Estimating Time-Varying Parameters by the Kalman Filter Based Algorithm: Stability and Convergence

LEI GUO, MEMBER, IEEE

Abstract—**By introducing new techniques, in this paper we establish convergence and stability properties of the Kalman filter based parameter estimator for linear stochastic time-varying regression models. The main features are: 1) both the variances and the sample path averages of the parameter tracking error are shown to be bounded; 2) the regression vector includes both stochastic and deterministic signals, and no assumptions of stationarity or independence are required; and 3) the unknown parameters are only assumed to have bounded variations in an average sense.**

I. Introduction

LET us consider the following time-varying stochastic linear regression model:

$$y_k = \varphi_k^\tau \theta_k + v_k, \qquad k \geq 0 \tag{1.1}$$

where y_k and v_k are the observation and the noise, respectively, and φ_k and θ_k are, respectively, the p-dimensional stochastic regression vector and the unknown time-varying parameter. Denote the parameter variation at time k by w_k

$$w_k = \theta_k - \theta_{k-1}, \qquad k \geq 1, \ E\|\theta_0\|^3 < \infty. \tag{1.2}$$

For estimating the unknown parameter θ_k, we introduce the following Kalman filter based adaptive estimator:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \frac{P_k \varphi_k}{R + \varphi_k^\tau P_k \varphi_k}(y_k - \varphi_k^\tau \hat{\theta}_k), \tag{1.3}$$

$$P_{k+1} = P_k - \frac{P_k \varphi_k \varphi_k^\tau P_k}{R + \varphi_k^\tau P_k \varphi_k} + Q \tag{1.4}$$

where $P_o > 0$, $R > 0$, and $Q > 0$ as well as $\hat{\theta}_0$ are deterministic and can be arbitrarily chosen (here R and Q may be regarded as the *a priori* estimates for the variances of v_k and w_k, respectively).

Note that if we take $R = 1$ and $Q = 0$, then (1.3), (1.4) become the standard least-squares algorithm which is commonly used in the special case where the parameter process is constant, i.e., $w_k = 0$ for all k.

In a "classical" Bayesian analysis of linear regression models (e.g., Lindley and Smith [1]), Q is a hyperparameter of prior distributions of the unknown parameters. With Gaussian assumptions and hyperparameter Q, Kitagawa and Gersh [2] presented a Kalman filter algorithm for the estimation of time-varying linear models with a worked example.

Manuscript received December 15, 1988; revised June 12, 1989. Paper recommended by Associate Editor, A. Benveniste.

The author was with the Department of Systems Engineering, The Australian National University, Canberra, Australia. He is now with the Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China.

IEEE Log Number 8932847.

It is known that if the regression vector φ_k belongs to F_{k-1}, the σ-algebra generated by $\{y_o, \cdots, y_{k-1}\}$, and the random process $\{w_k, v_k\}$ is Gaussian and white, then $\hat{\theta}_k$ generated by (1.3), (1.4) is the best estimate for θ_k, and P_k is the estimation error covariance, i.e.,

$$\hat{\theta}_k = E[\theta_k | F_{k-1}], \ P_k = E[\tilde{\theta}_k \tilde{\theta}_k^\tau | F_{k-1}], \qquad (\tilde{\theta}_k = \theta_k - \hat{\theta}_k) \tag{1.5}$$

provided that $Q = E w_k w_k^\tau$, $R = E v_k v_k^\tau$, $\hat{\theta}_o = E\theta_o$, and $P_o = E[\tilde{\theta}_o \tilde{\theta}_o^\tau]$ (see, e.g., Mayne [3], Astrom and Wittenmark [4], Kitagawa and Gersch [2], and Solo [5]).

A natural question now arises: is the tracking error $\tilde{\theta}_k$ bounded in some sense?

Unfortunately, general conditions in the case of ***stochastic*** *regressors have been difficult to find* even for the case where (1.5) holds. This problem is related to the stability issue in Kalman filtering theory, however, for that study a commonly used condition is that the regression vector φ_k is deterministic, and satisfies

$$\alpha I \leq \sum_{k=n}^{n+N} \varphi_k \varphi_k^\tau \leq \beta I, \qquad \forall n \tag{1.6}$$

for some deterministic positive constants α, β, and N (see, e.g., Jazwinski [6]). This condition is a uniformly completely observable requirement for the associated time-varying linear system, and is also known as "persistence of excitation" in the adaptive control literature (see, e.g., Anderson *et al.* [7]). *It is immediately seen that condition* (1.6) *is a mainly deterministic hypothesis, and* ***unsuitable*** *for general stochastic models, since it fails for possibly unbounded regressors (e.g., Gaussian signals), and* ***even*** *fails for a bounded independent and identically distributed (i.i.d.) signal!* A condition which is weaker than (1.6) and allows the regressor to be unbounded was presented in (Guo, Xia, and Moore [8]) by introducing stopping times. However, that condition is also imposed on the sample paths of $\{\varphi_k\}$, and is therefore also difficult to verify in the stochastic case.

Before pursuing further, some related work in the area of adaptive signal processing should be mentioned, although the algorithms considered there are mainly the least mean square (LMS) algorithm. This algorithm is formed by simply taking the adaptation gain $(P_k \varphi_k)/(R + \varphi_k^\tau P_k \varphi_k)$ in (1.3) as $\mu \varphi_k$, where μ is a stepsize. As far as the tracking aspect is concerned, Widow *et al.* [9] produced insightful heuristic analysis, Benveniste and Ruget [10] used the methods of continuous-time model approximation and gave bounds for vanishing small μ, Eweda and Macchi [11] studied the case of deterministic parameter variation where the joint regression vector output process $\{\varphi_k, y_k\}$ are M-dependent, Macchi [12] assumed that the regression vector is stationary, M-dependent, and independent of $\{\theta_k, v_k\}$, and in Benveniste [13] multistep schemes were analyzed and a complete design methodology of adaptive algorithms was presented. For some other in-

0018-9286/90/0200-0141$01.00 © 1990 IEEE

teresting studies see, e.g., Bitmead and Anderson [14], Shi and Kozin [15], Benveniste *et al.* [16], and the recent work of Solo [17].

Besides the Kalman filtering algorithm and the LMS algorithm mentioned above, there is also a number of estimation algorithms used for identifying/tracking time-varying parameters in the area of system identification, e.g., the forgetting factor algorithm, the gain resetting algorithm, the projection algorithm, etc. Again, precise theoretical analyses for stochastic models are difficult to find.

In this paper, we study properties of the estimation algorithm (1.3), (1.4) applied to stochastic regression model (1.1). *The main contributions of the paper are the investigation of stability properties of Kalman filter based algorithms when the regressors are stochastic and nonstationary, and the establishment of tracking error bounds for the unknown time-varying parameters.* Both the conditions and the techniques for analysis are different from the traditional ones used in the areas of system identification and adaptive signal processing.

The paper is organized as follows. In Section II we introduce the new excitation condition and state the main results. The proof of these results is given in Section III. Section IV concludes the paper with remarks.

II. Main Results

In the following, by the norm $\|X\|$ of a real matrix X, we always mean that $\|X\| = \{\lambda_{\max}(XX^\tau)\}^{1/2}$, and by $\lambda_{\max}(X)[\lambda_{\min}(X)]$ we mean the maximum (minimum) eigenvalue of X.

We now introduce the assumptions of the paper.

A1: $\{v_k, w_k\}$ is a random or deterministic process satisfying

$$\sigma_r \triangleq \sup_k E\{\|v_k\|^r + \|w_k\|^r\} < \infty, \qquad \text{for some } r > 4,$$

$$\mu_4 \triangleq \limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\{\|v_k\|^4 + \|w_k\|^4\} < \infty, \qquad \text{a.s.}$$

A2: $\{\varphi_k, F_k\}$ is an adapted sequence (i.e., φ_k is F_k-measurable for any k, where F_k is a sequence of nondecreasing σ-algebras), and there exist a constant $\delta > 0$ and an integer $h > 0$ such that

$$E\left\{\sum_{k=mh}^{(m+1)h-1} \frac{\varphi_k\varphi_k^\tau}{1+\|\varphi_k\|^2} | F_{mh-1}\right\} \geq \delta I, \qquad \text{a.s., } \forall m \geq 0. \tag{2.1}$$

Note that **no** assumptions of independence and stationarity are made on the signals $\{\varphi_i, v_i, w_i\}$. In particular, 1) the unknown parameter $\{\theta_i\}$ is allowed to be, e.g., a stationary process, a random walk, or a bounded deterministic sequence; 2) the assumptions on the regression vector $\{\varphi_k\}$ do not exclude signals derived from feedback.

It is evident that condition A2 is weaker than (1.6). Let us further illustrate this condition by considering the following examples where the regressor is stochastic.

Example 1: Let $\{\varphi_k\}$ be an M-dependent sequence (i.e., there exists some integer M such that, for any k, $\{\varphi_j, j \leq k\}$ and $\{\varphi_j, j > k+M\}$ are independent) which satisfies

$$\inf_k \lambda_{\min}(E[\varphi_k\varphi_k^\tau]) > 0, \qquad \text{and } \sup_k E\|\varphi_k\|^4 < \infty$$

then condition A2 holds.

Proof: Take $F_k = \sigma\{\varphi_j, j \leq k\}$, $h = M+1$, then by the M-dependency assumption, we have for any $m \geq 0$ ($t = (m+1)h - 1$)

$$E\left[\frac{\varphi_t\varphi_t^\tau}{1+\|\varphi_t\|^2}|F_{mh-1}\right] = E\left[\frac{\varphi_t\varphi_t^\tau}{1+\|\varphi_t\|^2}\right]$$

but, by the Schwarz inequality it is easy to verify that

$$\lambda_{\min}\left(E\frac{\varphi_t\varphi_t^\tau}{1+\|\varphi_t\|^2}\right) \geq \frac{\{\lambda_{\min}(E[\varphi_t\varphi_t^\tau])\}^2}{E[(1+\|\varphi_t\|^2)\|\varphi_t\|^2]}$$

hence condition A2 is true.

Example 2: Let $\{\varphi_k\}$ be generated by a linear model

$$\varphi_k = F\varphi_{k-1} + G\xi_k, \qquad k \geq 0$$

where F is a stable matrix, (F, G) is controllable, and $\{\xi_k\}$ is an i.i.d. sequence with $E\xi_k = 0$, $E\xi_k\xi_k^\tau > 0$, and $\|\xi_k\| \leq M$, for some constant M. Then condition A2 holds.

Proof: Since for any $m \geq 0$ and $h > 0$, $k \geq mh$

$$\varphi_k = F^{(k-mh+1)}\varphi_{mh-1} + \sum_{i=mh}^{k} F^{(k-i)}G\xi_i$$

we have by taking $F_k = \sigma\{\xi_j, j \leq k\}$ and using the orthogonality and controllability that

$$E\{\varphi_{(m+1)h-1}\varphi_{(m+1)h-1}^\tau|F_{mh-1}\} \geq \sum_{i=1}^{h-1} F^iG\{E[\xi_0\xi_0^\tau]\}G^\tau F^{i\tau} \geq \delta I$$

for some $\delta > 0$, provided that h is suitably large. From this and the boundedness of $\{\varphi_k\}$ we see that condition A2 is also true. □

We now proceed to present the results of the paper.

As one would have expected, the adaptive estimator has an attractive convergence rate in the ideal noise-free, constant parameter case. This property is addressed in Theorem 1.

Theorem 1: If in (1.1), $\theta_k \equiv \theta$, $v_k \equiv 0$, and $\{\varphi_k\}$ satisfies condition A2, then for $\{\theta_n\}$ given by (1.3), (1.4), as $n \to \infty$,

1) $$E\|\hat{\theta}_n - \theta\|^2 \to 0, \qquad \text{exponentially fast, and}$$

2) $$\hat{\theta}_n \to \theta, \qquad \text{a.s., exponentially fast.}$$

The proof is given in the next section.

In the general case, the parameter variation process $\{w_i\}$ and the noise process $\{v_i\}$ may not be zero, and the boundedness of the tracking error is a natural and realistic performance criterion.

Theorem 2: For $\{\theta_i\}$ given by (1.3), (1.4), if conditions A1 and A2 hold, then

1) $$\limsup_{n\to\infty} E\|\hat{\theta}_n - \theta_n\|^2 \leq A[\sigma_r]^{2/r}, \quad \text{and,}$$

2) $$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\|\hat{\theta}_i - \theta_i\| \leq B[\mu_4]^{1/4}, \qquad \text{a.s.}$$

where σ_r, μ_4, and r are defined in condition A1, and A and B are finite deterministic constants.

We remark that in Theorem 2, the constants A and B are functions of p, Q, r, h, and δ, the precise expressions may be found in the proof (see the next section).

As a simple example, let us take $\theta_o = 0$ and assume that $\{w_k, \iota_k\}$ is a nondegenerate i.i.d. sequence with zero mean and fifth moment. Then it is obvious that condition A1 holds and that that

$$E\|\theta_n\|^2 = nE\|w_0\|^2 \to \infty, \qquad \text{as } n \to \infty.$$

Hence, Theorem 2 implies that the adaptive algorithm (1.3), (1.4) can indeed perform the nontrivial task of tracking rapidly varying unknown parameters.

III. Analysis

We first prove some lemmas.

Lemma 1: Let $Q_k = P_k - (P_k\varphi_k\varphi_k^\tau P_k)/(R+\varphi_k^\tau P_k\varphi_k)$. Then for any $m \geq 0$ and any $k \in [mh, (m+1)h]$, the following inequality holds:

$$\operatorname{tr}[Q_k]^4 \leq \operatorname{tr}[P_{mh}]^4 - \frac{\operatorname{tr}[(P_{mh}+hQ)^5\varphi_k\varphi_k^\tau]}{R+\varphi_k^\tau(P_{mh}+hQ)\varphi_k} + O(\operatorname{tr}[P_{mh}]^3) + O(1).$$

Proof: We will need the following fact. For any nonnegative definite matrices G and H, if $G \leq H$, then

$$\operatorname{tr}G^4 \leq \operatorname{tr}GH^3 \leq \operatorname{tr}H^4. \tag{3.1}$$

The proof follows from the following chain of inequalities:

$$\begin{aligned}\operatorname{tr}G^4 &= \operatorname{tr}G^{3/2}GG^{3/2} \leq \operatorname{tr}G^{3/2}HG^{3/2} = \operatorname{tr}H^{1/2}GGGH^{1/2}\\ &\leq \operatorname{tr}H^{1/2}GHGH^{1/2} = \operatorname{tr}G^{1/2}HGHG^{1/2}\\ &\leq \operatorname{tr}G^{1/2}HHHG^{1/2} \leq \operatorname{tr}GH^3\\ &\leq \operatorname{tr}H^{3/2}GH^{3/2} \leq \operatorname{tr}H^4.\end{aligned}$$

Now, by the matrix inverse formula, it follows that for any $k \geq 0$:

$$Q_k = P_k - \frac{P_k\varphi_k\varphi_k^\tau P_k}{R+\varphi_k^\tau P_k\varphi_k} = [(P_k)^{-1} + R^{-1}\varphi_k\varphi_k^\tau]^{-1} \geq 0. \tag{3.2}$$

Note also that by (1.4), for any $k \in [mh, (m+1)h]$,

$$P_k \leq P_{k-1} + Q \leq \cdots\cdots \leq P_{mh} + hQ.$$

Hence by this, (3.1), and (3.2) we have

$$\begin{aligned}\operatorname{tr}[Q_k]^4 &= \operatorname{tr}\{[(P_k)^{-1} + R^{-1}\varphi_k\varphi_k^\tau]^{-1}\}^4\\ &\leq \operatorname{tr}\{[(P_{mh}+hQ)^{-1} + R^{-1}\varphi_k\varphi_k^\tau]^{-1}\}^4\\ &\leq \operatorname{tr}\{[(P_{mh}+hQ)^{-1} + R^{-1}\varphi_k\varphi_k^\tau]^{-1}[P_{mh}+hQ]^3\}\\ &= \operatorname{tr}\left\{[P_{mh}+hQ]^3\left[(P_{mh}+hQ) - \frac{(P_{mh}+hQ)\varphi_k\varphi_k^\tau(P_{mh}+hQ)}{R+\varphi_k^\tau(P_{mh}+hQ)\varphi_k}\right]\right\}\\ &= \operatorname{tr}[P_{mh}+hQ]^4 - \frac{\operatorname{tr}[(P_{mh}+hQ)^5\varphi_k\varphi_k^\tau]}{R+\varphi_k^\tau(P_{mh}+hQ)\varphi_k}.\end{aligned} \tag{3.3}$$

By Holder's inequality, we know that for any p-dimensional nonnegative definite matrix G,

$$\operatorname{tr}G \leq \{p\}^{1/2}\{\operatorname{tr}G^2\}^{1/2}, \quad \operatorname{tr}G^2 \leq \{p\}^{1/3}\{\operatorname{tr}G^3\}^{2/3}.$$

Therefore, by a direct expansion it is easy to show that

$$\operatorname{tr}[P_{mh}+hQ]^4 = \operatorname{tr}[P_{mh}]^4 + O(\operatorname{tr}[P_{mh}]^3) + O(1).$$

Then the result of the lemma follows from this and (3.3). □

Lemma 2: Under condition A2,

$$\sup_k E\|P_k\|^4 < \infty.$$

Proof: Let us first observe the following facts. For any p-dimensional nonnegative definite matrix G,

$$(\operatorname{tr}G)\operatorname{tr}[G]^4 \leq p^4\operatorname{tr}[G]^5 \text{and } \operatorname{tr}G^3 \leq p^{1/4}\{\operatorname{tr}[G]^4\}^{3/4}. \tag{3.4}$$

The second inequality follows from the Holder inequality; while for the first one, we have by letting $\lambda_i, i = 1,\cdots,p$ be the eigenvalues of G,

$$\begin{aligned}(\operatorname{tr}G)\operatorname{tr}[G]^4 &= \left(\sum_{i=1}^p\lambda_i\right)\left[\sum_{i=1}^p(\lambda_i)^4\right]\\ &\leq \left(\sum_{i=1}^p\lambda_i\right)^5 \leq p^4\sum_{i=1}^p(\lambda_i)^5 = p^4\operatorname{tr}[G]^5.\end{aligned}$$

Now, let us consider the following Lyapunov function:

$$T_m = \sum_{k=(m-1)h}^{mh-1}\operatorname{tr}(P_{k+1})^4, \qquad m \geq 1.$$

By (1.4) and Lemma 1, we have

$$\begin{aligned}T_{m+1} &= \sum_{k=mh}^{(m+1)h-1}\operatorname{tr}(P_{k+1})^4 = \sum_{k=mh}^{(m+1)h-1}\operatorname{tr}(Q_k+Q)^4\\ &\leq \sum_{k=mh}^{(m+1)h-1}\{\operatorname{tr}[Q_k]^4 + O(\operatorname{tr}[Q_k]^3) + O(1)\}\\ &\leq \sum_{k=mh}^{(m+1)h-1}\left\{\operatorname{tr}[P_{mh}]^4 - \frac{\operatorname{tr}[(P_{mh}+hQ)^5\varphi_k\varphi_k^\tau]}{R+\varphi_k^\tau(P_{mh}+hQ)\varphi_k} + O(\operatorname{tr}[P_{mh}]^3) + O(1)\right\}\\ &\quad + O\left(\sum_{k=mh}^{(m+1)h-1}\operatorname{tr}[P_{mh}+hQ]^3\right) + O(1)\\ &\leq h\operatorname{tr}[P_{mh}]^4 - \frac{1}{R+\lambda_{\max}(P_{mh}+hQ)}\\ &\quad\cdot\operatorname{tr}\left\{(P_{mh}+hQ)^5\sum_{k=mh}^{(m+1)h-1}\frac{\varphi_k\varphi_k^\tau}{1+\|\varphi_k\|^2}\right\}\\ &\quad + O(\operatorname{tr}[P_{mh}]^3) + O(1).\end{aligned}$$

Thus, by taking conditional expectations and using (3.4) and the fact that $(x/(R+x))$ is an increasing function of $x \geq 0$, we obtain

$$\begin{aligned}E[T_{m+1}|F_{mh-1}] &\leq h\operatorname{tr}[P_{mh}]^4 - \frac{\delta\operatorname{tr}(P_{mh}+hQ)}{p^4[R+\lambda_{\max}(P_{mh}+hQ)]}\\ &\quad\cdot\operatorname{tr}(P_{mh}+hQ)^4 + O(\operatorname{tr}[P_{mh}]^3) + O(1)\\ &\leq h\operatorname{tr}[P_{mh}]^4 - \frac{\delta h\|Q\|}{p^4(R+h\|Q\|)}\operatorname{tr}[P_{mh}]^4\\ &\quad + O(\operatorname{tr}[P_{mh}]^3) + O(1)\\ &= \left(1 - \frac{\delta\|Q\|}{p^4(R+h\|Q\|)}\right)h\operatorname{tr}[P_{mh}]^4\\ &\quad + O(\operatorname{tr}[P_{mh}]^3) + O(1).\end{aligned} \tag{3.5}$$

However, it is evident that

$$\begin{aligned}h\operatorname{tr}[P_{mh}]^4 &= \sum_{k=(m-1)h}^{mh-1}\operatorname{tr}(P_{mh})^4\\ &\leq \sum_{k=(m-1)h}^{mh-1}\operatorname{tr}[P_{k+1}+(mh-k)Q]^4\\ &\leq T_m + O\left(\sum_{k=(m-1)h}^{mh-1}\operatorname{tr}[P_{k+1}]^3\right) + O(1).\end{aligned}$$

Again by invoking (3.4) and the Holder inequality, it is easy to conclude from this that

$$h\,\mathrm{tr}\,[P_{mh}]^4 \le T_m + O(\{T_m\}^{3/4}) + O(1)$$

substituting this into (3.5), it follows that:

$$E[T_{m+1}|F_{mh-1}] \le \left(1 - \frac{\delta\|Q\|}{p^4(R+h\|Q\|)}\right)T_m + O(\{T_m\}^{3/4}) + O(1). \quad (3.6)$$

Applying the following elementary inequality:

$$x^{3/4} \le \epsilon x + \left(\frac{3}{4\epsilon}\right)^3, \qquad \forall x \ge 0, \forall \epsilon > 0$$

for appropriately small ϵ to (3.6) we get

$$E[T_{m+1}|F_{mh-1}] \le \left(1 - \frac{\delta\|Q\|}{2p^4(R+h\|Q\|)}\right)T_m + O(1). \quad (3.7)$$

Consequently,

$$ET_{m+1} \le \left(1 - \frac{\delta\|Q\|}{2p^4(R+h\|Q\|)}\right)ET_m + O(1).$$

From this it is obvious that

$$\sup_m ET_m < \infty.$$

Hence, the assertion of Lemma 2 is true. □

We remark that if in condition A2, the conditional expectation $E\{\cdot|F_{mh-1}\}$ is replaced by the nonconditional expectation $E\{\cdot\}$, then the result of Lemma 2 may not hold unless additional conditions are imposed. This can be illustrated by simply taking $\varphi_k = \varphi$, where φ is a random vector satisfying $E\varphi\varphi^\tau > 0$. In this case, it is easy to see that $E[\varphi\varphi^\tau/(1+\|\varphi\|^2)] > 0$. Furthermore, by (1.4) and (3.2) it is evident that

$$P_{k+1} = [(P_k)^{-1} + R^{-1}\varphi\varphi^\tau]^{-1} + Q \ge [(P_{k-1})^{-1} + R^{-1}\varphi\varphi^\tau]^{-1} + Q = P_k$$

provided that $P_k \ge P_{k-1} > 0$. Therefore, if P_o satisfies $0 < P_o \le Q$, then the sequence $\{P_k\}$ is monotonically increasing. Let $P = \lim_{k\to\infty} P_k$, if $\mathrm{tr}\,P < \infty$, then by taking limits on both sides of (1.4) we have $Q = P\varphi\varphi^\tau P/(R + \varphi^\tau P\varphi)$, consequently $\mathrm{tr}\,P = \infty$ when $\mathrm{rank}(Q) > 1$. Hence, by the monotone convergence theorem, $\lim_{k\to\infty} E\|P_k\| = \infty$, and so $\lim_{k\to\infty} E\|P_k\|^4 = \infty$.

Lemma 3: Under condition A2,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \|P_k\|^2 < \infty, \qquad \text{a.s.}$$

Proof: Let Q_k be the same as in Lemma 1; it follows from a similar argument as used in Lemma 1 that for any $k \in [mh, (m+1)h]$, $m \ge 0$

$$\mathrm{tr}\,[Q_k]^2 \le \mathrm{tr}\,[P_{mh}]^2 - \frac{\mathrm{tr}\,[(P_{mh}+hQ)^3\varphi_k\varphi_k^\tau]}{R + \varphi_k^\tau(P_{mh}+hQ)\varphi_k} + O(\mathrm{tr}\,[P_{mh}]) + O(1)$$

consequently, similar to the proof of (3.7) in Lemma 2, we have

$$E[M_{m+1}|F_{mh-1}] \le \left(1 - \frac{\delta\|Q\|}{2p^2(R+h\|Q\|)}\right)M_m + O(1), \qquad \forall m \ge 0 \quad (3.8)$$

where

$$M_m = \sum_{k=(m-1)h}^{mh-1} \mathrm{tr}\,(P_{k+1})^2.$$

Let us denote

$$g_{m+1} = M_{m+1} - E[M_{m+1}|F_{mh-1}], \qquad m \ge 0 \quad (3.9)$$

then it is easy to see that $\{g_m, F_{mh-1}, m \ge 0\}$ is a martingale difference sequence, and satisfies

$$\sup_m E[g_m]^2 < \infty$$

by Lemma 2. Hence, by the martingale convergence theorem (Chow [18]), we know that

$$\sum_{k=1}^{\infty} \frac{g_k}{k} \qquad \text{converges almost surely.}$$

Therefore, by the Kronecker lemma we have

$$\frac{1}{n}\sum_{k=1}^{n-1} g_k \to 0, \qquad \text{a.s. as } n \to \infty. \quad (3.10)$$

Now by (3.8) and (3.9) it follows that

$$M_{m+1} = E[M_{m+1}|F_{mh-1}] + g_{m+1} \le \left(1 - \frac{\delta\|Q\|}{2p^2(R+h\|Q\|)}\right)M_m + O(1) + g_{m+1}.$$

Summing up from 0 to $n-1$ we obtain

$$M_n \le M_0 - \frac{\delta\|Q\|}{2p^2(R+h\|Q\|)}\sum_{m=0}^{n-1} M_m + O(n) + \sum_{m=0}^{n-1} g_{m+1}$$

and so

$$\frac{1}{n}\sum_{m=0}^{n-1} M_m \le \frac{2p^2(R+h\|Q\|)}{\delta\|Q\|}\left\{\frac{M_0}{n} + \frac{1}{n}\sum_{m=0}^{n-1} g_{m+1} + O(1)\right\}.$$

Thus, by (3.10) we have

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{m=0}^{n-1} M_m < \infty, \qquad \text{a.s.}$$

From this, it is easy to conclude the desired result. □

The following result plays a key role in the paper. In Lemma 5 this result will be somewhat generalized, and then in the proof of Theorem 1 below, the modified result will be used to connect the excitation condition A2 with the stability of the homogeneous part of the recursion (1.3).

Lemma 4: Let $\{a_k, F_k\}$ be an adapted sequence, $a_k \ge 1$, $\forall k \ge 0$, $Ea_o < \infty$, and

$$E[a_k|F_{k-1}] \le \alpha a_{k-1} + \beta, \quad \forall k \ge 1, \ 0 < \alpha < 1, \ 0 < \beta < \infty.$$

Then there exist constants $\gamma \in (0, 1)$, $M < \infty$ such that

$$E\prod_{k=m}^{n}\left(1 - \frac{1}{a_k}\right) \le M\gamma^{n-m+1}, \qquad \forall n \ge m, \forall m \ge 0.$$

Proof: Without loss of generality assume that $\beta \ge 1$. Let us take a constant $c > \beta/\alpha$, so that

$$b_k \triangleq \frac{a_k + c}{c(1-\alpha)+\beta} > \frac{c}{c(1-\alpha)+\beta} > 1.$$

It is immediately verified that

$$E[b_k|F_{k-1}] \le \alpha b_{k-1} + 1. \tag{3.11}$$

Now, for any $n \ge m$, define a sequence $\{x_k,\ m \le k \le n\}$ recursively as follows:

$$x_k = \left(1 - \frac{1}{b_k}\right) x_{k-1}, \qquad x_{m-1} = 1. \tag{3.12}$$

Then x_k is F_k-measurable, and

$$b_k x_k = b_k x_{k-1} - x_{k-1}.$$

Consequently, by (3.11),

$$\begin{aligned} E[b_k x_k|F_{k-1}] &= E[b_k|F_{k-1}]x_{k-1} - x_{k-1} \\ &\le (\alpha b_{k-1} + 1)x_{k-1} - x_{k-1} \\ &= \alpha b_{k-1} x_{k-1}. \end{aligned}$$

Note that by (3.11), $Eb_m \le Eb_o + 1/(1-\alpha)$, $\forall m \ge 0$, so we have

$$\begin{aligned} E[b_n x_n] &\le \alpha E[b_{n-1}x_{n-1}] \le \cdots \le \alpha^{n-m+1} E[b_{m-1}x_{m-1}] \\ &= \alpha^{n-m+1} E[b_{m-1}] \le \left[Eb_o + \frac{1}{1-\alpha}\right] \alpha^{n-m+1} \end{aligned}$$

thus by (3.12) and the fact that $b_k > 1$,

$$\begin{aligned} E\prod_{k=m}^{n}\left(1 - \frac{1}{b_k}\right) &= Ex_n \le E[b_n x_n] \\ &\le \left[Eb_o + \frac{1}{1-\alpha}\right] \alpha^{n-m+1}. \end{aligned} \tag{3.13}$$

Next, by standard methods in calculus, it is easy to verify the following inequality:

$$1 - x \le (1 - dx)^{(1-r)/d}, \qquad 0 \le dx \le r < 1,\ d > 1.$$

By this, the Holder inequality, and (3.13), we finally obtain ($d = c(1-\alpha) + \beta$, $r = 1 - \alpha + (\beta/c)$)

$$\begin{aligned} E\prod_{k=m}^{n}\left(1 - \frac{1}{a_k}\right) &\le E\prod_{k=m}^{n}\left(1 - \frac{1}{a_k + c}\right) \\ &\le E\prod_{k=m}^{n}\left\{\left(1 - \frac{c(1-\alpha)+\beta}{a_k + c}\right)\right\}^{(1-r)/d} \\ &= E\left\{\prod_{k=m}^{n}\left(1 - \frac{1}{b_k}\right)\right\}^{(1-r)/d} \\ &\le \left\{E\prod_{k=m}^{n}\left(1 - \frac{1}{b_k}\right)\right\}^{(1-r)/d} \\ &\le \left[Eb_o + \frac{1}{1-\alpha}\right]^{(1-r)/d} [\alpha^{(1-r)/d}]^{n-m+1}. \end{aligned}$$

□

Lemma 5: Let $\{a_k, F_k\}$ be an adapted sequence, $a_k \ge 1$, $\forall k \ge 0$. If for some integer $h > 0$, and constants $0 < \alpha < 1$, $\beta < \infty$,

$$E[S_{k+1}|F_{kh-1}] \le \alpha S_k + \beta, \qquad \forall k \ge 0,\ ES_o < \infty$$

where

$$S_k = \sum_{j=(k-1)h}^{kh-1} a_j$$

then there exist constants $\gamma \in (0, 1)$ and $M < \infty$, such that

$$E\prod_{k=m}^{n}\left(1 - \frac{1}{a_k}\right) \le M\gamma^{n-m+1}, \qquad \forall n \ge m,\ \forall m \ge 0.$$

Proof: By Lemma 4 we know that there exist constants $0 < \gamma_o < 1$ and $M_o < \infty$, such that

$$E\sum_{k=m}^{n}\left(1 - \frac{1}{S_k}\right) \le M_o(\gamma_o)^{n-m+1}, \qquad \forall n \ge m,\ \forall m \ge 0.$$

Clearly, for the result of the lemma we need only to consider the case of $n - m > h$. Let i and j be two integers such that

$$ih \le n < (i+1)h, \qquad (j-1)h < m \le jh.$$

It then follows that

$$\begin{aligned} E\prod_{k=m}^{n}\left(1 - \frac{1}{a_k}\right) &\le E\prod_{k=jh}^{ih}\left(1 - \frac{1}{a_k}\right) \le E\prod_{t=j}^{i}\left(1 - \frac{1}{a_{th}}\right) \\ &\le E\prod_{t=j}^{i}\left(1 - \frac{1}{S_{t+1}}\right) \le M_o(\gamma_o)^{i-j+1} \\ &= M_o[(\gamma_o)^{1/h}]^{h(i-j)+h} \\ &\le M_o[(\gamma_o)^{1/h}]^{n-h-m-h+h} \\ &= [M_o(\gamma_o)^{-1-(1/h)}][(\gamma_o)^{1/h}]^{n-m+1}. \end{aligned}$$

□

Let us now denote $\tilde{\theta}_k = \theta_k - \hat{\theta}_k$, and consider the following stochastic Lyapunov function V_k:

$$V_k = \tilde{\theta}_k^\tau P_k^{-1} \tilde{\theta}_k. \tag{3.14}$$

We have the following.

Lemma 6: For any $k \ge 0$,

$$V_{k+1} \le V_k - \frac{V_k}{4 + a\,\mathrm{tr}(P_k)} + O(\|P_k\|\{\|v_k\|^2 + \|w_{k+1}\|^2\})$$

where $a = 2\|Q^{-1}\|$.

Proof: Let us denote

$$K_k = \frac{P_k\varphi_k}{R + \varphi_k^\tau P_k \varphi_k}, \qquad G_k = I - K_k\varphi_k^\tau$$

and rewrite (1.4) as

$$P_{k+1} = G_k P_k G_k^\tau + RK_k K_k^\tau + Q. \tag{3.15}$$

Note that by (1.1)–(1.3) the error equation is

$$\tilde{\theta}_{k+1} = G_k\tilde{\theta}_k + z_{k+1}, \qquad z_{k+1} = -K_k v_k + w_{k+1}. \tag{3.16}$$

So we have by (3.14),

$$\begin{aligned} V_{k+1} &= [G_k\tilde{\theta}_k + z_{k+1}]^\tau [P_{k+1}]^{-1} [G_k\tilde{\theta}_k + z_{k+1}] \\ &= \tilde{\theta}_k^\tau G_k^\tau P_{k+1}^{-1} G_k \tilde{\theta}_k + 2z_{k+1}^\tau P_{k+1}^{-1} G_k\tilde{\theta}_k + z_{k+1}^\tau P_{k+1}^{-1} z_{k+1}. \end{aligned} \tag{3.17}$$

By (3.15) and the matrix inverse formula, we know that

$$\begin{aligned} G_k^\tau P_{k+1}^{-1} G_k &= G_k^\tau \{G_k P_k G_k^\tau + K_k R K_k^\tau + Q\}^{-1} G_k \\ &= P_k^{-1} - [P_k + P_k G_k^\tau (K_k R K_k^\tau + Q)^{-1} G_k P_k]^{-1} \\ &= P_k^{-1/2}\{I - [I + (P_k)^{1/2} G_k^\tau (K_k R K_k^\tau + Q)^{-1} \\ &\quad \cdot G_k (P_k)^{1/2}]^{-1}\} P_k^{-1/2} \\ &\le \{1 - [1 + \|(K_k R K_k^\tau + Q)^{-1} G_k P_k G_k^\tau\|]^{-1}\} P_k^{-1} \\ &\le \{1 - [1 + \|(K_k R K_k^\tau + Q)^{-1} P_{k+1}\|]^{-1}\} P_k^{-1} \\ &\le \{1 - [1 + \|Q^{-1}(P_k + Q)\|]^{-1}\} P_k^{-1} \\ &\le P_k^{-1} - \frac{1}{2 + \|Q^{-1}\| \, \|P_k\|} P_k^{-1}. \end{aligned} \tag{3.18}$$

Putting this into (3.17) we get

$$V_{k+1} \le V_k - \frac{1}{2 + \|Q^{-1}\| \, \|P_k\|} V_k + 2 z_{k+1}^\tau P_{k+1}^{-1} G_k \tilde{\theta}_k + z_{k+1}^\tau P_{k+1}^{-1} z_{k+1}. \tag{3.19}$$

By the elementary inequality $2|xy| \le x^2 + y^2$, it follows that:

$$\begin{aligned} 2|z_{k+1}^\tau P_{k+1}^{-1} G_k \tilde{\theta}_k| &\le 2\|z_{k+1}^\tau P_{k+1}^{-1/2}\| \, \|P_{k+1}^{-1/2} G_k \tilde{\theta}_k\| \\ &\le 2 z_{k+1}^\tau P_{k+1}^{-1} z_{k+1} (2 + \|Q^{-1}\| \, \|P_k\|) \\ &\quad + \frac{\tilde{\theta}_k^\tau G_k^\tau P_{k+1}^{-1} G_k \tilde{\theta}_k}{2(2 + \|Q^{-1}\| \, \|P_k\|)}. \end{aligned} \tag{3.20}$$

Recall that by (3.14) and (3.18),

$$\tilde{\theta}_k^\tau G_k^\tau P_{k+1}^{-1} G_k \tilde{\theta}_k \le V_k. \tag{3.21}$$

By (3.15) $P_{k+1} \ge R K_k K_k^\tau + Q$, then by (3.16) it follows that:

$$\begin{aligned} z_{k+1}^\tau P_{k+1}^{-1} z_{k+1} & \left\| P_{k+1}^{-1/2}(-K_k v_k + w_{k+1}) \right\|^2 \\ &\le O(K_k^\tau P_{k+1}^{-1} K_k \|v_k\|^2) + O(\|w_{k+1}\|^2) \\ &\le O(\|v_k\|^2 + \|w_{k+1}\|^2). \end{aligned} \tag{3.22}$$

Finally, substituting (3.20)-(3.22) into (3.19) we get

$$V_{k+1} \le V_k - \frac{V_k}{2(2 + \|Q^{-1}\| \, \|P_k\|)} + O(\|P_k\| \{\|v_k\|^2 + \|w_{k+1}\|^2\}).$$

Hence, the result of Lemma 6 is true. □

Proof of Theorem 1: Similar to the proof of (3.7) or (3.8) it can be shown that

$$E[S_{m+1} | F_{mh-1}] \le \left(1 - \frac{\delta \|Q\|}{p(R + h\|Q\|)}\right) S_m + O(1), \quad \forall m \ge 0$$

where

$$S_m = \sum_{k=(m-1)h}^{mh-1} \operatorname{tr}(P_{k+1}).$$

From this, it is easy to see that $a_k \triangleq 4 + a \operatorname{tr}(P_{k+1})$ satisfies the conditions in Lemma 5, therefore, if $\Phi(n, k)$ is defined as

$$\Phi(n+1, k) = \left(1 - \frac{1}{4 + a \operatorname{tr}(P_n)}\right) \Phi(n, k), \quad \Phi(k, k) = 1, \ \forall n \ge k \ge 0 \tag{3.23}$$

then

$$E\Phi(n+1, k) \le M \gamma^{n-k+1}, \quad \forall n \ge k \ge 0, \ 0 < \gamma < 1, \ M < \infty. \tag{3.24}$$

Now, under the conditions of Theorem 1, it follows from Lemma 6 that:

$$V_{n+1} \le \Phi(n+1, 0) V_0$$

so by the Holder inequality

$$\begin{aligned} E[V_n]^{4/3} &\le O(E[\Phi(n, 0)]^{4/3} \|\tilde{\theta}_0\|^{8/3}) \\ &\le O(\{E[\Phi(n, 0)]^{12}\}^{1/9} \{E\|\tilde{\theta}_o\|^3\}^{8/9}) \\ &= O(\{E\Phi(n, 0)\}^{1/9}) \underset{n \to \infty}{\longrightarrow} 0, \quad \text{exponentially fast.} \end{aligned}$$

From this and Lemma 2, it follows that:

$$\begin{aligned} E\|\tilde{\theta}_n\|^2 &\le E\|P_n^{1/2}\|^2 \|P_n^{-1/2} \tilde{\theta}_n\|^2 \\ &\le \{E\|P_n\|^4\}^{1/4} \{E[V_n]^{4/3}\}^{3/4} \to 0, \end{aligned}$$

exponentially fast.

This proves the first assertion 1), while the second assertion can be easily proved by using 1) and the Borel-Cantelli Lemma. □

Proof of Theorem 2: With $\Phi(n, k)$ defined as in (3.23), it follows from Lemma 6 that

$$V_n \le \Phi(n, 0) V_0 + O\left(\sum_{k=0}^{n-1} \Phi(n, k) \|P_k\| [\|v_k\|^2 + \|w_{k+1}\|^2]\right)$$

so by the Minkowski inequality we have

$$\begin{aligned} \{E[V_n]^{4/3}\}^{3/4} &\le \{E[\Phi(n, 0) V_0]^{4/3}\}^{3/4} \\ &\quad + O\left(\sum_{k=0}^{n-1} \{E[\Phi(n, k) \|P_k\| (\|v_k\|^2 + \|w_{k+1}\|^2)]^{4/3}\}^{3/4}\right). \end{aligned} \tag{3.25}$$

Now, by the Holder inequality, Lemma 2, condition A1, and the fact that $\Phi(n, k) \le 1$, we know that ($q = 3r/2(r-4)$)

$$\begin{aligned} & E[\Phi(n, k) \|P_k\| (\|v_k\|^2 + \|w_{k+1}\|^2)]^{4/3} \\ &\le 2^{1/3} E[\Phi(n, k)]^{4/3} \|P_k\|^{4/3} (\|v_k\|^{8/3} + \|w_{k+1}\|^{8/3}) \\ &\le O(\{E[\Phi(n, k)]^{4q/3}\}^{1/q} \{E\|P_k\|^4\}^{1/3} \\ &\quad \cdot \{E[\|v_k\|^r + \|w_{k+1}\|^r]\}^{8/3r}) \\ &\le O(\{E\Phi(n, k)\}^{1/q} \{E[\|v_k\|^r + \|w_{k+1}\|^r]\}^{8/3r}) \\ &\le O([\sigma_r]^{8/3r} \{E\Phi(n, k)\}^{1/q}). \end{aligned}$$

Hence, it follows from (3.24) and (3.25) that

$$\limsup_{n \to \infty} E[V_n]^{4/3} \le O([\sigma_r]^{8/3r}).$$

Therefore, we have

$$\begin{aligned} \limsup_{n \to \infty} E[\|\tilde{\theta}_n\|^2] &\le \limsup_{n \to \infty} E\|P_n^{1/2}\|^2 \|P_n^{-1/2} \tilde{\theta}_n\|^2 \\ &= \limsup_{n \to \infty} E\|P_n\| V_n \\ &\le \limsup_{n \to \infty} \{E\|P_n\|^4\}^{1/4} \{E[V_n]^{4/3}\}^{3/4} \\ &\le O([\sigma_r]^{2/r}). \end{aligned}$$

We now proceed to prove the second conclusion 2) of the theorem.

By Lemma 6, it is evident that

$$\sum_{k=0}^{n-1} \frac{V_k}{4 + a\mathrm{tr}\,(P_k)} = O(1) + O\left(\sum_{k=0}^{n-1} \|P_k\| \{\|v_k\|^2 + \|w_{k+1}\|^2\}\right)$$

so by the Schwarz inequality, condition A1, and Lemma 3,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \frac{V_k}{4 + a\mathrm{tr}\,(P_k)} \le O(\{\mu_4\}^{1/2}).$$

Consequently, by this and Lemma 3 it follows that ($b = 1 + a$)

$$\begin{aligned}\sum_{k=0}^{n-1} \|\tilde{\theta}_k\| &= \sum_{k=0}^{n-1} \frac{\|\tilde{\theta}_k\|}{[4 + b(\mathrm{tr}\,P_k)^2]^{1/2}} [4 + b(\mathrm{tr}\,P_k)^2]^{1/2} \\ &\le \left\{\sum_{k=0}^{n-1} \frac{\|\tilde{\theta}_k\|^2}{4 + b(\mathrm{tr}\,P_k)^2}\right\}^{1/2} \left\{\sum_{k=0}^{n-1} [4 + b(\mathrm{tr}\,P_k)^2]\right\}^{1/2} \\ &\le O\left(n^{1/2}\left\{\sum_{k=0}^{n-1} \frac{\|\tilde{\theta}_k\|^2}{[4 + a(\mathrm{tr}\,P_k)](\mathrm{tr}\,P_k)}\right\}^{1/2}\right) \\ &\le O\left(n^{1/2}\left\{\sum_{k=0}^{n-1} \frac{V_k}{4 + a\mathrm{tr}\,(P_k)}\right\}^{1/2}\right) \\ &= O(\{\mu_4\}^{1/4} n) \qquad \square\end{aligned}$$

IV. Conclusion

Most of the work done in the area of system identification is concerned with the estimation of constant parameters. In the time-varying case, few precise theoretical results are available, although various estimation methods have already been proposed. Among these methods, the Kalman filtering algorithm, due to its optimality in some sense, is one of the most attractive estimation algorithms (see, e.g., Ljung [19]), and has applications in stochastic adaptive control (see, Meyn and Caines [20], Guo and Meyn [21]).

In this paper we have presented a theoretical analysis of the Kalman filter based adaptive estimator applied to a time-varying stochastic linear regression model. In particular, by introducing a suitable excitation condition, we have shown that the parameter tracking errors $\limsup_{n\to\infty} E\|\hat{\theta}_n - \theta_n\|^2$ and $\limsup_{n\to\infty} (1/n)\sum_{i=0}^{n-1} \|\hat{\theta}_i - \theta_i\|$ are small when the parameter variation process $\{w_i\}$ and the noise process $\{v_i\}$ are small. It is worth noting that as no assumptions of stationarity or independence are imposed on the regression vector $\{\varphi_k\}$, the results of this paper do not exclude applications to feedback control systems.

Acknowledgment

The author would like to thank Dr. S. P. Meyn for his valuable discussions and comments.

References

[1] D. V. Lindley and A. F. M. Smith, "Bayes estimate for the linear model," *J. Royal Statist. Soc., B*, vol. 34, pp. 1-14, 1972.

[2] G. Kitagawa and W. Gersh, "A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 48-56, 1985.

[3] D. Q. Mayne, "Optimal nonstationary estimation of the parameters of a linear system with Gaussian inputs," *J. Electron. Conf.*, vol. 14, p. 101, 1963.

[4] K. J. Astrom and B. Wittenmark, "Problems of identification and control," *J. Math. Anal. Appl.*, vol. 34, pp. 90-113, 1971.

[5] V. Solo, "Stochastic adaptive control and martingale limit theory," *IEEE Trans. Automat. Contr.*, to be published.

[6] A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* New York: Academic, 1970.

[7] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis.* Boston, MA: M.I.T. Press, 1986

[8] L. Guo, L. G. Xia, and J. B. Moore, "Tracking randomly varying parameters—Analysis of a standard algorithm," in *Proc. 27th IEEE Conf. Decision Contr.*, Austin, TX, Dec. 1988.

[9] B. Widow *et al.*, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151-1162, 1976.

[10] A. Benveniste and G. Ruget, "A measure of the tracking capability of recursive stochastic algorithms with constant gains," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 639-649, 1982.

[11] E. Eweda and O. Macchi, "Tracking error bounds of adaptive nonstationary filtering," *Automatica*, vol. 21, no. 3, pp. 293-302, 1985.

[12] O. Macchi, "Optimization of adaptive identification for time varying filters," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 283-287, 1986.

[13] A. Benveniste, "Design of adaptive algorithms for the tracking of time-varying systems," *Int. J. Adaptive Contr. Signal Processing*, vol. 1, pp. 3-29, 1987.

[14] R. R. Bitmead and B. D. O. Anderson, "Performance of adaptive estimation algorithm in dependent random environments," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 788-794, 1980.

[15] K. H. Shi and F. Kozin, "On almost sure convergence of adaptive algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 471-474, 1986.

[16] A. Benveniste, M. Metivier, and P. Priouret, *Algorithmes Adaptatifs et Approximations Stochastiques, Techniques Stochastiques.* Paris: Masson, 1987.

[17] V. Solo, "The limiting behaviour of LMS," *IEEE Trans. Acoust., Speech, Signal Processing*, to be published.

[18] Y. S. Chow, "Local convergence of martingales and the law of large numbers," *Ann. Math. Stat.*, vol. 36, pp. 552-558, 1965.

[19] L. Ljung, "Adaptation and tracking in system identification," in *Preprints 8th IFAC/IFORS Symp. Ident. Syst. Parameter Estimation*, Beijing, 1988, vol. 1, pp. 1-10.

[20] S. P. Meyn and P. E. Caines, "A new approach to stochastic adaptive control," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 220-226, 1987.

[21] L. Guo and S. P. Meyn, "Adaptive control for time-varying systems: A combination of martingale and Markov chain techniques," *Int. J. Adaptive Contr. Signal Processing*, vol. 3, no. 1, pp. 1-14, 1989.

Lei Guo (M'88) was born in China in 1961. He received the B.S. degree in mathematics from Shandong University in 1982, and the M.S. and Ph.D. degrees in control theory from the Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China, in 1984 and 1987, respectively.

From June 1987 to June 1989, he was with the Department of Systems Engineering, Australian National University, Canberra. Presently, he is a Research Associate with the Institute of Systems Science, Academia Sinica, Beijing. His research interests are in adaptive control, system identification, time series analysis, and statistics of random processes.

Dr. Guo (with H. F. Chen) received the 1987 National Natural Science Prize of China.