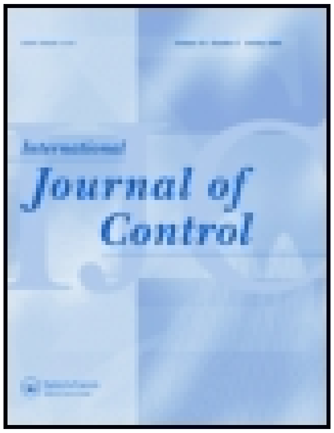

International Journal of Control
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcon20

Convergence rate of least-squares identification and adaptive control for stochastic systems†

HAN-FU CHEN [a] & LEI GUO [a]
[a] Institute of Systems Science, Academia Sinica , Beijing, People's Republic of China
Published online: 30 Jun 2010.

INT. J. CONTROL, 1986, VOL. 44, NO. 5, 1459–1476

Convergence rate of least-squares identification and adaptive control for stochastic systems†

HAN-FU CHEN‡ and LEI GUO‡

The strong consistency and the convergence rate of least-squares identification for the multidimensional ARMAX model are established under some decaying excitation conditions which are satisfied if both input and output do not grow too fast and the attenuating excitation technique is applied. The parameter-identification results are applied to adaptive-control systems with a quadratic loss function. The rate of convergence of the loss function to its minimum is also obtained.

1. Introduction

Consider a stochastic system with m- and l-dimensional output and input respectively that is driven by an m-dimensional martingale difference sequence $\{w_n, \mathscr{F}_n\}$, where $\{\mathscr{F}_n\}$ is a family of non-decreasing σ-algebras:

$$\left.\begin{aligned} A(z)y_n &= B(z)u_n + C(z)z_n, & n \geqslant 0 \\ y_n &= u_n = w_n = 0, & n < 0 \end{aligned}\right\} \tag{1}$$

where $A(z)$, $B(z)$ and $C(z)$ are matrix polynomials in the shift-back operator z:

$$A(z) = I + A_1 z + \dots + A_p z^p$$

$$B(z) = B_1 z + B_2 z^2 + \dots + B_q z^q$$

$$C(z) = I + C_1 z + \dots + C_r z^r$$

with unknown matrix coefficient

$$\theta = [-A_1 \quad \dots \quad -A_p \quad B_1 \quad \dots \quad B_q \quad C_1 \quad \dots \quad C_r]^\tau \tag{2}$$

which may be estimated by various recursive algorithms.

For consistency analysis of the estimate θ_n, the matrix $\sum_{i=1}^{n} \phi_i \phi_i^\tau$ consisting of the stochastic regressors for system (1) is of great importance. In earlier works (e.g. Ljung 1976, Moore 1978, Solo 1979) the persistent excitation condition—that the ratio of the maximum to the minimum eigenvalues of $\sum_{i=1}^{n} \phi_i \phi_i^\tau$ is bounded—is the key to guaranteeing strong consistency. In Chen and Guo (1985 a) we have shown that for strong consistency of θ_n given by the stochastic-gradient algorithm this ratio should not grow faster than $\left(\log \sum_{i=1}^{n} \|\phi_i\|^2\right)^{1+\delta}$ with $\delta > 0$, and if it does not grow faster than $\left(\log \sum_{i=1}^{n} \|\phi_i\|^2\right)^{1/4}$ then θ_n converges to θ almost surely under some reasonable

Received 7 January 1986.
Work supported by the Science Fund of the Chinese Academy of Sciences.
‡ Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China.

conditions on the system noise $C(z)w_n$ (Chen and Guo 1985 b). But, as pointed out in Chen and Guo (1985 a), this does not exclude the possibility of strong consistency of estimates given by algorithms other than the stochastic-gradient one under a growth rate faster than that mentioned above. In fact, the least-squares estimation is such an algorithm (Chen 1982, Lai and Wei 1982). Lai and Wei showed that for a single-input–single-output system with uncorrelated noise the least-squares estimate is strongly consistent if the minimum eigenvalue of $\sum_{i=1}^{n} \phi_i\phi_i^\tau$ goes to infinity faster than the logarithm of the maximum eigenvalue of $\sum_{i=1}^{n} \phi_i\phi_i^\tau$. This is probably the weakest condition for strong consistency of estimates for θ (Lai and Wei 1982).

In this paper we first give results similar to those given in Lai and Wei (1982), but for the multidimensional system and with correlated noise. Then, using these results, we show that the least-squares estimate converges to the true parameter if both the input and output do not diverge too fast and if the attenuating excitation technique is applied to the control. Finally, we apply parameter-estimation results to an adaptive-control system with quadratic cost and give the rate of convergence of the cost to its optimal value.

In the present paper we mean by the least-squares estimate the one given by the following recursive algorithm (Chen 1985):

$$\theta_{n+1} = \theta_n + a_n P_n \phi_n (y_{n+1}^\tau - \phi_n^\tau \theta_n) \tag{3}$$

$$P_{n+1} = P_n - a_n P_n \phi_n \phi_n^\tau P_n, \quad a_n = (1 + \phi_n^\tau P_n \phi_n)^{-1} \tag{4}$$

$$\phi_n^\tau = [y_n^\tau \quad \cdots \quad y_{n-p+1}^\tau \quad u_n^\tau \quad \cdots \quad u_{n-q+1}^\tau \quad y_n^\tau - \phi_{n-1}^\tau\theta_n \quad \cdots \quad y_{n-r+1}^\tau - \phi_{n-r}^\tau\theta_{n-r+1}] \tag{5}$$

with $P_0 = dI$, $d = mp + lq + mr$ for convenience and with θ_0 arbitrary.

It is clear that

$$P_n = \left(\sum_{i=0}^{n-1} \phi_i\phi_i^\tau + \frac{1}{d}I\right)^{-1} \tag{6}$$

We introduce the vector

$$\phi_n^0 = [y_n^\tau \quad \cdots \quad y_{n-p+1}^\tau \quad u_n^\tau \quad \cdots \quad u_{n-q+1}^\tau \quad w_n^\tau \quad \cdots \quad w_{n-r+1}^\tau]^\tau \tag{7}$$

which, in contrast with ϕ_n, is unavailable but is free of the estimate $\{\theta_n\}$.

Denote by $\lambda_{\min}^n$ and $\lambda_{\min}^{0n}$ the minimum eigenvalues of $\sum_{i=0}^{n-1} \phi_i\phi_i^\tau + I/d$ and $\sum_{i=0}^{n-1} \phi_i^0\phi_i^{0\tau} + I/d$ respectively, and set

$$r_n = 1 + \sum_{i=0}^{n-1} \|\phi_i\|^2, \quad r_n^0 = 1 + \sum_{i=0}^{n-1} \|\phi_i^0\|^2 \tag{8}$$

2. Convergence rate

In this section we first express the estimation error $\theta_n - \theta$ in terms of r_n and $\lambda_{\min}^n$ with no condition imposed on the growth rate of $\lambda_{\min}^n$ or r_n. Then we derive the similar expressions with $\lambda_{\min}^n$ and r_n replaced by $\lambda_{\min}^{0n}$ and r_n^0.

Theorem 1

For system (1) assume that (*a*) the driven noise $\{w_n, \mathscr{F}_n\}$ is a martingale difference sequence with respect to a family of non-decreasing σ-algebras $\{\mathscr{F}_n\}$ and

$$\sup_{n\geqslant 0} E[\|w_{n+1}\|^\beta \mid \mathscr{F}_n] < \infty \quad \text{a.s.} \quad \text{with } \beta \geqslant 2$$

(*b*) u_n is $\mathscr{F}_n$-measurable; and (*c*) $C^{-1}(z) - \frac{1}{2}I$ is strictly positive-real. Then as $n \to \infty$ the estimation error produced by (3)–(5) is expressed by

$$\text{(i)} \quad \|\theta_n - \theta\| = 0\left(\left(\frac{\log r_n}{\lambda_{\min}^n}\right)^{1/2}\right) \quad \text{a.s.} \qquad \text{if } \beta > 2 \tag{9}$$

$$\text{(ii)} \quad \|\theta_n - \theta\| = 0\left(\left(\frac{\log r_n (\log\log r_n)^c}{\lambda_{\min}^n}\right)^{1/2}\right) \quad \text{a.s.,} \quad \forall c > 1, \quad \text{if } \beta = 2 \tag{10}$$

Proof

Set

$$\tilde{\theta}_n = \theta - \theta_n \tag{11}$$

It is clear that

$$\|\tilde{\theta}_{n+1}\|^2 \leqslant \frac{1}{\lambda_{\min}^{n+1}} \operatorname{tr} \tilde{\theta}_{n+1}^\tau P_{n+1}^{-1} \tilde{\theta}_{n+1} \tag{12}$$

so to prove the theorem it is sufficient to show that

$$\operatorname{tr} \tilde{\theta}_{n+1}^\tau P_{n+1}^{-1} \tilde{\theta}_{n+1} = \begin{cases} 0(\log r_{n+1}) & \text{a.s.} \quad \text{if } \beta > 2 \\ 0(\log r_{n+1} (\log\log r_{n+1})^c) & \text{a.s.} \quad \text{if } \beta = 2 \end{cases} \tag{13}$$

The proof is divided into 3 lemmas.

Lemma 1

Under the conditions of Theorem 1 there is a constant $k_0 > 0$ such that

$$\operatorname{tr} \tilde{\theta}_{n+1}^\tau P_{n+1}^{-1} \tilde{\theta}_{n+1} \leqslant 0(1) - k_0 \sum_{i=0}^n \|\tilde{\theta}_{i+1}^\tau \phi_i\|^2 - 2\sum_{i=0}^n w_{i+1}^\tau \tilde{\theta}_{i+1}^\tau \phi_i \quad \text{a.s. for } n \geqslant 0$$

Proof

Set

$$\xi_{n+1} = y_{n+1} - w_{n+1} - \theta_{n+1}^\tau \phi_n \tag{14}$$

By (3) and (4) it is easy to see that

$$\begin{aligned} y_{n+1}^\tau - \phi_n^\tau \theta_{n+1} &= y_{n+1}^\tau - \phi_n^\tau [\theta_n + a_n P_n \phi_n (y_{n+1}^\tau - \phi_n^\tau \theta_n)] \\ &= (1 - a_n \phi_n^\tau P_n \phi_n)(y_{n+1}^\tau - \phi_n^\tau \theta_n) = a_n (y_{n+1}^\tau - \phi_n^\tau \theta_n) \end{aligned} \tag{15}$$

By (1) and (5) we have

$$\begin{aligned} C(z)\xi_{n+1} &= y_{n+1} + (C(z) - I)(y_{n+1} - \theta_{n+1}^\tau \phi_n) - \theta_{n+1}^\tau \phi_n - C(z) w_{n+1} \\ &= -(A(z) - I) y_{n+1} + B(z) u_n + (C(z) - I)(y_{n+1} - \theta_{n+1}^\tau \phi_n) - \theta_{n+1}^\tau \phi_n \\ &= \theta^\tau \phi_n - \theta_{n+1}^\tau \phi_n = \tilde{\theta}_{n+1}^\tau \phi_n \end{aligned} \tag{16}$$

Since $C^{-1}(z)-\frac{1}{2}I$ is strictly positive-real, there are constants $k_0>0$ and $k_1\geqslant 0$ such that

$$s_n \triangleq \sum_{i=0}^{n} \phi_i^\tau \tilde{\theta}_{i+1}(\xi_{i+1}-\tfrac{1}{2}(1+k_0)\tilde{\theta}_{i+1}^\tau \phi_i)+k_1 \geqslant 0 \quad \forall n \geqslant 0 \tag{17}$$

By (15) we can rewrite (3) as

$$\tilde{\theta}_{n+1}=\tilde{\theta}_n-P_n\phi_n(\xi_{n+1}^\tau+w_{n+1}^\tau) \tag{18}$$

From this and (6) it then follows that

$$\begin{aligned}
\operatorname{tr}\tilde{\theta}_{k+1}^\tau P_{k+1}^{-1}\tilde{\theta}_{k+1} &= \operatorname{tr}\tilde{\theta}_{k+1}^\tau \phi_k\phi_k^\tau\tilde{\theta}_{k+1}+\operatorname{tr}\tilde{\theta}_{k+1}^\tau P_k^{-1}\tilde{\theta}_{k+1}\\
&= \|\phi_k^\tau\tilde{\theta}_{k+1}\|^2+\operatorname{tr}[\tilde{\theta}_k-P_k\phi_k(\xi_{k+1}^\tau+w_{k+1}^\tau)]^\tau P_k^{-1}\\
&\quad\times[\tilde{\theta}_k-P_k\phi_k(\xi_{k+1}^\tau+w_{k+1}^\tau)]\\
&= \|\phi_k^\tau\tilde{\theta}_{k+1}\|^2-2(\xi_{k+1}^\tau+w_{k+1}^\tau)\tilde{\theta}_k^\tau\phi_k+\phi_k^\tau P_k\phi_k\|\xi_{k+1}+w_{k+1}\|^2\\
&\quad+\operatorname{tr}\tilde{\theta}_k^\tau P_k^{-1}\tilde{\theta}_k\\
&= \|\phi_k^\tau\tilde{\theta}_{k+1}\|^2-2(\xi_{k+1}^\tau+w_{k+1}^\tau)[\tilde{\theta}_{k+1}+P_k\phi_k(\xi_{k+1}^\tau+w_{k+1}^\tau)]^\tau\phi_k\\
&\quad+\phi_k^\tau P_k\phi_k\|\xi_{k+1}+w_{k+1}\|^2+\operatorname{tr}\tilde{\theta}_k^\tau P_k^{-1}\tilde{\theta}_k\\
&\leqslant \operatorname{tr}\tilde{\theta}_k^\tau P_k^{-1}\tilde{\theta}_k+\|\tilde{\theta}_{k+1}^\tau\phi_k\|^2-2\xi_{k+1}^\tau\tilde{\theta}_{k+1}^\tau\phi_k-2w_{k+1}^\tau\tilde{\theta}_{k+1}^\tau\phi_k\\
&= \operatorname{tr}\tilde{\theta}_k^\tau P_k^{-1}\tilde{\theta}_k-2[\phi_k^\tau\tilde{\theta}_{k+1}(\xi_{k+1}-\tfrac{1}{2}(1+k_0)\tilde{\theta}_{k+1}^\tau\phi_k)]-k_0\|\tilde{\theta}_{k+1}^\tau\phi_k\|^2\\
&\quad-2w_{k+1}^\tau\tilde{\theta}_{k+1}^\tau\phi_k
\end{aligned} \tag{19}$$

Summing both sides of (19) from 0 to n and using (17), we conclude that

$$\begin{aligned}
\operatorname{tr}\tilde{\theta}_{n+1}^\tau P_{n+1}^{-1}\tilde{\theta}_{n+1} &\leqslant \operatorname{tr}\tilde{\theta}_0^\tau P_0^{-1}\tilde{\theta}_0-2s_n+2k_1-k_0\sum_{i=0}^{n}\|\tilde{\theta}_{i+1}^\tau\phi_i\|^2-2\sum_{i=0}^{n}w_{i+1}^\tau\tilde{\theta}_{i+1}^\tau\phi_i\\
&\leqslant 0(1)-k_0\sum_{i=0}^{n}\|\tilde{\theta}_{i+1}^\tau\phi_i\|^2-2\sum_{i=0}^{n}w_{i+1}^\tau\tilde{\theta}_{i+1}^\tau\phi_i
\end{aligned}$$

□

Lemma 2

Under the conditions of Theorem 1,

$$\operatorname{tr}\tilde{\theta}_{n+1}^\tau P_{n+1}^{-1}\tilde{\theta}_{n+1}\leqslant 0(1)+0\left(\sum_{i=0}^{n}a_i\phi_i^\tau P_i\phi_i\|w_{i+1}\|^2\right) \quad \text{a.s.} \tag{20}$$

Proof

From Lemma 2 of Chen and Guo (1985 c) for any $\mathscr{F}_n$-measurable matrices M_n, we have

$$\sum_{i=0}^{n}M_iw_{i+1}=0\left(\left(\sum_{i=0}^{n}\|M_i\|^2\right)^{1/2}\left(\log\left(\sum_{i=0}^{n}\|M_i\|^2+e\right)\right)^{1/2+\eta}\right) \quad \text{a.s,} \quad \eta>0 \tag{21}$$

(see also Lai and Wei 1982). This estimate will be used in the sequel without explanation.

Set

$$\eta_n=y_{n+1}-\theta_n^\tau\phi_n-w_{n+1}$$

Obviously, η_n is $\mathscr{F}_n$-measurable and

$$\tilde{\theta}_{n+1} = \tilde{\theta}_n - a_n P_n \phi_n (w_{n+1}^\tau + \eta_n^\tau)$$

Then we estimate as follows:

$$\begin{aligned}
\left|\sum_{i=0}^n w_{i+1}^\tau \tilde{\theta}_{i+1}^\tau \phi_i\right| &= \left|\sum_{i=0}^n w_{i+1}^\tau (\tilde{\theta}_i^\tau - a_i(w_{i+1}+\eta_i)\phi_i^\tau P_i)\phi_i\right| \\
&\leqslant \sum_{i=0}^n a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 + \left|\sum_{i=0}^n w_{i+1}^\tau (\tilde{\theta}_i^\tau - a_i \eta_i \phi_i^\tau P_i)\phi_i\right| \\
&= \sum_{i=0}^n a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 + 0\left(\left[\sum_{i=0}^n \|(\tilde{\theta}_i^\tau - a_i \eta_i \phi_i^\tau P_i)\phi_i\|^2\right]^\alpha\right) \\
&= \sum_{i=0}^n a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 + 0\left(\left[\sum_{i=0}^n \|(\tilde{\theta}_{i+1}^\tau + a_i w_{i+1} \phi_i^\tau P_i)\phi_i\|^2\right]^\alpha\right) \\
&= \sum_{i=0}^n a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 + 0\left(\left[\sum_{i=0}^n \|\tilde{\theta}_{i+1}^\tau \phi_i\|^2\right]^\alpha\right) \\
&\quad + 0\left(\left[\sum_{i=0}^n (a_i \phi_i^\tau P_i \phi_i)^2 \|w_{i+1}\|^2\right]^\alpha\right)
\end{aligned}$$

whenever $\alpha \in (\frac{1}{2}, 1)$. Since $a_i \phi_i^\tau P_i \phi_i \leqslant 1$, from this we have

$$\left|\sum_{i=0}^n w_{i+1}^\tau \tilde{\theta}_{i+1}^\tau \phi_i\right| = 0\left(\left[\sum_{i=0}^n \|\tilde{\theta}_{i+1}^\tau \phi_i\|^2\right]^\alpha\right) + 0\left(\sum_{i=0}^n a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2\right) \tag{22}$$

and by Lemma 1 and (22) it follows that

$$\begin{aligned}
&\operatorname{tr} \tilde{\theta}_{n+1}^\tau P_{n+1}^{-1} \tilde{\theta}_{n+1} \\
&\quad \leqslant 0(1) - k_0 \sum_{i=0}^n \|\tilde{\theta}_{i+1}^\tau \phi_i\|^2 + 0\left(\left[\sum_{i=0}^n \|\tilde{\theta}_{i+1}^\tau \phi_i\|^2\right]^\alpha\right) + 0\left(\sum_{i=0}^n a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2\right)
\end{aligned} \tag{23}$$

Noting that $\alpha < 1$ and $k_0 > 0$, (20) follows from (23). □

Lemma 3

Under the conditions of Theorem 1, (13) is true.

Proof

By Lemma 2, we only need to estimate the last term of (20). For a matrix X we denote det X by $|X|$. From (6) we have

$$P_{i+1}^{-1} = P_i^{-1} + \phi_i \phi_i^\tau = P_i^{-1}(I + P_i \phi_i \phi_i^\tau)$$
$$|P_{i+1}^{-1}| = |P_i^{-1}|\,|I + P_i \phi_i \phi_i^\tau| = |P_i^{-1}|(1 + \phi_i^\tau P_i \phi_i)$$

Hence

$$\phi_i^\tau P_i \phi_i = \frac{|P_{i+1}^{-1}| - |P_i^{-1}|}{|P_i^{-1}|} \tag{24}$$

Then, from the definition of a_i and (24) we see that

$$\sum_{i=0}^{n} a_i \phi_i^\tau P_i \phi_i = \sum_{i=0}^{n} \frac{|P_{i+1}^{-1}| - |P_i^{-1}|}{|P_{i+1}^{-1}|} = \sum_{i=0}^{n} \int_{|P_i^{-1}|}^{|P_{i+1}^{-1}|} \frac{dx}{|P_{i+1}^{-1}|}$$

$$\leqslant \int_{|P_0^{-1}|}^{|P_{n+1}^{-1}|} \frac{dx}{x} = \log |P_{n+1}^{-1}| + d \log d \tag{25}$$

since $P_0^{-1} = I/d$.

A similar calculation leads to the conclusions

$$\sum_{i=0}^{\infty} \frac{a_i \phi_i^\tau P_i \phi_i}{\log^c |P_{i+1}^{-1}|} < \infty \quad \text{for any } c > 1 \tag{26}$$

$$\sum_{i=0}^{\infty} \frac{a_i \phi_i^\tau P_i \phi_i}{\log |P_{i+1}^{-1}| (\log \log |P_{i+1}^{-1}|)^c} < \infty \quad \text{for any } c > 1 \tag{27}$$

Also, we note that

$$|P_{n+1}^{-1}| \geqslant \lambda_{\max}^{n+1} \left(\frac{1}{d}\right)^{d-1} \geqslant r_{n+1} \left(\frac{1}{d}\right)^{d}$$

Hence

$$\log r_{n+1} - d \log d \leqslant \log |P_{n+1}^{-1}| \leqslant d \log r_{n+1} \tag{28}$$

Case 1

If $\lim_{n\to\infty} r_n < \infty$ then it follows from (25) and (28) that

$$\sum_{i=0}^{\infty} a_i \phi_i^\tau P_i \phi_i < \infty$$

Then by the martingale convergence theorem we have

$$\sum_{i=0}^{\infty} a_i \phi_i^\tau P_i \phi_i (\|w_{i+1}\|^2 - E[\|w_{i+1}\|^2 \mid \mathscr{F}_i]) < \infty$$

Hence

$$\sum_{i=0}^{\infty} a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 = 0(1)$$

Then from Lemma 2 we know that

$$\operatorname{tr} \tilde{\theta}_{n+1}^\tau P_{n+1}^{-1} \tilde{\theta}_{n+1} = 0(1)$$

which verifies (13).

Case 2

We now consider the case where $\lim_{n\to\infty} r_n = \infty$. If $\beta > 2$ then it follows from (26) that

$$\sum_{i=0}^{\infty} E\left[\left|\frac{a_i \phi_i^\tau P_i \phi_i}{\log |P_{i+1}^{-1}|} (\|w_{i+1}\|^2 - E[\|w_{i+1}\|^2 \mid \mathscr{F}_i])\right|^{\beta'/2} \middle| \mathscr{F}_i\right] < \infty$$

whenever $\beta' \in (2, \min(\beta, 4)]$. Then by the martingale convergence theorem

$$\sum_{i=0}^{\infty} \frac{a_i \phi_i^\tau P_i \phi_i}{\log |P_{i+1}^{-1}|} (\|w_{i+1}\|^2 - E[\|w_{i+1}\|^2 | \mathscr{F}_i]) < \infty$$

From this, by the Kronecker lemma, we have

$$\sum_{i=0}^{n} a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 = \sum_{i=0}^{n} a_i \phi_i^\tau P_i \phi_i E[\|w_{i+1}\|^2 | \mathscr{F}_i] + o(\log |P_{n+1}^{-1}|)$$

Hence, by (25) and (28), we conclude

$$\sum_{i=0}^{n} a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 = O(\log r_{n+1}) \tag{29}$$

which together with Lemma 2 implies (13) for the case where $\beta > 2$.

If $\beta = 2$ then by (27) and the martingale convergence theorem we see that

$$\sum_{i=0}^{\infty} \frac{a_i \phi_i^\tau P_i \phi_i}{(\log |P_{i+1}^{-1}|)(\log \log |P_{i+1}^{-1}|)^c} (\|w_{i+1}\|^2 - E[\|w_{i+1}\|^2 | \mathscr{F}_i]) < \infty$$

Noting that $\sup_n E[\|w_{n+1}\|^2 | \mathscr{F}_n] < \infty$ and (25), then again by the Kronecker lemma we have

$$\sum_{i=0}^{n} a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2 = o(\log |P_{n+1}^{-1}| (\log \log |P_{n+1}^{-1}|)^c) \tag{30}$$

which, together with Lemma 2 and (28), yields (13) for the case $\beta = 2$.

This completes the proof of both Lemma 3 and Theorem 1. □

Denote by ϕ_n^ξ the difference $\phi_n - \phi_n^0$; then by (14)

$$\phi_n^\xi = [0 \quad \dots \quad 0 \quad \xi_n^\tau \quad \dots \quad \xi_{n-r+1}^\tau]^\tau$$

Theorem 2

Under the conditions of Theorem 1 as $n \to \infty$:

(i) if $\beta > 2$ and $\log r_n^0 = o(\lambda_{\min}^{0n})$ then

$$\|\theta_n - \theta\| = O\left(\left(\frac{\log r_n^0}{\lambda_{\min}^{on}}\right)^{1/2}\right) \quad \text{a.s.}$$

(ii) If $\beta = 2$ and $\log r_n^0 (\log \log r_n^0)^c = o(\lambda_{\min}^{0n})$ for some $c > 1$ then

$$\|\theta_n - \theta\| = O\left(\left(\frac{\log r_n^0 (\log \log r_n^0)^c}{\lambda_{\min}^{on}}\right)^{1/2}\right) \quad \text{a.s.}$$

Proof

Since $C^{-1}(z)$ is strictly positive-real it must be stable; then by (16) and (23), and noting that $\operatorname{tr} \tilde{\theta}_{n+1}^\tau P_{n+1}^{-1} \tilde{\theta}_{n+1} \geq 0$, we have

$$\sum_{i=0}^{n} \|\xi_{i+1}\|^2 = O\left(\sum_{i=0}^{n} \|\tilde{\theta}_{i+1}^\tau \phi_i\|^2\right) = O(1) + O\left(\sum_{i=0}^{n} a_i \phi_i^\tau P_i \phi_i \|w_{i+1}\|^2\right) \tag{31}$$

(i) Let $\beta > 2$; by (29) we see that

$$\sum_{i=0}^{n-1} \|\phi_i^\xi\|^2 = O(\log r_n)$$

Hence

$$r_n \leqslant 2r_n^0 + 2\sum_{i=0}^{n-1} \|\phi_i^\xi\|^2 = 2r_n^0 + O(\log r_n)$$

and

$$r_n = O(r_n^0) \tag{32}$$

Further, for any $x \in R^d$ with $\|x\| = 1$ it is clear that

$$\begin{aligned}\sum_{i=0}^{n-1} (x^\tau \phi_i^0)^2 &= \sum_{i=0}^{n-1} (x^\tau \phi_i - x^\tau \phi_i^\xi)^2 \leqslant 2\sum_{i=0}^{n-1} (x^\tau \phi_i)^2 + 2\sum_{i=0}^{n-1} \|\phi_i^\xi\|^2 \\ &\leqslant 2\sum_{i=0}^{n-1} (x^\tau \phi_i)^2 + O(\log r_n^0)\end{aligned} \tag{33}$$

Hence

$$\lambda_{\min}^{0n} \leqslant 2\lambda_{\min}^n + o(\lambda_{\min}^{0n})$$

and then

$$\lambda_{\min}^{0n} = O(\lambda_{\min}^n) \tag{34}$$

By (32) and (34) we have

$$\frac{\log r_n}{\lambda_{\min}^n} = O\left(\frac{\log r_n^0}{\lambda_{\min}^{0n}}\right)$$

The first assertion then follows from Theorem 1.

(ii) If $\beta = 2$ then from (28), (30) and (31) we have

$$\sum_{i=0}^{n-1} \|\phi_i^\xi\|^2 = O(\log r_n (\log\log r_n)^c)$$

Hence (32) remains valid, and (33) becomes

$$\sum_{i=0}^{n-1} (x^\tau \phi_i^0)^2 \leqslant 2\sum_{i=0}^{n-1} (x^\tau \phi_i)^2 + O(\log r_n^0 (\log\log r_n^0)^c)$$

Therefore, under the conditions of the theorem, (34) holds true. Then

$$\frac{\log r_n (\log\log r_n)^c}{\lambda_{\min}^n} = O\left(\frac{\log r_n^0 (\log\log r_n^0)^c}{\lambda_{\min}^{0n}}\right)$$

and the second conclusion follows from Theorem 1.

3. Convergence rate of parameter estimation for systems with attenuately excited control

In a stochastic adaptive-control system a performance index of the long-run average type is frequently used (see e.g. Goodwin *et al.* 1981, Chen 1984, Chen and Caines 1985), for which an external decaying disturbance added to the input or to the output of the system does not change the performance index. This is one of the reasons why we introduced a random dither with covariance matrix tending to zero (Chen and Guo 1985 b, c) to the system in order to get both optimality of the control and consistency of the estimate. To be precise, this treatment, called the attenuating excitation technique, consists of the following. Let $\{v_n\}$ be an l-dimensional mutually independent random vector sequence and let $\{v_n\}$ be independent of $\{w_n\}$ with properties

$$Ev_n = 0, \quad Ev_n v_n^\tau = \frac{1}{n^\varepsilon} I, \quad \|v_n\|^2 \leqslant \frac{\sigma^2}{n^\varepsilon}, \quad \varepsilon \in \left[0, \frac{1}{2(t+1)}\right), \quad t = \max(p, q, r) + mp - 1 \tag{35}$$

where σ^2 is a constant. $\{v_n\}$ will serve as the attenuating excitation source.

Without loss of generality, we assume $\mathscr{F}_n = \sigma\{w_i, v_i, 0 \leqslant i \leqslant n\}$. Set

$$\mathscr{F}'_{n-1} = \sigma\{w_i, 0 \leqslant i \leqslant n, v_j, 0 \leqslant j \leqslant n-1\}$$

Let $\mathscr{F}'_{n-1}$-measurable u_n^s be the desired control. The attenuating excitation technique suggests that we take

$$u_n = u_n^s + v_n$$

instead of $u_n = u_n^s$.

Theorem 3

Suppose that

(*a*) $(w_n, \mathscr{F}_n)$ is a martingale difference sequence with

$$\sup_n E(\|w_{n+1}\|^\beta \mid \mathscr{F}_n) < \infty, \quad \beta \geqslant 2 \tag{36}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} w_i w_i^\tau = R > 0 \tag{37}$$

(*b*) $C^{-1}(z) - \frac{1}{2}I$ is strictly positive-real;

(*c*) $A(z)$, $B(z)$ and $C(z)$ have no common left factor and A_p is of full rank with $A_0 = I$ by definition;

(*d*) The output of system (1) under control

$$u_n = u_n^s + v_n$$

has growth rate

$$\frac{1}{n} \sum_{i=1}^{n} \|y_i\|^2 = 0(n^\delta) \tag{38}$$

where u_n^s in $\mathscr{F}'_{n-1}$-measurable and

$$\frac{1}{n} \sum_{i=1}^{n} \|u_i^s\|^2 = 0(n^\delta) \tag{39}$$

and where v_n is defined by (35) and

$$\delta \in \left[0, \frac{1 - 2\varepsilon(t+1)}{2t+3}\right) \tag{40}$$

Then

$$\|\theta_n - \theta\| = \begin{cases} 0\left(\left(\dfrac{\log n}{n^\alpha}\right)^{1/2}\right) & \text{a.s. if } \beta > 2 \\ 0\left(\left(\dfrac{\log n(\log\log n)^c}{n^\alpha}\right)^{1/2}\right) \quad \forall c > 1 & \text{a.s. if } \beta = 2 \end{cases}$$

for any $\alpha \in (\frac{1}{2}(1+\delta), 1-(t+1)(\varepsilon+\delta)]$.

Proof

We first note that

$$(t+1)(\varepsilon+\delta) + \tfrac{1}{2}\delta < (t+1)\left[\varepsilon + \frac{1-2\varepsilon(t+1)}{2t+3}\right] + \frac{1-2\varepsilon(t+1)}{2(2t+3)} = \tfrac{1}{2}$$

so the interval $(\frac{1}{2}(1+\delta), 1-(t+1)(\varepsilon+\delta)$ is not empty. We also note that

$$1 < \frac{1}{1-\varepsilon} < \frac{2(t+1)}{2t+1} < 2$$

so we can take $\gamma \in (1/(1-\varepsilon), 2]$ such that

$$\sum_{i=1}^{\infty} E\left[\frac{\|v_i v_i^\tau - \frac{1}{i^\varepsilon} I\|^\gamma}{i^{(1-\varepsilon)\gamma}} \middle| \mathscr{F}'_{i-1}\right] < \infty$$

Hence we have

$$\sum_{i=1}^{\infty} \frac{v_i v_i^\tau - \frac{1}{i^\varepsilon} I}{i^{1-\varepsilon}} < \infty$$

From this, by the Kronecker lemma and by the fact that

$$\frac{1}{n^{1-\varepsilon}} \sum_{i=1}^{n} \frac{1}{i^\varepsilon} \underset{n\to\infty}{\to} \frac{1}{1-\varepsilon}$$

we conclude that

$$\frac{1-\varepsilon}{n^{1-\varepsilon}} \sum_{i=1}^{n} v_i v_i^\tau \underset{n\to\infty}{\to} I \tag{41}$$

By (21) and (39) it follows that

$$\sum_{i=2}^{n} u_i^{\varepsilon\tau} v_i = 0\left(\left(\sum_{i=2}^{n} \|u_i^\varepsilon\|^2\right)^{1/2}\left(\log\left(\sum_{i=2}^{n} \|u_i^\varepsilon\|^2 + e\right)\right)^{1/2+\eta}\right)$$
$$= 0(n^{(1+\delta)/2}(\log n)^{1/2+\eta}), \quad \eta > 0 \tag{42}$$

Then from (37)–(39), (41) and (42) we know that

$$r_n^0 = 0(n^{1+\delta}) \tag{43}$$

If we can show that for sufficiently large n_0

$$\lambda_{\min}^{0n} \geq c_0 n^\alpha, \quad c_0 > 0, \quad \forall n \geq n_0 \tag{44}$$

then (43) and (44) guarantee the conditions in Theorem 2, and the conclusion of the theorem follows from Theorem 2.

Clearly, (44) is equivalent to

$$\liminf_{n\to\infty} n^{-\alpha}\lambda_{\min}^{0n} \neq 0 \tag{45}$$

We now prove (45) along the lines of the argument used in Chen and Guo (1985 c).

Note that the full rank of A_p implies $\deg A(z)=p$, $\deg[\det A(z)]=mp$ and $\deg[\operatorname{Adj} A(z)]=mp-p$, since $A(z)[\operatorname{Adj} A(z)]=\det A(z)I$. Set

$$\psi_n = (\det A(z))\phi_n^0$$

$$\det A(z) = a_0 + a_1 z + \dots + a_{mp}z^{mp}$$

By the Schwarz inequality and the fact that $\phi_i^0=0$ for $i<0$, it is easy to see that

$$\lambda_{\min}\left(\sum_{i=1}^{n}\psi_i\psi_i^{\tau}\right)=\inf_{\|x\|=1}\sum_{i=1}^{n}(x^{\tau}\psi_i)^2 \leqslant (mp+1)\sum_{j=0}^{mp}a_j^2\lambda_{\min}^{0n}$$

where $\lambda_{\min}(X)$ denotes the minimum eigenvalue of a matrix X. So for (45) it suffices to show that

$$\liminf_{n\to\infty} n^{-\alpha}\lambda_{\min}\left(\sum_{i=1}^{n}\psi_i\psi_i^{\tau}\right)\neq 0 \tag{46}$$

If (46) were not true then there would exist a vector sequence $\{\eta_{n_k}\}$,

$$\eta_{n_k}=(\alpha_{n_k}^{0\tau}\ \ \dots\ \ \alpha_{n_k}^{(p-1)\tau}\ \ \beta_{n_k}^{0\tau}\ \ \dots\ \ \beta_{n_k}^{(q-1)\tau}\ \ \gamma_{n_k}^{0\tau}\ \ \dots\ \ \gamma_{n_k}^{(r-1)\tau})^{\tau}\in R^d$$

such that $\|\eta_{n_k}\|=1$ and

$$\lim_{n\to\infty} n_k^{-\alpha}\left(\sum_{i=1}^{n_k}(\eta_{n_k}^{\tau}\psi_i)^2\right)\underset{k\to\infty}{\to} 0 \tag{47}$$

Let

$$\begin{aligned} H_{n_k}(z) &= \sum_{i=0}^{p-1}\alpha_{n_k}^{i\tau}z^i(\operatorname{Adj} A(z))[B(z)\quad C(z)] \\ &\quad + \sum_{i=0}^{q-1}\beta_{n_k}^{i\tau}z^i[\det A(z)I_l\quad 0] + \sum_{i=0}^{r-1}\gamma_{n_k}^{i\tau}z^i[0\quad \det A(z)I_m] \\ &\triangleq \sum_{j=0}^{t}[h_{n_k}^{j\tau}\quad g_{n_k}^{j\tau}]z^j;\quad t=\max(p,q,r)+mp-1 \end{aligned} \tag{48}$$

where $h_{n_k}^i$ and $g_{n_k}^j$ are l- and m-dimensional vectors respectively. Clearly, $h_{n_k}^i$ and $g_{n_k}^j$ are bounded in k and ω (sampling points).

Then (47) means

$$\lim_{k\to\infty} n_k^{-\alpha}\sum_{i=1}^{n_k}(h_{n_k}^{0\tau}u_i+\dots+h_{n_k}^{t\tau}u_{i-t}+g_{n_k}^{0\tau}w_i+\dots+g_{n_k}^{t\tau}w_{i-t})^2=0 \tag{49}$$

By (21) and (39), it is easy to see that

$$\left\|\sum_{i=1}^{n}u_{i-j}^{s}v_i^{\tau}\right\| = O(n^{(1+\delta)/2}(\log n)^{1/2+\eta}),\quad j\geqslant 0 \tag{50}$$

$$\left\|\sum_{i=1}^{n}u_{i-j}v_i^{\tau}\right\| = O(n^{(1+\delta)/2}(\log n)^{1/2+\eta}),\quad j>0 \tag{51}$$

$$\left\| \sum_{i=1}^{n} w_{i-j} v_i^{\tau} \right\| = 0(n^{1/2} (\log n)^{1/2+\eta}), \quad j \geq 0 \tag{52}$$

Because $g_{n_k}^j$ and $h_{n_k}^i$ are bounded and $\alpha > \frac{1}{2}(1+\delta)$, from (50)–(52) we know that

$$n_k^{-\alpha} \left\{ h_{n_k}^{0\tau} \sum_{i=1}^{n_k} u_i^s v_i^{\tau} h_{n_k}^0 + \sum_{j=1}^{t} h_{n_k}^{j\tau} \sum_{i=1}^{n_k} u_{i-j} v_i^{\tau} h_{n_k}^0 + \sum_{j=0}^{t} g_{n_k}^{j\tau} \sum_{i=1}^{n_k} w_{i-j} v_i^{\tau} h_{n_k}^0 \right\} \underset{k\to\infty}{\to} 0 \tag{53}$$

From this and (49) we have

$$n_k^{-\alpha} \sum_{i=1}^{n_k} (h_{n_k}^{0\tau} v_i)^2 \underset{k\to\infty}{\to} 0 \tag{54}$$

$$n_k^{-\alpha} \sum_{i=1}^{n_k} (h_{n_k}^{0\tau} u_i^s + h_{n_k}^{1\tau} u_{i-1} + \dots + h_{n_k}^{t\tau} u_{i-t} + g_{n_k}^{0\tau} w_i + \dots + g_{n_k}^{t\tau} w_{i-t})^2 \underset{k\to\infty}{\to} 0 \tag{55}$$

Relationships (39), (41), (54) and (55) imply that

$$\|h_{n_k}^0\|^2 = o(n_k^{-(1-\varepsilon-\alpha)}) \tag{56}$$

$$n_k^{-(1+\delta)+1-\varepsilon-\alpha} \sum_{i=1}^{n_k} (h_{n_k}^{0\tau} u_i^s)^2 \underset{k\to\infty}{\to} 0 \tag{57}$$

$$n_k^{-\alpha-(\varepsilon+\delta)} \sum_{i=1}^{n_k} (h_{n_k}^{1\tau} u_{i-1} + \dots + h_{n_k}^{t\tau} u_{i-t} + g_{n_k}^{0\tau} w_i + \dots + g_{n_k}^{t\tau} w_{i-t})^2 \underset{k\to\infty}{\to} 0 \tag{58}$$

Comparing (58) with (49), we note that $n_k^{-\alpha}$ in (49) is replaced by $n_k^{-\alpha-(\varepsilon+\delta)}$ and the term $h_{n_k}^{0\tau} u_i$ disappears. Continuing the same argument, we obtain

$$\|h_{n_k}^i\|^2 = o(n_k^{-(1-\varepsilon-\alpha-i(\varepsilon+\delta))}), \quad 0 \leq i \leq t \tag{59}$$

$$n_k^{-\alpha-s(\varepsilon+\delta)} \sum_{i=1}^{n_k} (h_{n_k}^{s\tau} u_{i-s} + \dots + h_{n_k}^{t\tau} u_{i-t} + g_{n_k}^{0\tau} w_i + \dots + g_{n_k}^{t\tau} w_{i-t})^2 \underset{k\to\infty}{\to} 0$$

$$\forall \text{ integers } s \in [1, t+1] \tag{60}$$

Since $\alpha \leq 1-(t+1)(\varepsilon+\delta)$, it immediately follows from (59) that

$$\|h_{n_k}^i\| \underset{k\to\infty}{\to} 0, \quad 0 \leq i \leq t \tag{61}$$

For $s = t+1$ we have, from (60),

$$n_k^{-\alpha-(t+1)(\varepsilon+\delta)} \sum_{i=1}^{n_k} (g_{n_k}^{0\tau} w_i + \dots + g_{n_k}^{t\tau} w_{i-t})^2 \underset{k\to\infty}{\to} 0 \tag{62}$$

By (37), from (62) it is easy to conclude that

$$g_{n_k}^j \underset{k\to\infty}{\to} 0 \quad \forall \text{ integers } j \in [0, t] \tag{63}$$

and by (48), (61) and (63) we see

$$H_{n_k}(z) \underset{k\to\infty}{\to} 0$$

From here, by use of Condition (*c*), exactly the same argument as used in Chen and Guo (1985 c) leads to a contradiction. Thus (46) is verified and the proof is completed. □

4. Application to adaptive control

We now apply the results obtained to the following adaptive-control problem with quadratic loss function:

$$J(u)=\overline{\lim_{n\to\infty}}\, J_n(u), \quad J_n(u)=\frac{1}{n}\sum_{i=0}^{n-1}(y_i^\tau Q_1 y_i+u_i^\tau Q_2 u_i) \tag{64}$$

where $Q_1 \geqslant 0$, $Q_2>0$.

We present System (1) in the state-space form

$$x_{k+1}=Ax_k+Bu_k+Cw_{k+1} \tag{65}$$

$$y_k=Hx_k, \quad x_0^\tau=[y_0^\tau \quad 0 \quad \dots \quad 0] \tag{66}$$

where

$$A=\begin{bmatrix} -A_1 & I & & \\ -A_2 & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & I \\ -A_s & 0 & \dots & 0 \end{bmatrix} \quad B=\begin{bmatrix} B_1 \\ \vdots \\ B_s \end{bmatrix} \tag{67}$$

$$C^\tau=[I \quad C_1^\tau \quad \dots \quad C_{s-1}^\tau], \quad H=[\underbrace{I \quad 0 \quad \dots \quad 0}_{ms}]\}m \tag{68}$$

with $s=p\vee q\vee(r+1)$ and $A_i=0$, $B_j=0$, $C_k=0$ for $i>p, j>q, k>r$.

Denote by A_i^n, B_j^n and C_k^n the estimates given by θ_n for A_i, B_j and C_k respectively, $i=1,\dots,p, j=1,\dots,q$ and $k=1,\dots,r$, and estimate the state x_n by the adaptive filter

$$\left.\begin{aligned} \hat{x}_{n+1}&=\hat{A}_n\hat{x}_n+\hat{B}_n u_n+\hat{C}_n(y_{n+1}-H\hat{A}_n\hat{x}_n-H\hat{B}_n u_n) \\ \hat{x}_0&=[y_0^\tau \quad 0 \quad \dots \quad 0]^\tau \end{aligned}\right\} \tag{69}$$

where $\hat{A}_n$, $\hat{B}_n$ and $\hat{C}_n$ are defined by (67) and (68) with A_i, B_j and C_k replaced by their estimates A_i^n, B_j^n and C_k^n respectively, $i=1,\dots,p, j=1,\dots,q$ and $k=1,\dots,r$.

Set

$$L_n=-(\hat{B}_n^\tau S_n\hat{B}_n+Q_2)^{-1}\hat{B}_n^\tau S_n\hat{A}_n \tag{70}$$

where S_n is defined recursively by

$$S_n=\hat{A}_n^\tau S_{n-1}\hat{A}_n-\hat{A}_n^\tau S_{n-1}\hat{B}_n(Q_2+\hat{B}_n^\tau S_{n-1}\hat{B}_n)^{-1}\hat{B}_n^\tau S_{n-1}\hat{A}_n+H^\tau Q_1 H \tag{71}$$

with an arbitrary initial value $S_0 \geqslant 0$.

Define stopping times $\{\tau_k\}$, $\{\sigma_k\}$ with

$$1=\tau_1<\sigma_1<\tau_2<\sigma_2<\dots$$

such that

$$\sigma_k=\sup\left\{t>\tau_k: \sum_{i=\tau_k}^{j-1}\|L_i\hat{x}_i\|^2 \leqslant (j-1)^{1+\delta}+\|L_{\tau_k}\hat{x}_{\tau_k}\|^2 \quad \forall j\in(\tau_k,t]\right\} \tag{72}$$

$$\tau_{k+1}=\inf\left\{t>\sigma_k: \sum_{i=\tau_k}^{\sigma_k-1}\|L_i\hat{x}_i\|^2 \leqslant \frac{t^{1+\delta}}{2^k}, \quad \sum_{j=1}^{t}\|\hat{x}_j\|^2 \leqslant t^{1+\delta/2}, \frac{\|L_t\hat{x}_t\|^2}{t^{1+\delta}} \leqslant 1\right\} \tag{73}$$

where $\delta\in\left(0, \dfrac{1-2\varepsilon(t+1)}{2t+3}\right)$

We now define the adaptive control u_n^a as

$$u_n^a = L_n^0 \hat{x}_n + v_n \tag{74}$$

with

$$L_n^0 = \begin{cases} L_n & \text{if } n \text{ belongs to some } [\tau_k, \sigma_k) \\ 0 & \text{if } n \text{ belongs to some } [\sigma_k, \tau_{k+1}) \end{cases} \tag{75}$$

where v_n is defined by (35).

It is worth noting that u_n^a can really be recursively computed in real time.

It is shown in Chen and Guo (1985 c) that

$$\inf_{u \in U} J(u) = \operatorname{tr} SCRC^\tau \tag{76}$$

where S is a positive-definite matrix satisfying

$$S = A^\tau SA - A^\tau SB(Q_2 + B^\tau SB)^{-1} B^\tau SA + H^\tau Q_1 H \tag{77}$$

and by definition

$$U = \left\{ u : \sum_{i=1}^{n} \|u_i\|^2 = 0(n),\ \|u_n\|^2 = o(n) \quad \text{a.s. as } n \to \infty \right\} \tag{78}$$

Theorem 4

Suppose that for System (1) the following conditions are fulfilled:

(*a*) $\{w_n, \mathscr{F}_n\}$ is a martingale difference sequence with properties

$$\sup_n E(\|w_n\|^\beta \mid \mathscr{F}_{n-1}) < \infty \quad \text{a.s.} \quad \text{if } \beta \geq 2$$

$$\left\| \frac{1}{n} \sum_{i=1}^{n} w_i w_i^\tau - R \right\| = 0(n^{-\rho}) \quad \text{a.s.} \quad \text{as } n \to \infty$$

where R is a positive-definite matrix and $\rho > 0$;

(*b*) A_p is of full rank ($A_0 = I$ by definition) and $A(z)$ is stable;

(*c*) $C^{-1}(z) - \frac{1}{2}I$ is strictly positive real;

(*d*) (A, B, D) is controllable and observable, where D is any matrix such that $D^\tau D = H^\tau Q_1 H$.

Then under the adaptive control $\{u_n^a\}$ given by (74), the following convergence rates hold as $n \to \infty$:

$$\|\theta_n - \theta\| = \begin{cases} 0\left(\left(\dfrac{\log n}{n^\alpha}\right)^{1/2}\right) & \text{a.s.} & \text{if } \beta > 2 \quad (79) \\ 0\left(\left(\dfrac{\log n (\log \log n)^c}{n^\alpha}\right)^{1/2}\right) & \text{a.s.}, \ \forall c > 1, & \text{if } \beta = 2 \quad (80) \end{cases}$$

for any $\alpha \in (\frac{1}{2}(1 + \delta),\ 1 - (t + 1)(\varepsilon + \delta)]$, and

$$\|J_n(u^a) - \operatorname{tr} SCRC^\tau\| = 0(n^{-(\rho \wedge \varepsilon)}) \quad \text{a.s.} \tag{81}$$

where $\rho \wedge \varepsilon = \min(\rho, \varepsilon)$.

Proof

By an argument similar to that used in the proof of Theorem 1 of Chen and Guo (1985 c), it can be shown that the following properties hold true:

(i) $$\frac{1}{n}\sum_{i=1}^{n}\|L_i^0\hat{x}_i\|^2 = 0(n^\delta) \quad \text{a.s.} \tag{82}$$

(see Lemma 4 of Chen and Guo 1985 c);

(ii) $$S_n \underset{n\to\infty}{\to} S \quad \text{a.s.} \tag{83}$$

(see Lemma 5 of Chen and Guo (1985 c);

(iii) there is a τ_{k_0} such that

$$L_n^0 \equiv L_n \quad \text{a.s.} \quad \forall\, n \geqslant \tau_{k_0} \tag{84}$$

(see Lemma 6 of Chen and Guo 1985 c);

(iv) $$\|\hat{x}_{k+1}\|^2 + \|\hat{x}_{k+1} - x_{k+1}\|^2 = 0(1) + 0\left(\sum_{i=1}^{k}\mu^{k-i}(\|w_{i+1}\|^2 + \|v_i\|^2)\right) \tag{85}$$

with $0 < \mu < 1$ (see (57) of Chen and Guo 1985 c);

(v) $$\frac{1}{n}\sum_{i=1}^{n}\|x_i\|^2 = 0(1) \quad \text{a.s.} \tag{86}$$

(see (62) of Chen and Guo 1985 c);

and finally from both (49) and (59) of Chen and Guo (1985 c) we have

(vi) $$\frac{1}{n}\sum_{i=1}^{n}\|\hat{x}_i - x_i\|^2 = 0(\tfrac{1}{n}) + 0\left(\frac{1}{n}\sum_{i=1}^{n}\|\tilde{\theta}_i\|^2(\|\hat{x}_i\|^2 + \|w_{i+1}\|^2 + \|v_i\|^2)\right) \tag{87}$$

Since $A(z)$ is stable, from (82), (74) and (1) we see that

$$\frac{1}{n}\sum_{i=1}^{n}\|y_i\|^2 = 0(n^\delta)$$

From Theorem 6.2-6 of Kailath (1980, p. 366) it can be concluded that the controllability of (A, B) implies that $A(z)$ and $B(z)$ have no common left factor. Hence Theorem 3 is applicable and (79) and (80) are verified. Thus we only need to establish (81).

By a standard treatment (see e.g. Chen 1985), from (65), (66) and (77) we have

$$\begin{aligned} J_n(u) = \sum_{i=0}^{n-1}(y_i^\tau Q_1 y_i + u_i^\tau Q_2 u_i) &= x_0^\tau S x_0 - x_n^\tau S x_n + \sum_{i=0}^{n-1} w_{i+1}^\tau C^\tau S C w_{i+1} \\ &+ 2\sum_{i=0}^{n-1}(Ax_i + Bu_i)^\tau S C w_{i+1} + \sum_{i=0}^{n-1}(u_i - Lx_i)^\tau (Q_2 + B^\tau S B)(u_i - Lx_i) \end{aligned} \tag{88}$$

where

$$L = -(B^\tau S B + Q_2)^{-1} B^\tau S A$$

From (35) and (85), we have

$$\begin{aligned} \|x_{k+1}\|^2 &= 0(1) + 0\left(\sum_{i=1}^{k}\mu^{k-i} i\left(\frac{\Sigma_{j=1}^{i+1}\|w_j\|^2}{i} - \frac{\Sigma_{j=1}^{i}\|w_j\|^2}{i}\right)\right) \\ &= 0(1) + 0\left(\sum_{i=0}^{k}\mu^{k-i} i \frac{1}{i^\rho}\right) = 0(1) + 0(k^{1-\rho}) \end{aligned}$$

hence

$$\frac{\|x_n\|^2}{n} = 0(n^{-\rho}) \quad \text{a.s.} \tag{89}$$

and

$$\frac{1}{n}\sum_{i=0}^{n-1} w_{i+1}^\tau C^\tau SCw_{i+1} - \operatorname{tr} SCRC^\tau$$

$$= \operatorname{tr} C^\tau SC\left(\frac{1}{n}\sum_{i=0}^{n-1} w_{i+1}w_{i+1}^\tau - R\right) = 0(n^{-\rho}) \tag{90}$$

Since $\theta_n \underset{n\to\infty}{\to} \theta$ and S_n is bounded, it is not difficult to see that

$$S_{n+1} - S = (A + BK_n)^\tau (S_n - S)(A + BL) + \varepsilon_n$$

Here

$$K_n = -(Q_2 + B^\tau S_n B)^{-1} B^\tau S_n A$$

$$\|\varepsilon_n\| = 0(\|\tilde{\theta}_n\|) = o\left(\frac{1}{n^{1/4}}\right)$$

where the last estimate follows from (79), (80) and the fact that $\alpha > \frac{1}{2}$.

Noting that $A + BL$ is stable and $K_n \underset{n\to\infty}{\to} L$ by (83), we have

$$\|S_n - S\| = o\left(\frac{1}{n^{1/4}}\right) \tag{91}$$

$$\|L_i^0 - L\| = o\left(\frac{1}{n^{1/4}}\right) \tag{92}$$

by (84).

We now estimate the last term in (88):

$$\left\|\sum_{i=0}^{n-1} (u_i - Lx_i)^\tau (Q_2 + B^\tau SB)(u_i - Lx_i)\right\| = 0\left(\sum_{i=0}^{n-1} \|u_i - Lx_i\|^2\right)$$

$$= 0\left(\sum_{i=0}^{n-1} \|(L_i^0 - L)x_i + L_i^0(\hat{x}_i - x_i) + v_i\|^2\right) \tag{93}$$

in which

$$\sum_{i=1}^{n-1} \|v_i\|^2 = 0(n^{1-\varepsilon}) \tag{94}$$

by (41), and

$$\sum_{i=1}^{n} \|(L_i^0 - L)x_i\|^2 = 0\left(\sum_{i=1}^{n} \frac{1}{i^{1/2}} \|x_i\|^2\right) = 0\left(\sum_{i=1}^{n} \frac{1}{i^{1/2}}\left(\sum_{j=0}^{i} \|x_j\|^2 \sum_{j=0}^{i-1} \|x_j\|^2\right)\right)$$

$$= 0\left(\frac{\Sigma_{j=0}^{n} \|x_j\|^2}{n^{1/2}} + \sum_{i=1}^{n-1}\left(\frac{\Sigma_{j=0}^{i} \|x_j\|^2}{i^{1/2}} - \frac{\Sigma_{j=0}^{i} \|x_j\|^2}{(i+1)^{1/2}}\right) - \|x_0\|^2\right)$$

$$= 0\left(n^{1/2} + \sum_{i=1}^{n-1}\left(\frac{1}{i^{1/2}} - \frac{1}{(i+1)^{1/2}}\right)i - \|x_0\|^2\right)$$

$$= 0\left(n^{1/2} + \sum_{i=1}^{n-1} \frac{i}{i^{1+1/2}}\right) + 0(1) = 0(n^{1/2}) \tag{95}$$

Similarly, we have

$$\sum_{i=0}^{n-1} \|L_i^0(\hat{x}_i - x_0)\|^2 = 0\left(\sum_{i=0}^{n-1} \|\hat{x}_i - x_i\|^2\right) = o(n^{1/2}) \tag{96}$$

by (87) and $\sum_{i=1}^{n} \|\tilde{\theta}_i\|^2 \|\hat{x}_i\|^2 = o\left(\sum_{i=1}^{n} i^{-1/2} \|\hat{x}_i\|^2\right) = o(n^{1/2})$.

Combining (88)–(96), we conclude that

$$J_n(u^a) - \text{tr } SCRC^\tau = 0(n^{-\rho} + n^{-\varepsilon} + n^{-1/2})$$
$$= 0(n^{-\rho} + n^{-\varepsilon}) = 0(n^{-(\rho \wedge \varepsilon)}) \quad \text{a.s.}$$

5. Conclusion

We have presented the convergence rates of both the parameter estimate and the quadratic index when the least-squares (LS) algorithm is applied. Comparing with the stochastic-gradient (SG) algorithm (Chen and Guo 1985 b, c), we have found that the LS algorithm is not as simple as the SG one, but gives better results in the following sense.

(i) We have not given the convergence rate for SG algorithm; but we have done so for the LS algorithm.

(ii) For strong consistency of the SG algorithm it is supposed that

$$\frac{\lambda_{\max}^n}{\lambda_{\min}^n} = 0((\log \lambda_{\max}^n)^{1/4})$$

which means that $\log \lambda_{\max}^n$ and $\log \lambda_{\min}^n$ are of the same order, while for the LS algorithm we only require that

$$\frac{\log \lambda_{\max}^n}{\lambda_{\min}^n} \underset{n \to \infty}{\to} 0 \quad \text{for } \beta > 2$$

(iii) When adaptive control with the attenuating excitation technique is applied for strong consistency of parameter estimates, the growth rate of $(1/n)\Sigma_{i=1}^n (\|y_i\|^2 + \|u_i\|^2)$ should be limited by $0((\log n)^\delta)$ for the SG algorithm, but for the LS algorithm the order is increased to n^δ. The covariance matrix of the excitation source from $I/\log^\varepsilon n$ for the SG algorithm is reduced to I/n^ε for the LS algorithm. This means that the LS algorithm can give a better approach to the optimal value of the quadratic cost.

It is clear that the LS algorithm can be used to solve the stochastic adaptive tracking and pole–zero assignment problems.

References

CHEN, H. F., 1982, *Scientia sin.*, **25**, 771; 1984, *SIAM J. Control Optim.*, **22**, 758; 1985, *Recursive Estimation and Control for Stochastic Systems* (New York: Wiley).

CHEN, H. F., and CAINES, P. E., 1985, *I.E.E.E. Trans. autom. Control*, **30**, 189.

CHEN, H. F., and GUO, L., 1985 a, The limit of stochastic gradient algorithm for identifying systems excited not persistently. *Kexue Tongbao (Science Bulletin)*, to be published; 1985 b, Adaptive control with recursive identification for stochastic linear systems. *Advances in Control and Dynamic Systems*, Vol. 24, edited by C. T. Leondes (New York: Academic Press); 1985 c, Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost. *SIAM J. Control Optim.*, to be published.

GOODWIN, G. C., RAMADGE, P. J., and CAINES, P. E., 1981, *SIAM J. Control Optim.*, **19**, 829.

KAILATH, T., 1980, *Linear Systems* (Englewood Cliffs, NJ: Prentice-Hall).

LAI, T. L., and WEI, C. Z., 1982, *Ann. Statist.*, **10**, 154.

LJUNG, L., 1976, *I.E.E.E. Trans. autom. Control*, **22**, 551.

MOORE, J. B., 1978, *Automatica*, **14**, 505.

SOLO, V., 1979, *I.E.E.E. Trans. autom. Control*, **24**, 958.