a_{ni} are nonrandom constants (depending also on j) such that $\Sigma_{i=1}^n a_{ni}a_{mi} = \Sigma_{i=1}^m a_{mi}^2$ for $n \geq m$. The a_{ni} are no longer nonrandom if the ψ_i are sequentially determined random variables, and this is why stronger assumptions as in (17) are needed for the strong consistency of $\widehat{\theta}_n$.

Like SA schemes that are defined by recursions, the least squares estimate $\widehat{\theta}_n$ also has a recursive representation of the form

$$\widehat{\theta}_n = \widehat{\theta}_{n-1} + P_n \psi_n (y_n - \widehat{\theta}_{n-1}^T \psi_n), \tag{18a}$$

$$P_n = P_{n-1} - P_{n-1}\psi_n \psi_n^T P_{n-1}/(1 + \psi_n^T P_{n-1} \psi_n), \tag{18b}$$

where $P_n = (\Sigma_{i=1}^n \psi_i \psi_i^T)^{-1}$. The recursion (18) leads to a corresponding recursive inequality (9) for $V_n := (\widehat{\theta}_n - \theta)^T P_n^{-1} (\widehat{\theta}_n - \theta)$ with $\alpha_n = 0$, $\beta_n = \psi_n^T P_n \psi_n \epsilon_n^2$, $\gamma_n = \{(\widehat{\theta}_{n-1} - \theta)^T \psi_n\}^2 (1 - \psi_n^T P_n \psi_n) \geq 0$ and $w_{n-1} = 2\{(\widehat{\theta}_{n-1} - \theta)^T \psi_n\}(1 - \psi_n^T P_n \psi_n)$. Therefore, assuming that $\{\epsilon_n, \mathcal{F}_n, n \geq 1\}$ is a martingale difference sequence with $\sup_n E(\epsilon_n^2 | \mathcal{F}_{n-1}) < \infty$ a.s., it follows from (10) that $\max(V_n, \Sigma_{i=1}^n \gamma_i) = O(\Sigma_{i=1}^n \psi_i^T P_i \psi_i \epsilon_i^2) + O((\Sigma_{i=1}^n \gamma_i)^{2/3})$ a.s. If we assume the stronger moment condition $\sup_n E(|\epsilon_n|^{2+\delta} | \mathcal{F}_{n-1}) < \infty$ a.s. for some $\delta > 0$, then $\Sigma_{i=1}^n \psi_i^T P_i \psi_i \epsilon_i^2 = O(\Sigma_{i=1}^n \psi_i^T P_i \psi_i) = O(\log \lambda_{\max}(\Sigma_{i=1}^n \psi_i \psi_i^T))$ a.s. by Lemma 2 of Lai and Wei (1982). Therefore

$$\max\left(V_n, \sum_{t=1}^n \gamma_t\right) = O\left(\log \lambda_{\max}\left(\sum_{i=1}^n \psi_i \psi_i^T\right)\right) \text{ a.s.} \tag{19}$$

Since $V_n \geq \|\widehat{\theta}_n - \theta\|^2 \lambda_{\min}(\Sigma_{i=1}^n \psi_i \psi_i^T)$, (19) implies (17). Lai and Wei (1982) essentially proved the strong consistency of $\widehat{\theta}_n$ for stochastic regressors ψ_n that are $\mathcal{F}_{n-1}$-measurable by this argument, for which Lai (1989) later provided a more general framework extending the Robbins-Siegmund theorem (8). Moreover, by the definition of γ_n, (19) also yields

$$\sum_{t=1}^n \left(\widehat{\theta}_{t-1}^T \psi_t - \theta^T \psi_t\right)^2 I_{\{\psi_t^T P_t \psi_t \leq \delta\}} = O\left(\log \lambda_{\max}\left(\sum_{i=1}^n \psi_i \psi_i^T\right)\right) \text{ a.s.} \tag{20}$$

for every $0 < \delta < 1$. Note that $\theta^T \psi_t$ is the minimum-variance $\mathcal{F}_{t-1}$-measurable predictor of y_t when θ is known and that $\widehat{\theta}_{t-1}^T \psi_t$ is the adaptive predictor of y_t using the least squares estimate $\widehat{\theta}_{t-1}$ to replace θ. Therefore (20) gives an asymptotic bound on the cumulative squared difference between these adaptive and optimal predictors, without consistency assumptions on the least squares estimates $\widehat{\theta}_t$.

For the special case of the simple linear regression model $y_n = \alpha + \beta x_n + \epsilon_n$ considered by Anderson and Taylor (1976), $\widehat{\theta}_t = (\widehat{\alpha}_t, \widehat{\beta}_t)^T$, $\psi_t = (1, x_t)^T$ and the certainty equivalence rule (11) (with $y^* = 0$) is tantamount to $\widehat{\theta}_n^T \psi_{n+1} = 0$, modulo truncation of x_{n+1} by K_1 and K_2. Strictly

speaking, (11) is undefined when $\widehat{\beta}_n = 0$, but this is a zero-probability event in the simulation studies of Anderson and Taylor involving continuous (normal) ϵ_n. Lai and Robbins (1982b) showed that (11) does not converge a.s. to θ by exhibiting an event which has positive probability and on which x_n gets stuck at one of the end-points K_1, K_2 for $n \geq 2$. They also showed that for a recursive scheme of the type $x_{n+1} = \bar{x}_n - \bar{y}_n/b_n$, $x_n \to \theta$ a.s. if $b_n \to b$ a.s., where b has the same sign as β; moreover, (14) still holds for this scheme if $0 < b/\beta < 2$. The next step, therefore, was to modify (11) so that the design has enough information for $\widehat{\beta}_n$ to estimate β consistently. Instead of assuming *a priori* bounds on θ, Lai and Robbins (1982a) assumed that the sign of β is known (say, to be positive) and that there are *a priori* bounds B_1, B_2 such that $0 < B_1 < \beta < B_2$. Using $\widehat{\sigma}_n^2 = n^{-1}\Sigma_{i=1}^n (y_i - \bar{y}_n)^2$ to estimate σ^2 and setting $b_n = b_{n-1}$ if $\Sigma_{i=1}^n (x_i - \bar{x}_n)^2 < (\widehat{\sigma}_n^2/2B_2^2)\log n$, and otherwise setting $b_n = B_1 \vee (\widehat{\beta}_n \wedge B_2)$, they showed that $b_n \to \beta$ a.s. for the modified least squares certainty equivalence rule $x_{n+1} = \bar{x}_n - \bar{y}_n/b_n$ which, therefore, satisfies (12).

Adaptive control has been an active area of research in the engineering literature. Unlike the static regression models considered above, control engineering involves time series models and stochastic dynamical systems, for which the input u_n at stage n affects not only the output y_{n+1} at stage $n+1$ but also the future dynamics of the system, as in the case of ARX model (autoregressive model with exogenous inputs) $A(q^{-1})y_n = B(q^{-1})u_{n-1} + \epsilon_n$, where $A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_p q^{-p}$, $B(q^{-1}) = b_1 + \ldots + b_r q^{-(r-1)}$ and q^{-1} is the unit delay operator (defined by $q^{-1}x_n = x_{n-1}$). Rewriting the ARX model as a stochastic regression model $y_n = \theta^T \psi_n + \epsilon_n$ with $\theta = (-a_1, \ldots, -a_p, b_1, \ldots, b_r)^T$ and $\psi_n = (y_{n-1}, \ldots, y_{n-p}, u_{n-1}, \ldots, u_{n-r})^T$, we can apply (19) and (20) to analyze the least squares estimates $\widehat{\theta}_t$ and the adaptive least squares predictors $\widehat{\theta}_{t-1}^T \psi_t$. The adaptive control problem for the ARX model is to choose the inputs u_t sequentially so that the outputs y_t are as close as possible to some bounded reference signals y_t^*. Assuming that $b_1 \neq 0$, the minimum-variance controller u_{t-1} is given by $\theta^T \psi_t = y_t^*$ when θ is known, yielding outputs $y_t = y_t^* + \epsilon_t$. When θ is unknown, we define the *regret* R_N (at stage n) of a control rule by $R_N = \Sigma_{t=1}^N (y_t - y_t^* - \epsilon_t)^2$, as in Lai and Robbins (1979). Beginning with the landmark paper of Åström and Wittenmark (1973), there was considerable effort for nearly two decades to address the problem whether using the least squares estimate $\widehat{\theta}_{t-1}$ to substitute for the unknown θ in $\theta^T \psi_t = y_t^*$ can "self-tune" the ARX system in the sense that

$$R_n/n \to 0 \text{ a.s. and } \sum_{i=1}^n (u_i^2 + y_i^2) = O(n) \text{ a.s.} \tag{21}$$

(The first property in (21) is equivalent to minimizing the long-run average squared tracking error, while the second property says that the closed-loop system is globally stable.) An essentially affirmative answer was provided by Guo and Chen (1991) who slightly modified the least squares estimate $\widehat{\theta}_i$ by truncating $\widehat{b}_{1,i}$ when it falls below $1/\log r_{i-1}$, where $r_t = \Sigma_{j=1}^t \|\psi_j\|^2$ and $\widehat{b}_{1,i}$ is the component of $\widehat{\theta}_i$ estimating b_1. Important tools in their analysis are (20) and the Bellman-Gronwall inequality that uses the dynamics of y_t to bound $\|\psi_t\|^2$.

6. Recursive stochastic algorithms in system identification and control. The

ARX model corresponds to the so-called "white noise" case in the control systems literature, which usually deals with the more general ARMAX model (in which MA stands for "moving average" disturbances or "colored noise"):

$$A(q^{-1})y_n = B(q^{-1})u_{n-1} + C(q^{-1})\epsilon_n, \tag{22}$$

where $C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_k q^{-k}$, the polynomials $A(z), B(z)$ and $C(z)$ are relatively prime and $\{\epsilon_n, \mathcal{F}_n, n \geq 1\}$ is a martingale difference sequence satisfying $\sup_{n\geq 1} E(|\epsilon_n|^{2+\delta}|\mathcal{F}_{n-1}) < \infty$ a.s. for some $\delta > 0$. . Fuchs (1982) proposed to estimate the parameter vector $\theta = (-a_1, \ldots, -a_p, b_1, \ldots, b_r, c_1, \ldots, c_k)^T$ recursively by the *stochastic gradient* (SG) algorithm

$$\theta_n = \theta_{n-1} + (\alpha/r_n)\phi_n(y_n - \theta_{n-1}^T\phi_n), \quad r_n = r_{n-1} + \|\phi_n\|^2, \tag{23}$$

where $\phi_n = (y_{n-1}, \ldots, y_{n-p}, u_{n-r}, \ldots, u_{n-q}, \widehat{\epsilon}_{n-1}, \ldots, \widehat{\epsilon}_{n-k})^T$ and $\widehat{\epsilon}_t = y_t - \theta_{t-1}^T\phi_t$. Note that ϕ_n is a surrogate for the $\mathcal{F}_{n-1}$-measurable regressor $\psi_n = (y_{n-1}, \ldots, y_{n-p}, u_{n-1}, \ldots, u_{n-r}, \epsilon_{n-1}, \ldots, \epsilon_{n-k})^T$. A similar SG algorithm was introduced earlier by Goodwin, Ramadge and Caines (1981) after reparameterizing the model as $C(q^{-1})E(y_{n+1}|\mathcal{F}_n) = G(q^{-1})y_n + B(q^{-1})u_n$, where the polynomial division algorithm is applied to yield $C(z) = A(z) + zG(z)$. Assume the stability and positive real conditions

$$C(z) \text{ has all zeros outside the unit circle and } \min_{|z|=1} \mathrm{Re}\{C(z) - \alpha/2\} > 0. \tag{24}$$

By continuity and compactness, (24) also holds with α replaced by $\alpha + \rho$ for some $\rho > 0$. Let $\eta_n = \widehat{\epsilon}_n - \epsilon_n = \theta^T\psi_n - \theta_{n-1}^T\phi_n$, $h_n = \{C(q^{-1}) - \frac{1}{2}(\alpha + \rho)\}\eta_n$. Then there exists $K > 0$ such that $S_n := 2\rho\Sigma_{i=1}^n h_i\eta_i + K \geq 0$ for all n. Letting $V_n = \|\theta_n - \theta\|^2 + S_n/r_n$, the analysis of Goodwin et al. (1981) and Fuchs (1982) shows, among other things, that (7) holds with $\alpha_{n-1} = 0$, $\beta_{n-1} = (\alpha/r_n)^2\|\phi_n\|^2 E(\epsilon_n^2|\mathcal{F}_{n-1})$ and $\gamma_{n-1} = \rho\alpha\eta_n^2/r_n$. Therefore it follows from (8) that V_n converges a.s. and

$$\sum_{n=1}^{\infty}(\theta_{n-1}^T\phi_n - \theta^T\psi_n)^2/r_n < \infty \text{ a.s.} \tag{25}$$

Making use of (25) and the dynamics of y_t under a stability assumption on $B(z)$ (but not on $A(z)$), Fuchs (1982) showed that the certainty equivalence rule $\theta_{n-1}^T\phi_n = y_n^*$ satisfies that the self-tuning property (21), while Goodwin et al. (1981) obtained the same result for a reparameterized version of the rule.

The SG algorithm (23) is reminiscent of most multivariate extensions of the Robbins-Monro and Kiefer-Wolfowitz schemes in the statistics literature, taking the form $X_{n+1} = X_n - a_n Y_n$ where the X_i and Y_i are $d \times 1$ vectors and a_n is a scalar, as in (3) or (4) in the case $d = 1$; see Nevel'son and Has'minskii (1973). A common feature of these algorithms is that they have scalar gain sequences (α/r_n or a_n), which enable them to be analyzed via stochastic Lyapunov functions of the form $\|\theta_n - \theta\|^2$ or $\|X_n - \theta\|^2$. An exception to using scalar gains is the extension of the

adaptive SA scheme (13) to the multivariate case (with $M : \mathbf{R}^d \to \mathbf{R}^d$) by Wei (1987) who used a matrix-gain algorithm $X_{n+1} = X_n - (nB_n)^{-1}Y_n$, in which B_n is an estimate of the Jacobian matrix $\partial M/\partial x|_{x=\theta}$ that is assumed to be nonsingular. Matrix-gain SA schemes are also used in extending the Robbins-Monro scheme to derive recursive estimates of multidimensional parameters that are asymptotically as efficient as the maximum likelihood estimates in certain parametric models; see Chapters 8 and 9 of Nevel'son and Has'minskii (1973).

Returning to the ARMAX model (22), a matrix-gain alternative to the SG algorithm is the *extended least squares* (ELS) algorithm given by (18) with ϕ_t replacing the unobservable ψ_t. Note that SG with $\alpha = 1$ essentially replaces the matrix $P_n^{-1} = \Sigma_1^n \phi_i \phi_i^T$ in ELS by its trace r_n. Defining the $\widehat{\epsilon}_t$ (to form the ϕ_i) by $y_t - \theta_t^T \phi_t$ (instead of $y_t - \theta_{t-1}^T \phi_t$), Lai and Wei (1986) proved that (20) (with ψ_t replaced by ϕ_t) still holds for the ELS algorithm under the stability and positive real conditions

(26) $\qquad C(z)$ has zeros outside the unit circle and $\min_{|z|=1} \mathrm{Re}\{1/C(z) - 1/2\} > 0.$

Their argument essentially shows that (9) holds for the extended stochastic Lyapunov function $V_n = (\widehat{\theta}_n - \theta)^T P_n^{-1} (\widehat{\theta}_n - \theta)$ so that (10) can be applied. For the ELS algorithm $\widehat{\theta}_t$, (20) with ϕ_t in place of ψ_t was used by Guo and Chen (1991) to prove that (21) holds for the certainty equivalence rule $\widehat{\theta}_{n-1}^T \phi_n = y_n^*$ in the ARMAX model (22).

The self-tuning property (21) is considerably weaker than the order $\sigma^2 \log n$ for the regret established by Lai and Robbins (1979, 1981, 1982a,b) for the multiperiod control problem in simple linear regression models. Although the ELS algorithm has been sometimes called "approximate maximum likelihood" (AML), its statistical properties are actually unrelated to those of the maximum likelihood estimator (MLE) even when the ϵ_i are normal. Linear approximations to the estimating equations defining the MLE in some neighborhood of the true parameter lead to another recursive algorithm called RML2. Lai and Ying (1991) showed how parallel recursive algorithms can be used to develop an adaptive control rule with regret no larger than $\{C+o(1)\} \log n$, where the constant C can be shown to be the smallest possible under much more restrictive assumptions. These parallel recursive algorithms involve an SG component to stabilize the system so that occasional blocks of white-noise inputs can be used to generate enough information for consistent estimation of θ, and an RML2 component that is constrained to lie within diminishing neighborhoods of the auxiliary consistent estimates.

Closely related to the stochastic Lyapunov functions and extended stochastic Lyapunov functions described above is the ODE approach introduced by Ljung (1977) for the convergence analysis of SA and other recursive stochastic algorithms. To fix the ideas, consider the multivariate Robbins-Monro scheme $X_{n+1} = X_n - a_n\{M(X_n) + \epsilon_n\}$. Let $t_n = \Sigma_1^n a_i$ ($\to \infty$ as $n \to \infty$), $X(t) = X_{n+1}$ at $t = t_n$, and define $X(t)$ by linear interpolation when $t_n < t < t_{n+1}$. Noting that $t_n - t_{n-1} = a_n \to 0$ and that $\sup_{m \geq n} \|\Sigma_{i=n}^m a_i \epsilon_i\| \to 0$ as $n \to \infty$, the SA scheme can be approximated by the ODE $dX/dt = -M(X(t))$, suggesting that $X_n \to \theta$ if θ is a globally asymptotically stable equilibrium of the ODE. This argument can be made rigorous by using compactness arguments under additional

assumptions such as continuity of M and a.s. boundedness of X_n. For more general recursive stochastic algorithms, different sets of technical assumptions have been developed to ensure the validity of the ODE approach; see Kushner and Clark (1978), Ljung and Söderström(1983) and Benveniste, Métivier and Priouret (1987). However, the ODE approach is often used heuristically to obtain a quick understanding of the dynamics of the algorithm, especially how it is coupled to the dynamics of the underlying system (e.g. the ELS algorithm in the ARMAX model above). Although these technical assumptions are often either too stringent or too difficult to check, one often bypasses them and tries to find useful clues from the limiting ODE. A case in point is the least squares recursion (18) in the stochastic regression model $y_n = \theta^T \psi_n + \epsilon_n$, as pointed out in Lai (1989, page 213). Here Ljung's ODE approach (with $a_n = n^{-1}$) can be justified under the overly restrictive "persistent excitation" condition $n^{-1}\Sigma_1^n \psi_i \psi_i^T (= n^{-1} P_n^{-1}) \to Q$ a.s. for some positive definite matrix Q, and stability analysis of the limiting ODE involves the Lyapunov function $V(\theta(t), P(t)) = (\theta(t) - \theta)^T P^{-1}(t)(\theta(t) - \theta)$. Instead of working with the Lyapunov function of the limiting ODE, an obvious alternative is to work directly with its analog $V_n = (\widehat{\theta}_n - \theta)^T P_n^{-1} (\widehat{\theta}_n - \theta)$, which is an "extended stochastic Lyapunov function" that need not even converge. This yields the asymptotic behavior (19) not only for V_n but also for $\Sigma_1^n (\widehat{\theta}_{t-1}^T \psi_t - \theta^T \psi_t)^2 (1 - \psi_t^T P_t \psi_t)$ that has played an important role in the solution of the Åström-Wittenmark problem.

7. Some recent developments and concluding remarks. The subject SA founded by Robbins and Monro fifty years ago is still showing vigor and vibrancy. As evidenced by the relatively recent survey articles and monographs by Ruppert (1991), Yin (1991), Ljung, Pflug and Walk (1992) and Duflo (1997), the subject has been growing steadily in different directions over these decades. There is now a rich arsenal of theories and techniques which have found and been inspired by new applications from other fields. For example, interactions of SA with neural network modeling, neuro-dynamic programming, reinforcement learning, perturbation analysis of discrete-event dynamic systems, simulated annealing, spatial statistics and maximum likelihood estimation based on incomplete data have led to advances in these areas and in SA; see Spall (1992, 2000), Spall and Cristion (1994), Bertsekas and Tsitsiklis (1996), Tsitsiklis and Van Roy (1997), Chong and Ramadge (1993), Gelfand and Mitter (1991), Yin (1999), Younes (1989), Moyeed and Baddely (1991), Gu and Zhu (2001), Gu and Kong (1998) and Delyon, Lavielle and Moulines (1999).

An attractive feature of SA and other recursive stochastic algorithms is that they allow fast updating at each instant when new data arrive and therefore can be used to support decisions "on-line", i.e., during the operation of the system. Moreover, they can be asymptotically as efficient as their "off-line" counterparts that require much greater computational complexity and whose memory requirements grow to ∞ with the sample size. However, the loss of information in using only the previous estimate $\widehat{\theta}_{t-1}$ and the current observation y_t to form the current estimate $\widehat{\theta}_t$ of θ often results in unsatisfactory finite-sample performance except for linear problems in which off-line estimates have simple recursive forms such as (18). For example, in applying the Robbins-Monro scheme (3) to estimate the pth quantile θ_p of a distribution function F, the nonlinearity of F and the binary nature of the response suggest that there is substantial loss of information to use an

SA scheme instead of regression modeling with all the data collected so far. The simulation study of Cochran and Davis (1965) involving sample size $n = 24$ for estimating the median lethal dose ($p = \frac{1}{2}$) confirms this. In fact, it seems too ambitious to use a nonparametric approach to model F for samples of such size, and initializing with a parametric model as in Wu (1985) for moderate sample sizes before switching to SA is a better idea. Another way to improve the finite-sample performance of SA schemes, proposed by Kesten (1958), is to keep the gains a_n constant until $x_n - x_{n-1}$ changes sign. Alternatively, using parallel recursive algorithms as in Yin (1991) and Lai and Ying (1991) can take advantage of parallel computation to incorporate more features of efficient off-line procedures into an on-line estimation or control scheme. As pointed out in Section 7 of Gu and Zhu (2001), it is often important to "tweak" the Robbins-Monro procedure appropriately for the application being considered in order to obtain satisfactory finite-sample performance besides attaining the desired asymptotic properties.

Robbins (1952) mentioned the two-armed bandit problem (introduced in that paper) and SA as two new problems in sequential design of experiments, "different from those usually met in statistical literature." Whereas the Kiefer-Wolfowitz scheme (4) chooses design points sequentially from $\mathbf{R}$ (continuous action space) in order to maximize $E(\Sigma_{i=1}^N y_i)$, with y_i given by the regression model (1), the k-armed bandit problem is concerned with sampling $y_1, \ldots, y_N$ sequentially from k populations so that $E(\Sigma_{i=1}^N y_i)$ is maximized, where the conditional distribution of y_i given that it is sampled from population Π_j is the same as that of Π_j. Section 6 of Lai (2001) gives a survey of the multi-armed bandit (MAB) problem. Like SA, MAB has undergone many far-reaching generalizations and has become another important topic (concerning finite action spaces) in the field of stochastic adaptive control. Through SA and MAB, Robbins has indeed greatly broadened the scope of sequential experimentation and pushed it to new heights and in new directions.

REFERENCES

ALBERT, A.E. and GARDNER, L.A. (1967). *Stochastic Approximation and Nonlinear Regression.* MIT Press, Cambridge, MA.

ANDERSON, T.W. and TAYLOR, J.B. (1976). Some experimental results on the statistical properties of least squares estimates in control problems. *Econometrica* **44**, 1289-1302.

ANDERSON, T.W. and TAYLOR, J.B. (1979). Strong consistency of least squares estimates in dynamic models. *Ann. Statist.* **7**, 484-489.

ÅSTRÖM, K.J. and WITTENMARK, B. (1973). On self-tuning regulators. *Automatica* **9**, 195-199.

BENVENISTE, A., MÉTIVIER, M. and PRIOURET, P. (1987). *Adaptive Algorithms and Stochastic Approximations.* Springer-Verlag, New York.

BERTSEKAS, D.P. and TSITSIKLIS, J.N. (1996). *Neuro-dynamic Programming.* Athena, Belmont, MA.

BLUM, J.R. (1954). Approximation methods which converge with probability one. *Ann. Math. Statist.* **25**, 382-386.

BURKHOLDER, D.L. (1956). On a class of stochastic approximation processes. *Ann. Math. Statist.* **27**, 1044-1059.

CHONG, E.K.P. and RAMADGE, P.J. (1993). Optimization of queues using infinitesimal perturbation analysis-based stochastic algorithm with general update times. *SIAM J. Contr. Optimiz.* **31**, 698-732.

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

COCHRAN, W.G. and DAVIS, M. (1965). The Robbins-Monro method for estimating the median lethal dose. *J. Roy. Statist. Soc. Ser. B* **27**, 28-44.

DELYON, B., LAVIELLE, E. and MOULINES, E. (1999). Convergence of a stochastic approximation version of EM algorithm. *Ann. Statist.* **27**, 94-128.

DIXON, W.J. and MOOD, A.M. (1948). A method of obtaining and analyzing sensitivity data. *J. Amer. Statist. Assoc.* **43**, 109-126.

DUFLO, M. (1997). *Random Iterative Models.* Springer-Verlag, New York.

DVORETZKY, A. (1956). On stochastic approximation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1**, 39-56. Univ. California Press.

FABIAN, V. (1971). Stochastic approximation. In *Optimizing Methods in Statistics* (J.S. Rustagi, ed.), 439-470. Academic Press, New York.

FUCHS, J.J. (1982). Indirect stochastic adaptive control: the general delay-colored noise case. *IEEE Trans. Automat. Contr.* **27**, 470-472.

GELFAND, S.B. and MITTER, S.K. (1991). Recursive stochastic algorithms for global optimization in $\mathbf{R}^d$. *SIAM J. Contr. Optimiz.* **29**, 999-1018.

GLADYSHEV, E.G. (1965). On stochastic approximation. *Theory Probab. Appl.* **10**, 275-278.

GOODWIN, G.C., RAMADGE, P.J. and CAINES, P.E. (1981). Discrete time stochastic adaptive control. *SIAM J. Contr. Optimiz.* **19**, 829-853.

GU, M.G. and KONG, F.H. (1998). A stochastic approximation algorithm with Markov chain Monte-Carlo method for incomplete data estimation problems. *Proc. Natl. Acad. Sci. USA* **95**, 7270-7274.

GU, M.G. and ZHU, H.T. (2001). Maximum likelihood estimation for spatial models by Markov chain Monte Carlo stochastic approximation. *J. Roy. Statist. Soc. Ser. B* **63**, 329-355.

GUO, L. and CHEN, H.F. (1991). The Åström–Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers. *IEEE Trans. Automat. Contr.* **36**, 802-812.

HODGES, J.L. and LEHMANN, E.L. (1956). Two approximations to the Robbins-Monro process. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1**, 95-104. Univ. California Press.

HOTELLING, H. (1941). Experimental determination of the maximum of a function. *Ann. Math. Statist.* **12**, 20-45.

KESTEN, H. (1958). Accelerated stochastic approximation. *Ann. Math. Statist.* **29**, 41-59.

KIEFER, J. and WOLFOWITZ, J. (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.* **23**, 462-466.

KUSHNER, H.J. and CLARK, D.S. (1978). *Stochastic Approximation for Constrained and Unconstrained Systems.* Springer-Verlag, New York.

LAI, T.L. (1989). Extended stochastic Lyapunov functions and recursive algorithms in linear stochastic systems. In *Stochastic Differential Systems* (N. Christopeit, K. Helmes and N. Kolmann, eds.), 206-220. Springer-Verlag, New York.

LAI, T.L. (2001). Sequantial analysis: Some classical problems and new challenges (with discussion). *Statistica Sinica* **11**, 303-408.

LAI, T.L. and ROBBINS, H. (1979). Adaptive design and stochastic approximation. *Ann. Statist.* **7**, 1196-1221.

LAI, T.L. and ROBBINS, H. (1981). Consistency and asymptotic efficiency of slope estimates in stochastic approximation schemes. *Z. Wahrsch. Verw. Gebiete* **56**, 329-360.

LAI, T.L. and ROBBINS, H. (1982a). Adaptive design and the multiperiod control problem. In *Statistical Decision Theory and Related Topics III* **2**, (S.S. Gupta, ed.), 103-120. Academic Press, New York.

LAI, T.L. and ROBBINS, H. (1982b). Iterated least squares in multiperiod control. *Adv. Appl. Math.* **3**, 50-73.

LAI, T.L., ROBBINS, H. and WEI, C.Z. (1979). Strong consistency of least squares estimates in multiple regression II. *J. Multivariate Anal.* **9**, 343-361.

LAI, T.L. and WEI, C.Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Ann. Statist.* **10**, 154-166.

LAI, T.L. and WEI, C.Z. (1986). Extended least squares and their applications to adaptive control and prediction in linear systems. *IEEE Trans. Automat. Contr.* **31**, 898-906.

LAI, T.L. and YING, Z. (1991). Parallel recursive algorithms in asymptotically efficient adaptive control of linear stochastic systems. *SIAM J. Contr. Optimiz.* **29**, 1091-1127.

LJUNG, L. (1977). Analysis of recursive stochastic algorithms. *IEEE Trans. Automat. Contr.* **22**, 551-575.

LJUNG. L., PFLUG, G. and WALK, H. (1992). *Stochastic Approximation and Optimization of Random Systems.* Birkhäuser, Boston.

LJUNG. L. and SÖDERSTÖM, T. (1983). *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA.

MOYEED, R.A. and BADDELEY, A.J. (1991). Stochastic approximation of the MLE for a spatial point pattern. *Scand. J. Statist.* **18**, 39-50.

NEVEL'SON, M.B. and HAS'MINSKII, R.Z. (1973). *Stochastic Approximation and Recursive Estimation.* Amer. Math. Soc. Transl., Providence, RI.

ROBBINS, H. (1952). Some aspects of the sequential design of experiments. *Bull. Amer. Math.*

Soc. **58**, 527-535.

ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22**, 400-407.

ROBBINS, H. and SIEGMUND, D. (1971). A convergence theorem for nonnegative almost supermartingales and some applications. In *Optimizing Methods in Statistics* (J.S. Rustagi, ed.), 233-257. Academic Press, New York.

RUPPERT, D. (1991). Stochastic approximation. In *Handbook of Sequential Analysis* (B.K. Ghosh and P.K. Sen, eds.), 503-529. Dekker, New York.

SACKS, J. (1958). Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.* **29**, 375-405.

SCHMETTERER, L. (1968). Multidimensional stochastic approximation. In *Multivariate Analysis II* (P.R. Krishnaiah, ed.), 443-460. Academic Press, New York.

SPALL, J.C. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automat. Contr.* **37**, 332-341.

SPALL, J.C. (2000). Adaptive stochastic approximation by the simultaneous perturbation method. *IEEE Trans. Automat. Contr.* **45**, 1839-1853.

SPALL, J.C. and CRISTION, J.A. (1994). Nonlinear adaptive control using neural networks: Estimation with a smoothed form of simultaneous perturbation gradient approximation. *Statistica Sinica* **4**, 1-27.

TSITSIKLIS, J.N. and VAN ROY, B. (1997). An analysis of temporal-difference learning with function approximation. *IEEE Trans. Automat. Contr.* **42**, 674-690.

TSYPKIN, J.Z. (1971). *Adaptation and Learning in Automatic Systems.* Academic Press, New York.

VENTER, J.H. (1967). An extension of the Robbins-Monro procedure. *Ann. Math. Statist.* **38**, 181-190.

WEI, C.Z. (1987). Multivariate adaptive stochastic approximation. *Ann. Statist.* **15**, 1115-1130.

WU, C.J.F. (1985). Efficient sequential design with binary data. *J. Amer. Statist. Assoc.* **80**, 974-984.

YIN, G. (1991). Recent progress in parallel stochastic approximations. In *Topics in Stochastic Systems, Modeling, Estimation and Adaptive Control* (L. Gerencsér and P.E. Caines. eds.), 159-184.

YIN. G. (1999). Rates of convergence for a class of global stochastic optimization algorithms. *SIAM J. Optimiz.* **10**, 99-120.

YOUNES, L. (1989). Parameter estimation for imperfectly observed Gibbsian fields. *Probab. Theory Rel. Fields* **82**, 625-645.