STOCHASTIC APPROXIMATION AND ADAPTIVE CONTROL

Tze Leung Lai*

Columbia University

The pioneering work of Robbins and Monro in stochastic approximation initiated a rich and important field of recursive procedures in estimation and control for stochastic models. Herein the control aspects of Robbins-Monro schemes are discussed. In this connection, we also review some recent results on (i) adaptive stochastic approximation, (ii) its extensions to general multivariate stochastic regression models, and (iii) the application of these extensions to the classical problem of adaptive control of linear dynamic systems and time series models.

1. Stochastic approximation.

In 1951, Robbins and Monro introduced the subject of stochastic approximation in their seminal paper on the problem of finding the root of a regression function by successive approximations. Consider the regression model

$$y_i = M(x_i) + \epsilon_i \quad (i=1,2,\ldots), \tag{1.1}$$

where y_i denotes the response at the design level x_i, M is an unknown regression function, and ϵ_i represents unobservable noise (error). In the deterministic case (where $\epsilon_i = 0$ for all i), Newton's method for finding the root λ of a smooth function M is a sequential scheme defined by the recursion

* Research supported by the Army Research Office and the National Science Foundation.

AMS 1980 Subject classifications. Primary 62L20, 62J05, 93E20; Secondary 60G42, 93E10.

Key words and phrases. Adaptive stochastic approximation, stochastic regression models, control, self-tuning regulators, linear dynamic systems.

(1.2) $$x_{n+1} = x_n - y_n/M'(x_n).$$

When errors ϵ_i are present, using Newton's method (1.2) entails that

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

Hence, if x_n should converge to λ (so that $M(x_n) \to 0$ and $M'(x_n) \to M'(\lambda)$), then (1.3) implies that $\epsilon_n \to 0$, which is not possible for typical models of random noise (e.g., i.i.d. ϵ_n with positive variance).

To dampen the effect of the errors ϵ_i, Robbins and Monro (1951) replaced $1/M'(x_n)$ in (1.2) by constants that converge to 0. Specifically, assuming that

(1.4) $$M(\lambda) = 0, \quad \inf_{\varepsilon \leq x-\lambda \leq 1/\varepsilon} M(x) > 0 \text{ and } \sup_{\varepsilon \leq \lambda - x \leq 1/\varepsilon} M(x) < 0$$

for all $0 < \varepsilon < 1$, the Robbins-Monro scheme is defined by the recursion

(1.5) $$x_{n+1} = x_n - a_n y_n \quad (x_1 = \text{initial guess of } \lambda),$$

where a_n are positive constants such that $\sum_1^\infty a_n^2 < \infty$ and $\sum_1^\infty a_n = \infty$.

Many stochastic models of random noise ϵ_n (e.g., L_2-bounded martingale difference sequences) have the property that

(1.6) $\sum_1^\infty \alpha_n \epsilon_n$ converges a.s. and in L_2 for all constants α_n such that $\sum_1^\infty \alpha_n^2 < \infty$.

For such models of random errors, Blum (1954) showed that the Robbins-Monro scheme (1.5) converges to λ a.s., under the assumptions (1.4) and

(1.7) $$|M(x)| \leq c|x| + d \text{ for some } c,d \text{ and all } x.$$

Earlier, Robbins and Monro (1951) showed that the scheme converges to λ in L_2.

Suppose that the random variables ϵ_n are i.i.d. with mean 0 and variance σ^2. Chung (1954) and Sacks (1958) established the asymptotic normality of the scheme (1.5) for various choices of a_n and showed that the smallest variance of the asymptotic normal distribution is attained by the choice $a_n \sim (nM'(\lambda))^{-1}$, assuming also that $M'(\lambda)$ exists and is positive. For this asymptotically optimal choice $a_n \sim (nM'(\lambda))^{-1}$, the Robbins-Monro scheme (1.5) is asymptotically normal with mean θ and variance $n^{-1}(\sigma/M'(\lambda))^2$.

In particular, consider the linear model

$$y_i = \beta(x_i - \lambda) + \epsilon_i \qquad (i=1,2,\ldots), \tag{1.8}$$

where $\beta > 0$ is known and the ϵ_i are i.i.d. with mean 0 and variance σ^2. The least squares estimate of λ based on $x_1, y_1, \ldots, x_n, y_n$ is

$$\lambda_n^* = \bar{x}_n - \bar{y}_n/\beta \; (= \lambda - \bar{\epsilon}_n/\beta). \tag{1.9}$$

Here and in the sequel, we use the notation $\bar{z}_n$ to denote the arithmetic mean $n^{-1}\sum_1^n z_i$. Thus, irrespective of how the levels x_i are chosen, $\lambda_n^* (= \lambda - \bar{\epsilon}_n/\beta)$ is asymptotically normal with mean λ and variance $n^{-1}(\sigma/\beta)^2$, and $\lambda_n^* \to \lambda$ a.s. Moreover, letting the least squares estimate $\lambda_n^* (= \bar{x}_n - \bar{y}_n/\beta)$ at stage n be the design level x_{n+1} at the next stage is equivalent to the Robbins-Monro scheme (1.5) with $a_n = (n\beta)^{-1}$, in view of

THEOREM 1 (Lai and Robbins, 1979). For every real number c and positive integer n, we have the equivalence

$$x_{i+1} = \bar{x}_i - c\bar{y}_i \quad (i=1,2,\ldots,n) \Leftrightarrow x_{i+1} = x_i - cy_i/i \quad (i=1,\ldots,n).$$

Hence, in the linear model (1.8) with i.i.d. errors ϵ_i, the asymptotically optimal Robbins-Monro scheme with $a_n = (n\beta)^{-1}$ has the same statistical properties as the sequence of least squares estimates λ_n^*. For nonlinear regression functions M such that $M'(\lambda) = \beta$ exists and is positive, since $M(x) \sim \beta(x-\lambda)$ as $x \to \lambda$, we still have the same asymptotic distribution

given by

$$n^{1/2}(x_n - \lambda) \xrightarrow{\mathcal{D}} N(0,\sigma^2/\beta^2), \tag{1.10}$$

provided that $x_n \to \lambda$ a.s. The latter a.s. convergence was established by Blum (1954) for general Robbins-Monro schemes.

In practice, $\beta = M'(\lambda)$ is usually unknown. To obtain strongly consistent estimates of β, Venter (1967) proposed the following modification of the Robbins-Monro scheme (1.5) with $a_n = (n\beta)^{-1}$. At the m^{th} stage (m=1,2,...), two observations y_m' and y_m'' are taken at levels $x_m' = \lambda_m - c_m$ and $x_m'' = \lambda_m + c_m$, where $\{c_m\}$ is a sequence of positive constants such that

$$c_m \sim cm^{-\gamma} \text{ for some constants } c > 0 \text{ and } \frac{1}{4} < \gamma < \frac{1}{2}, \tag{1.11}$$

and λ_m is the m^{th} approximation to λ defined recursively by

$$\lambda_{i+1} = \lambda_i - y_i/(ib_i) \qquad (\lambda_1 = \text{initial guess of } \lambda). \tag{1.12}$$

The quantity y_i in (1.12) estimates the (unobserved) response at the level λ_i and is defined by $y_i = \frac{1}{2}(y_i' + y_i'')$. Assuming that positive constants b and B are known such that $b < \beta < B$, Venter defines the slope estimate b_i in (1.12) by

$$b_i = b \vee \{B \wedge i^{-1} \sum_{j=1}^{i} (y_j'' - y_j')/(2c_j)\},$$

where the symbols $\vee$ and $\wedge$ denote maximum and minimum respectively. Note that by (1.1)

$$y_i' = M(\lambda_i - c_i) + \epsilon_i', \quad y_i'' = M(\lambda_i + c_i) + \epsilon_i'',$$

$$y_i = \frac{1}{2}\{M(\lambda_i - c_i) + M(\lambda_i + c_i)\} + \frac{1}{2}(\epsilon_i' + \epsilon_i'').$$

Suppose that the errors $\epsilon_1', \epsilon_2', \ldots, \epsilon_1'', \epsilon_2'', \ldots$ are i.i.d. with mean 0 and variance σ^2. Then $\mathrm{Var}\{\frac{1}{2}(\epsilon_i' + \epsilon_i'')\} = \frac{1}{2}\sigma^2$, and under certain assumptions on M, Venter showed that (1.10) with $n = 2m$ and $x_n = \lambda_m$ still holds for this modification of the Robbins-Monro scheme; moreover, $b_i \to \beta$ a.s.

2. The control aspects of stochastic approximation.

The preceding section shows that the successive approximations in the Robbins-Monro scheme (1.5) provide strongly consistent estimates of the root λ of an unknown regression function M. Moreover, by choosing a_n in (1.5) to be $a_n = (n\beta)^{-1}$, or to be $a_n = (nb_n)^{-1}$ as in Venter's modification (1.12), these recursive estimates are asymptotically equivalent (in the sense of (1.10)) to the least squares estimates under the linear model (1.8) with β known and with i.i.d. random errors. These results can be considered as the estimation aspects of stochastic approximation, and pave the way for a rich variety of subsequent developments in recursive estimation and identification algorithms (cf. Sakrison, 1965; Albert and Gardner, 1967; Nevel'son and Hasminski, 1972; Fabian, 1978; Ljung, 1977; Solo, 1981).

Let $\mathbf{F}_i$ denote the σ-field generated by $x_1, y_1, \ldots, x_i, y_i$. While the n^{th} approximation x_n in the Robbins-Monro scheme (1.5) provides an $\mathbf{F}_{n-1}$-measurable estimate of the unknown λ, it is also the design level at which the response y_n is to be observed. We now discuss the usefulness of the stochastic approximation design in adaptive control problems. Suppose that in the regression model (1.1) the objective is to regulate the outputs y_i so that they are as close as possible to some target value, which we assume to be 0 without loss of generality. Thus, the control problem is to minimize $\sum_1^n y_i^2$ in some sense, at least in the long run as $n \to \infty$. The linear case ($M(x) = \alpha+\beta x$) of this problem has been discussed in the econometrics literature; this is the so-called "multiperiod control problem under uncertainty" (cf. Zellner, 1971; Prescott, 1972; Anderson and Taylor, 1976).

Consider the linear regression model $y_i = \alpha + \beta x_i + \epsilon_i$ with i.i.d. random errors ϵ_i having mean 0 and variance σ^2. A common approach to formulate

the multiperiod control problem of choosing successive levels $x_1,\ldots,x_N$ is to minimize

$$(2.1)\quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\Big[\sum_{i=1}^{N} y_i^2 \,|\, \alpha,\beta\Big]\, d\pi(\alpha,\beta) = N\sigma^2 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\Big[\beta^2 \sum_{i=1}^{N} (x_i-\lambda)^2 \,|\, \alpha,\beta\Big]\, d\pi(\alpha),$$

where $\lambda = -\alpha/\beta$. However, because of the analytical and computational difficulties in studying the dynamic programming equations, not much is known about the Bayes rules and it is difficult to implement them in practice. Instead of this Bayesian approach, Anderson and Taylor (1976) noted that since the optimal level is λ when α and β are known and since the least squares estimate of (β,α) based on $x_1,y_1,\ldots,x_n,y_n$ is

$$(2.2)\qquad \hat{\beta}_n = \Big\{\sum_{i=1}^{n} (x_i-\bar{x}_n)y_i\Big\} \Big/ \sum_{i=1}^{n} (x_i-\bar{x}_n)^2, \quad \hat{\alpha}_n = \bar{y}_n - \hat{\beta}_n \bar{x}_n,$$

it seems reasonable to replace α and β in the optimal level $\lambda = -\alpha/\beta$ by $\hat{\alpha}_n$ and $\hat{\beta}_n$. Thus, assuming prior knowledge of bounds K_1 and K_2 such that $K_1 \leqslant \lambda \leqslant K_2$, Anderson and Taylor proposed the rule

$$(2.3)\qquad x_{n+1} = K_2 \wedge (-\hat{\alpha}_n/\hat{\beta}_n \vee K_1),\ n \geqslant 2,$$

where $\wedge$ and $\vee$ denote minimum and maximum respectively, and the initial values x_1,x_2 are distinct but otherwise arbitrary numbers between K_1 and K_2. When the ϵ_n are normal, this rule is tantamount to choosing the maximum likelihood estimate of λ at stage n to be the design level at stage n+1. On the basis of some Monte Carlo simulations, Anderson and Taylor (1976) conjectured that $\{x_n\}$ defined by (2.3) converges to λ a.s. However, Lai and Robbins (1982a) disproved this conjecture by exhibiting an event Ω_0 with positive probability on which x_n does not converge to λ. In fact, on Ω_0, there is so little information in the design to estimate β that $\hat{\beta}_n$ is of the wrong sign for all n.

While setting the current level as close to λ as the data suggest (like the Anderson-Taylor rule) seems to be consistent with the control

objective, there is also the conflicting need to set the design levels in a wide enough scatter so as to provide a reliable estimate of λ. This dilemna between information and control, however, disappears if we assume that $\beta \neq 0$ is in fact known. In this case, by (1.9), the least squares estimate $\lambda_n^* = \bar{x}_n - \bar{y}_n/\beta$ has the same distribution no matter how the levels x_i are chosen. In view of the control objective of setting the design levels as close to λ as possible, it is therefore natural to set

$$x_{n+1} = \bar{x}_n - \bar{y}_n/\beta \ (=\lambda_n^*). \tag{2.4}$$

For this choice, x_n has the asymptotic (normal) distribution given by (1.10). Moreover, by (1.9),

$$E[\beta^2 \sum_{i=2}^{n} (x_i-\lambda)^2] = \beta^2 E\{\sum_{i=1}^{n-1} \bar{\epsilon}_i^2/\beta^2\} \sim \sigma^2 \log n. \tag{2.5}$$

Since choosing the levels x_i to minimize $E(\sum_1^n y_i^2)$ is equivalent to minimizing $E[\beta^2 \sum_{i=1}^{n} (x_i-\lambda)^2]$ by (2.1), we call

$$C_n = \beta^2 \sum_{i=1}^{n} (x_i - \lambda)^2 \tag{2.6}$$

the "cost" of the control sequence $\{x_i\}$. Thus, (2.5) shows that $E(C_n)$ is of the order $\sigma^2 \log n$ for the control rule (2.4). Moreover, for this control rule,

$$C_n \sim \sigma^2 \log n \quad \text{a.s.}, \tag{2.7}$$

as was shown by Lai and Robbins (1979).

When the errors ϵ_i are normal, the least squares estimate $\lambda_n^* = \bar{x}_n - \bar{y}_n/\beta$ is the same as the maximum likelihood estimate. Putting a prior normal distribution π on λ, it can be shown that the control rule

$$x_{n+1} = E_\pi(\lambda | x_1, y_1, \ldots, x_n, y_n) = \text{Bayes estimate} \tag{2.8}$$

indeed minimizes the Bayes risk $\int_{-\infty}^{\infty} E(\sum_{n=1}^{N} y_n^2 | \lambda) d\pi(\lambda)$ for every $N \geq 1$. Since the maximum likelihood estimate λ_n^* and the Bayes estimate $E_\pi(\lambda | x_1, y_1, \ldots, x_n, y_n)$ are asymptotically equivalent, the Bayes rule (2.8) has the same asymptotic properties as (2.4). In particular, (2.7) also holds for the cost of the Bayes rule (2.8) (cf. Lai and Robbins, 1982b).

By Theorem 1, the control rule (2.4) is equivalent to the stochastic approximation scheme

$$x_{n+1} = x_n - y_n/(n\beta). \tag{2.9}$$

Now consider the general regression model (1.1), where M satisfies (1.4), (1.7) and the condition $M'(\lambda) = \beta > 0$. Suppose that β is unknown, and that we replace β in (2.7) by an $\mathcal{F}_{n-1}$-measurable estimate b_n such that $b_n \to \beta$ a.s., leading to the adaptive stochastic approximation scheme

$$x_{n+1} = x_n - y_n/(nb_n). \tag{2.10}$$

Defining the cost C_n of the design by (2.6), Lai and Robbins (1979) showed that the cost C_n of the adaptive stochastic approximation scheme (2.10) is still of the order $\sigma^2 \log n$, as in (2.7). Moreover, x_n defined by (2.10) converges to λ a.s. and has the same asymptotically normal distribution as that given in (1.10).

The construction of strongly consistent estimates b_n in the scheme (2.10) is given by Lai and Robbins (1981). When upper and lower bounds B and b > 0 for $\beta = M'(\lambda)$ are known, one can use the truncated version $b \vee (\hat{\beta}_n \wedge B)$ of the least squares estimate $\hat{\beta}_n$ defined in (2.2). Without assuming prior knowledge of these upper and lower bounds, it is also possible to modify the argument to construct a more elaborate slope estimate that is strongly consistent.

Although the asymptotic normality property (1.10) also holds for Venter's (1967) modification of the Robbins-Monro scheme, the cost of Venter's

design

$$C_n = \beta^2\{\sum_1^m (x_i' - \lambda)^2 + \sum_1^m (x_i'' - \lambda)^2\} \qquad (n = 2m)$$

is of the order $4^{\gamma}c^2(1-2\gamma)^{-1}n^{1-2\gamma}$, where c and γ are given by (1.11) (cf. Lai and Robbins, 1979), instead of the logarithmic order (2.7) for the adaptive stochastic approximation scheme (2.10).

3. Multivariate extensions and the adaptive control of linear dynamic systems.

A classical problem in the literature on stochastic adaptive control is the optimal regulation of the linear system

$$(3.1) \qquad y_n = \alpha_1 y_{n-1} + \ldots + \alpha_p y_{n-p} + \beta_1 x_{n-1} + \ldots + \beta_q x_{n-q} + \epsilon_n,$$

where the y's represent outputs and the x's represent inputs at various times and the ϵ's represent random disturbances. In the time series literature, (3.1) represents the autoregressive model with exogenous variables x_i. In the engineering literature, an important problem is how to choose the inputs adaptively to regulate the outputs, say, such that $\sum_1^n y_i^2$ is minimized in some sense in ignorance of the system parameters $\alpha_1,\ldots,\alpha_p,\beta_1,\ldots,\beta_q$, at least asymptotically as $n \to \infty$. Although one may in principle use a Bayesian approach, putting a prior distribution on the unknown parameters and applying dynamic programming when the disturbances ϵ_n are i.i.d. with a known common distribution and when there is a finite horizon, the dynamic programming problem is prohibitively difficult, both computationally and analytically.

A much more practical approach to this adaptive control problem is that of the self-tuning regulator, proposed by Åström and Wittenmark (1973) following an earlier idea due to Kalman (1958). The "self-tuning" idea is to start by considering the case where the system parameters are known, for which the optimal controller is given by

$$x_n = -(\alpha_1 y_n + \ldots + \alpha_p y_{n-p+1} + \beta_2 x_{n-1} + \ldots + \beta_q x_{n-q+1})/\beta_1, \tag{3.2}$$

assuming that $\beta_1 \neq 0$, or equivalently by the equation

$$\theta' \phi_n = 0, \tag{3.3}$$

where $\theta = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)'$ and $\phi_n = (y_n, \ldots, y_{n-p+1}, x_n, \ldots, x_{n-q+1})'$. Regarding (3.1) as the regression model $y_{n+1} = \theta' \phi_n + \epsilon_{n+1}$, let

$$\theta_n = \left(\sum_1^n \phi_i \phi_i' \right)^{-1} \sum_1^n \phi_i y_{i+1} \tag{3.4}$$

be the least squares of θ based on $\phi_1, y_2, \ldots, \phi_n, y_{n+1}$. Without assuming prior knowledge of the system parameters, the self-tuning regulator substitutes θ in (3.3) by its least squares estimate and defines x_n implicitly by the equation

$$\theta_{n-1}' \phi_n = 0. \tag{3.5}$$

When the random disturbances are i.i.d. normal with mean 0, θ_i is the maximum likelihood estimate of θ, and therefore the self-tuning regulator (3.5) is simply the maximum likelihood estimate of the (unobservable) optimal regulator (3.2) at every stage. While (3.5) may well represent one's best guess of (3.2), how good the guess is depends on how much information there is to estimate θ. A measure of the amount of information available to estimate θ on the basis of $\phi_1, y_2, \ldots, \phi_n, y_{n+1}$ is the information matrix $\sum_1^n \phi_i \phi_i'$. As shown by Lai and Wei (1982), the least squares estimate θ_n converges to θ a.s. if $\lambda_{\min}(\sum_1^n \phi_i \phi_i') \to \infty$ and

$$\log \lambda_{\max}\left(\sum_1^n \phi_i \phi_i' \right) = o\left(\lambda_{\min}\left(\sum_1^n \phi_i \phi_i' \right)\right) \quad \text{a.s.}, \tag{3.6}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ denote the maximum and minimum eigenvalues of a symmetric matrix.

This self-tuning idea is therefore analogous to the Anderson-Taylor rule (2.3) for the multiperiod control problem related to the simple linear model $y_i = \alpha + \beta x_i + \epsilon_i$ discussed in Section 2. A basic issue concerning the self-tuning approach of substituting the parameters by their sequential estimates is whether there is enough information for adequately estimating these parameters.

We now consider this issue and the dilemna between the need for information and the objective of efficient control in the more general context of a stochastic regression model

$$Y_n = AZ_n + BX_n + e_n, \tag{3.7}$$

where

(i) Y_n is a kxl vector of outputs at stage n;

(ii) A is a kxh matrix and B is a kxk nonsingular matrix of parameters;

(iii) e_n is a kxl vector of random disturbances such that $\{e_n\}$ is a martingale difference sequence with respect to an increasing sequence of σ-fields $\{\mathbf{F}_n\}$ and $\sup_n E(||e_n||^\gamma|\mathbf{F}_{n-1}) < \infty$ a.s. for some $\gamma > 2$;

(iv) X_n is a kxl vector of regulators depending on the current and past observations $Z_1, X_1, Y_1, \ldots, Z_{n-1}, X_{n-1}, Y_{n-1}, Z_n$;

(v) Z_n is an $\mathbf{F}_{n-1}$-measurable hxl vector of regressors.

The control problem is to choose the regulators $X_1, X_2, \ldots$ sequentially so that $\sum_1^n ||Y_i||^2$ is minimized in some sense, at least asymptotically as $n \to \infty$. A special case of the stochastic regression model (3.7) is the simple linear model $y_n = \alpha + \beta x_n + \epsilon_n$ in a multiperiod control problem discussed in Section 2, where $Z_n = 1$. The linear system (3.1) is also a special case of (3.7), where we set $Y_n = y_{n+1}$, $e_n = \epsilon_{n+1}$, $X_n = x_n$, and $Z_n = (y_n, \ldots, y_{n-p+1}, x_{n-1}, \ldots, x_{n-q+1})'$, so h = p+q-1 in this case.

Suppose that the parameter matrices A and B in (3.7) are known. Then the optimal regulators are given by $X_i = -B^{-1}AZ_i$, and the optimal outputs are $Y_i = \epsilon_i$. This leads us to define, in analogy with (2.6), the cost at stage n

of a sequence of regulators $\{X_i\}$ to be

$$C_n = \sum_{i=1}^{n} ||AZ_i + BX_i||^2 = \sum_{i=1}^{n} ||Y_i - e_i||^2. \tag{3.8}$$

A control rule is called "globally convergent" if

$$C_n/n \to 0 \quad \text{a.s.} \tag{3.9}$$

For the linear system (3.7), Goodwin, Ramadge and Caines (1981) made use of stochastic approximation techniques to estimate the system parameters and showed that the resultant rule is globally convergent. For the multiperiod control problem related to the simple linear model $y_n = \alpha + \beta x_n + \epsilon_n$, the "self-tuning" rule (2.3) proposed by Anderson and Taylor is not globally convergent, as has been shown by Lai and Robbins (1982a). On the other hand, the adaptive stochastic approximation scheme (2.10) with $b_n \to \beta$ a.s. is not only globally convergent but also has the much stronger property (2.7) of only logarithmic order for C_n. Moreover, this order of magnitude for C_n is shown in Section 2 to be asymptotically minimal by studying the auxiliary problem in which β is known.

To extend these ideas to the stochastic regression model (3.7), we therefore start by assuming that B is known. Moreover, in analogy with (1.8), we rewrite (3.7) as

$$Y_n = B(X_n - DZ_n) + e_n, \tag{3.10}$$

where $D = -B^{-1}A$. Noting that (3.10) corresponds to the regression model $X_n - B^{-1}Y_n = DZ_n - B^{-1}e_n$, the least squares estimate of D at stage n is given by

$$D_n^* = \{ \sum_{i=1}^{n} (X_i - B^{-1}Y_i)Z_i' \} (\sum_{i=1}^{n} Z_i Z_i')^{-1}, \tag{3.11}$$

where inverse denotes the Moore-Penrose generalized inverse. First suppose that

$$(3.12)\qquad Z_n = O(1) \text{ and } \liminf_{n\to 0} n^{-1}\lambda_{\min}\Big(\sum_1^n Z_i Z_i'\Big) > 0 \quad \text{a.s.},$$

and that

(3.13) the distribution of $\{Z_n\}$ does not depend on the choice of $\{X_n\}$.

This is the case when the Z_n are nonrandom, such as in multiperiod control problem of Section 2. Assume also that the e_n are i.i.d. with mean 0 and covariance matrix V. Then by (3.10) and (3.11), $D_n^* = D\text{-}B^{-1}\Big(\sum_{i=1}^n e_i Z_i'\Big)\Big(\sum_{i=1}^n Z_i Z_i'\Big)$, and therefore the distribution of D_n^* does not depend on the choice of $\{X_i\}$ by (3.13). In particular, in view of the control objective of setting X_n as close to DZ_n as possible, it is natural to use the control rule

$$(3.14)\qquad X_n = D_{n-1}^* Z_n,$$

which is the analogue of $x_n = \lambda_{n-1}^*$ in (2.4). Lai (1985) has shown that for the rule (3.14)

$$(3.15)\qquad C_n \sim (\text{tr } V)\, h \log n \text{ a.s.},$$

which provides an extension of (2.7).

The asymptotic behavior (3.15) of the cost assuming knowledge of B provides a bench mark that we try to achieve even when B is unknown. Without assuming knowledge of B, we make use of an approach similar to that of adaptive stochastic approximation described in Section 2. First, in analogy with Theorem 1 that expresses the sequence λ_n^* $(= \bar{x}_n - \bar{y}_n/\beta)$ in stochastic approximation form, we apply the Kalman-Plackett recursive representation of the least squares estimate D_n^*:

$$(3.16)\qquad D_n^* = D_{n-1}^* + (X_n - B^{-1}Y_n - D_{n-1}^* Z_n)Z_n' P_n,$$

(3.17) $$P_n = P_{n-1} - P_{n-1}Z_nZ_n'P_{n-1}/(1+Z_n'P_{n-1}Z_n),$$

(cf. Lai, 1985). Next we construct nonsingular kxk matrices B_n such that B_n is $\mathcal{F}_{n-1}$-measurable and $B_n \to B$ a.s. The construction of such B_n is discussed below. Replacing B by B_n in (3.16) leads to the recursive estimates

(3.18) $$D_n = D_{n-1} + (X_n - B_n^{-1}Y_n - D_{n-1}Z_n)Z_n'P_n,$$

with P_n defined recursively by (3.17).

To construct $\mathcal{F}_{n-1}$-measurable estimates B_n such that $B_n \to B$ a.s., we introduce white-noise probing inputs at certain prespecified times $n_1, n_2, \ldots$ to ensure that there is enough information to estimate B. Specifically, at stage $n \in I = \{n_1, n_2, \ldots\}$, the input X_n is a random vector independent of $Z_1, X_1, Y_1, \ldots, Z_{n-1}, X_{n-1}, Y_{n-1}, Z_n$ such that the components of X_n are uniformly bounded and independently distributed with mean 0 and common positive variance. The inputs at the other stages are given by the self-tuning rule

(3.19) $$X_n = D_{n-1}Z_n, \quad n \notin I.$$

The use of white-noise perturbations to improve the information content of the design is a well known idea in adaptive control of linear systems. It is similar in spirit to Venter's modification of the Robbins-Monro scheme discussed in Section 1. Dropping the i.i.d. assumption on e_n and the assumption (3.13) on Z_n, we are still able to do asymptotically as well as (3.15) by having the number of probing inputs up to stage ν diverge to ∞ but within the order of $o(\log \nu)$ (see (3.20) below), and by using strongly consistent estimates B_n of B given in Lai (1985, Section 4).

THEOREM 2 (Lai, 1985). Suppose that in the stochastic regression model (3.7), $\{Z_n\}$ satisfies (3.12). For the sequence of inputs $\{X_n\}$ described above, with B_n

suitably chosen and with

$$i/\log n_i \to 0, \tag{3.20}$$

we have

$$C_n \leq \{\limsup_{i\to\infty} E(||e_i||^2|\mathcal{F}_{i-1})+o(1)\}h \log n \text{ a.s.} \tag{3.21}$$

Although the linear system (3.1) is a special case of the regression model (3.7), condition (3.12) need not hold and Theorem 2 is not applicable. However, besides the regression structure, we can also make use of the system dynamics defined by (3.1) to construct adaptive regulators that satisfy (3.21) with $h = p+q-1$, which is the dimensionality of Z_n in this case. The details are given in Lai and Wei (1986).

REFERENCES

Albert, A.E. and Gardner, L.A. (1967). Stochastic Approximation and Nonlinear Regression. MIT Press, Cambridge, MA.

Anderson, T.W. and Taylor, J. (1976). Some experimental results in the statistical properties of least squares estimates in a control problem. Econometrica. 44 1289-1302.

Åström, K.J. and Wittenmark, B. (1973). On self-tuning regulators. Automatica. 9 195-199.

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

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

Fabian, V. (1978). On asymptotically efficient recursive estimation. Ann. Statist. 6 854-866.

Goodwin, G.C., Ramadge, P.J., and Caines, P.E. (1981). Discrete time stochastic adaptive control. SIAM J. Contr. and Optimization. 19 829-853.

Kalman, R.E. (1958). Design of a self-optimizing control system. Trans. ASME. 80 468-478.

Lai, T.L. (1985). Asymptotically efficient adaptive control in stochastic regression models. To appear in Adv. Appl. Math.

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). Iterated least squares in multiperiod control. Adv. Appl. Math. **3** 50-73.

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

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). Asymptotically efficient self-tuning regulators. To appear in SIAM J. Contr. and Optimization.

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

Nevel'son, M.B. and Has'minski, R.Z. (1972). Stochastic Approximation and Recursive Estimation. Amer. Math. Soc. Translations, Providence, RI.

Prescott, E.C. (1972). The multiperiod control problem under uncertainty. Econometrica **40** 1043-1058.

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

Sacks, J. (1958). Asymptotic distribution of stochastic approximation procedures. Ann. Math. Statist. **29** 373-405.

Sakrison, D.J. (1965). Efficient recursive estimation; application to estimating the parameters of a covariance function. Int. J. Eng. Sci. **3** 461-483.

Solo, V. (1981). The second order properties of a time series recursion. Ann. Statist. **9** 307-317.

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

Zellner, A. (1971). An Introduction to Bayesian Inference in Econometrics. Wiley, New York.