系统科学与数学
J. Sys. Sci. & Math. Scis.
5 (2) (1985), 81—93

STRONG CONSISTENCY OF PARAMETER ESTIMATES FOR DISCRETE-TIME STOCHASTIC SYSTEMS

CHEN HAN-FU GUO LEI
(Institute of Systems Science, Academia Sinica, Beijing)

1. Introduction

Strong consistency of parameter estimates for the linear stochastic system was first studied for the uncorrelated noise case on the basis of the least squares method[9,10] by invoking the persistent excitation-like conditions. When the system noise is an MA sequence and the parameters to be estimated are contained in the noise model as well, the sufficient conditions for strong consistency of the estimate were given in[12] for the approximate maximum likelihood algorithm, in[2] for the modified least squares algorithm and in[3] for the least squares algorithm.

The stochastic adaptive control problem is a topic closely related to parameter identification; its special case, the adaptive tracking, was considered by Goodwin, Ramadge and Caines[7] and Sin and Goodwin[11]. Later, the adaptive tracking was simultaneously solved with the parameter estimation problem by Caines and Lafortune[1], Chen[4], and Chen and Caines[5] with the help of a control disturbed by noise. In these recent papers conditions guaranteeing both strong consistency of the estimates and the suboptimality of adaptive control were given.

This paper considers the stochastic gradient algorithm, dealing with a class of system noise including martingale difference sequence and other correlated random sequences. We give a necessary and sufficient condition for strong consistency of the parameter estimates given by the algorithm. Then for the case where the noise is an MA sequence we give, probably the weakest, sufficient conditions for the strong consistency of the algorithm estimating the unknown parameters appearing in the system and in the noise models. A comparison between various sufficient conditions is also demonstrated in the paper.

2. Statement of Problem

Consider the stochastic system with l-dimensional input $\{u_n\}$ and m-dimensional output $\{y_n\}$:

$$y_n + A_1 y_{n-1} + \cdots + A_p y_{n-p} = B_1 u_{n-1} + \cdots + B_q u_{n-q} + \varepsilon_n, \tag{1}$$

where A_i, $B_j, i = 1, \cdots, p, j = 1, \cdots, q$, are unknown matrices to be estimated.

Received June 4, 1984.

Set

$$\theta^\tau = [-A_1 \cdots -A_p \ B_1 \cdots B_q], \tag{2}$$

$$\varphi_n^\tau = [y_n^\tau \cdots y_{n-p+1}^\tau \ u_n^\tau \cdots u_{n-q+1}^\tau], \tag{3}$$

$$r_n = 1 + \sum_{i=1}^{n} \|\varphi_i\|^2, \qquad r_0 = 1. \tag{4}$$

Given any deterministic φ_0 and θ_0, we estimate θ by the stochastic gradient algorithm

$$\theta_{n+1} = \theta_n + \frac{\varphi_n}{r_n}(y_{n+1}^\tau - \varphi_n^\tau \theta_n). \tag{5}$$

The system noise ε_n is not necessarily of zero-mean and is allowed to be correlated, but we require that the following Condition A be satisfied.

A. As $n \to \infty$, $\sum_{i=0}^{n} \frac{\varphi_i}{r_i} \varepsilon_{i+1}^\tau$ tends to a finite limit S, and there exist $c > 0$ and $\delta > 0$ which may depend upon ω such that

$$\left\| S - \sum_{i=0}^{n-1} \frac{\varphi_i}{r_i} \varepsilon_{i+1}^\tau \right\| \leqslant c r_n^{-\delta}. \tag{6}$$

We now give examples of $\{\varepsilon_n\}$ satisfying Condition A.

Example 1. Let $(\Omega, \mathscr{F}, P)$ be the basic probability space and let $\{\mathscr{F}_i\}$ be a family of nondecreasing sub-σ-algebras of $\mathscr{F}$ and let ε_i and φ_i be $\mathscr{F}_i$-measurable. Suppose that $E(\varepsilon_{i+1} | \mathscr{F}_i) = 0$ and that there exist $c_0 > 0$ and $\varepsilon \in [0,1)$ (c_0 and ε may depend on ω) such that $E(\|\varepsilon_{i+1}\|^2 | \mathscr{F}_i) \leqslant c_0 r_i^\varepsilon$. Then $\{\varepsilon_i\}$ satisfies Condition A a.s. (It is worth remarking that the last condition is more general than the uniform boundedness condition $E(\|\varepsilon_{i+1}\|^2 | \mathscr{F}_i) \leqslant \sigma^2$ with constant σ^2 since $r_i \geqslant 1, \forall i \geqslant 0$.)

Proof. We need the following fact[8]. Let $A_n \triangleq 1 + \sum_{i=1}^{n} a_i$ with $a_i \geqslant 0$. Then

$$\sum_{i=1}^{\infty} \frac{a_i}{A_i^\alpha} < \infty, \qquad \forall\ \alpha > 1, \tag{7}$$

and

$$\sum_{i=1}^{\infty} \frac{a_i}{A_i} = \infty \ \text{ iff } A_i \xrightarrow[i \to \infty]{} \infty. \tag{8}$$

Since $\varepsilon < 1$ we can find $\delta > 0$ so that $2 - \varepsilon - 2\delta > 1$. By (7) it is easy to see

$$\sum_{i=1}^{\infty} E\left[\left\| \frac{\varphi_i \varepsilon_{i+1}^\tau}{r_i^{1-\delta}} \right\|^2 \middle| \mathscr{F}_i \right] \leqslant c_0 \sum_{i=1}^{\infty} \frac{\|\varphi_i\|^2}{r_i^{2-\varepsilon-2\delta}} < \infty, \qquad \text{a.s.}$$

Then by the martingale convergence theorem[6] it follows that as $n \to \infty$, $\sum_{i=1}^{n} \frac{\varphi_i \varepsilon_{i+1}^\tau}{r_i^{1-\delta}}$ converges a. s. Hence for fixed ω for any $\eta > 0$, if n_0 is large enough, we have

$$\|S_n^\delta\| < \eta,$$

whenever $n \geqslant n_0$, where

$$S_n^\delta \triangleq \sum_{i=n}^{\infty} \frac{\varphi_i \varepsilon_{i+1}^\tau}{r_i^{1-\delta}}.$$

Then summation by parts yields the desired result

$$\left\| r_n^\delta \left(S - \sum_{i=1}^{n-1} \frac{\varphi_i \varepsilon_{i+1}^\tau}{r_i} \right) \right\| = \left\| r_n^\delta \sum_{i=n}^{\infty} \frac{\varphi_i \varepsilon_{i+1}^\tau}{r_i^{1-\delta}} \cdot \frac{1}{r_i^\delta} \right\|$$

$$= \left\| r_n^\delta \sum_{i=n}^{\infty} (S_i^\delta - S_{i+1}^\delta) \frac{1}{r_i^\delta} \right\| = \left\| S_n^\delta - r_n^\delta \sum_{i=n}^{\infty} S_{i+1}^\delta \left(\frac{1}{r_i^\delta} - \frac{1}{r_{i+1}^\delta} \right) \right\|$$

$$\leqslant \eta + \eta r_n^\delta \sum_{i=n}^{\infty} \left(\frac{1}{r_i^\delta} - \frac{1}{r_{i+1}^\delta} \right) \leqslant 2\eta, \qquad \forall n \geqslant n_0.$$

Example 2. If $\{\varepsilon_i\}$ is an arbitrary random sequence with

$$\|\varepsilon_i\| \leqslant c_1 \frac{\|\varphi_i\|}{r_i^\beta}, \qquad \forall i \geqslant 0,$$

where $c_1 \geqslant 0$, $\beta > 0$ and they may be ω-dependent, then Condition A is satisfied.

Proof. Setting

$$T_n = \sum_{i=n}^{\infty} \frac{\|\varphi_i\|^2}{r_i^{1+\beta/2}},$$

we have

$$\left\| r_n^{\beta/2} \sum_{i=n}^{\infty} \frac{\varphi_i \varepsilon_{i+1}^\tau}{r_i} \right\| \leqslant c_1 r_n^{\beta/2} \sum_{i=n}^{\infty} \frac{\|\varphi_i\|^2}{r_i^{1+\beta/2}} \cdot \frac{1}{r_i^{\beta/2}}$$

$$= c_1 r_n^{\beta/2} \sum_{i=n}^{\infty} (T_i - T_{i+1}) \frac{1}{r_i^{\beta/2}}$$

$$\leqslant c_1 r_n^{\beta/2} \left[T_n \frac{1}{r_n^{\beta/2}} - \sum_{i=n}^{\infty} T_{i+1} \left(\frac{1}{r_i^{\beta/2}} - \frac{1}{r_{i+1}^{\beta/2}} \right) \right] \leqslant c_1 T_n.$$

which goes to zero by (7). Hence Condition A holds with $\delta = \beta/2$.

Section 3 will only consider $\{\varepsilon_i\}$ satisfying Condition A, while Section 5 will deal with noise of other types.

3. A Necessary and Sufficient Condition

We now prove a necessary and sufficient condition for strong consistency when $\{\varepsilon_i\}$ satisfies Condition A.

Let matrix $\Phi(n, i)$ be recursively defined by

$$\Phi(n+1, i) = \left(I - \frac{\varphi_n \varphi_n^\tau}{r_n} \right) \Phi(n, i), \qquad \Phi(i, i) = I. \tag{9}$$

Lemma 1. *For any $n \geqslant 0$ the following inequality takes place*

$$\sum_{i=0}^{n-1} \frac{\|\Phi(n, i+1)\varphi_i\|^2}{r_i} \leqslant d \tag{10}$$

with $d = mp + lq$.

Proof. The assertion of the lemma is verified by the following chain of inequalities and equalities:

$$\begin{aligned} d &= \operatorname{tr}\Phi(n, n)\Phi^\tau(n,n) \geqslant \operatorname{tr} \sum_{i=0}^{n-1} [\Phi(n, i+1)\Phi^\tau(n,i+1) - \Phi(n,i)\Phi^\tau(n, i)] \\ &= \operatorname{tr} \sum_{i=0}^{n-1} \Phi(n,i+1)[I - \Phi(i+1,i)\Phi^\tau(i+1,i)]\Phi^\tau(n,i+1) \\ &= \operatorname{tr} \sum_{i=0}^{n-1} \Phi(n,i+1) \left[I - \left(I - \frac{\varphi_i\varphi_i^\tau}{r_i}\right)\left(I - \frac{\varphi_i\varphi_i^\tau}{r_i}\right)\right] \Phi^\tau(n,i+1) \\ &= \operatorname{tr} \sum_{i=0}^{n-1} \Phi(n,i+1) \left[\frac{\varphi_i\varphi_i^\tau}{r_i} + \frac{1}{r_i}\varphi_i\left(I - \frac{\varphi_i^\tau\varphi_i}{r_i} I\right)\varphi_i^\tau\right] \Phi^\tau(n, i+1) \\ &\geqslant \operatorname{tr} \sum_{i=0}^{n-1} \Phi(n,i+1) \frac{\varphi_i\varphi_i^\tau}{r_i} \Phi^\tau(n,i+1) = \sum_{i=0}^{n-1} \frac{\|\Phi(n, i+1)\varphi_i\|^2}{r_i}. \end{aligned}$$

Theorem 1. *Let* $\{\varepsilon_i\}$ *satisfy Condition A on a subset* Ω' *of* Ω. *Then for any initial value* θ_0 *the estimate* θ_n *defined by* (5) *converges to* θ *on* Ω' *if and only if*

$$\Phi(n, 0) \xrightarrow[n\to\infty]{} 0 \qquad \forall \omega \in \Omega'. \tag{11}$$

Proof. Set

$$\tilde{\theta}_n = \theta - \theta_n. \tag{12}$$

Since

$$y_{n+1}^\tau = \varphi_n^\tau\theta + \varepsilon_{n+1}^\tau,$$

from (5) and (12) it follows that

$$\tilde{\theta}_{n+1} = \tilde{\theta}_n - \frac{\varphi_n}{r_n}(\varphi_n^\tau\tilde{\theta}_n + \varepsilon_{n+1}^\tau)$$

or equivalently,

$$\tilde{\theta}_{n+1} = \left(I - \frac{\varphi_n\varphi_n^\tau}{r_n}\right)\tilde{\theta}_n - \frac{\varphi_n}{r_n}\varepsilon_{n+1}^\tau. \tag{13}$$

Then

$$\tilde{\theta}_{n+1} = \Phi(n+1, 0)\tilde{\theta}_0 - \sum_{j=0}^{n} \Phi(n+1, j+1)\frac{\varphi_j}{r_j}\varepsilon_{j+1}^\tau. \tag{14}$$

Necessity.

On Ω' for any θ_0, θ_n converges to θ. Hence for any $\tilde{\theta}_0, \tilde{\theta}_n$ tends to zero for any $\omega \in \Omega'$. Notice that the second term on the right-hand side of (14) is independent of $\tilde{\theta}_0$; so for any $\tilde{\theta}_0$ we have

$$\Phi(n+1, 0)\tilde{\theta}_0 \xrightarrow[n\to\infty]{} 0 \qquad \forall \omega \in \Omega'$$

and this means

$$\Phi(n+1,0)\xrightarrow[n\to\infty]{}0 \qquad \forall\omega\in\Omega'.$$

Sufficiency.

It is clear that we only need to prove

$$\sum_{j=0}^{n}\Phi(n+1,j+1)\frac{\varphi_j}{r_j}\varepsilon_{j+1}^{\tau}\xrightarrow[n\to\infty]{}0 \qquad \forall\omega\in\Omega'. \tag{15}$$

Set

$$S_n=\sum_{i=0}^{n}\frac{\varphi_i}{r_i}\varepsilon_{i+1}^{\tau},\ \tilde S_n=S-S_n=\sum_{i=n+1}^{\infty}\frac{\varphi_i\varepsilon_{i+1}^{\tau}}{r_i}, \qquad S_{-1}=0. \tag{16}$$

Assume $\omega\in\Omega'$. Then $\tilde S_n\xrightarrow[n\to\infty]{}0$ and $\|\tilde S_{j-1}\|\leqslant cr_j^{-\delta}$. Hence we have

$$\left\|\sum_{j=0}^{n}\Phi(n+1,j+1)\frac{\varphi_j}{r_j}\varepsilon_{j+1}^{\tau}\right\|=\left\|\sum_{j=0}^{n}\Phi(n+1,j+1)(S_j-S_{j-1})\right\|$$

$$=\left\|S_n-\sum_{j=0}^{n}[\Phi(n+1,j+1)-\Phi(n+1,j)]S\right.$$

$$\left.+\sum_{j=0}^{n}[\Phi(n+1,j+1)-\Phi(n+1,j)]\tilde S_{j-1}\right\|$$

$$=\left\|S_n-S+\Phi(n+1,0)S+\sum_{j=0}^{n}\Phi(n+1,j+1)[I-\Phi(j+1,j)]\tilde S_{j-1}\right\|$$

$$\leqslant\|\tilde S_n\|+\|\Phi(n+1,0)S\|+c\sum_{j=0}^{n}\frac{\|\Phi(n+1,j+1)\varphi_j\|}{r_j^{1/2}}\cdot\frac{\|\varphi_j\|}{r_j^{1/2+\delta}}$$

$$\leqslant\|\tilde S_n\|+\|\Phi(n+1,0)S\|+c\sum_{j=0}^{N}\frac{\|\Phi(n+1,j+1)\varphi_j\|}{r_j^{1/2}}\cdot\frac{\|\varphi_j\|}{r_j^{1/2+\delta}}$$

$$+c\left(\sum_{j=N+1}^{n}\frac{\|\Phi(n+1,j+1)\varphi_j\|^2}{r_j}\right)^{1/2}\left(\sum_{j=N+1}^{n}\frac{\|\varphi_j\|^2}{r_j^{1+2\delta}}\right)^{1/2},$$

which tends to zero by Lemma 1 and (7) if we first let $n\to\infty$ and then let $N\to\infty$.

4. A Comparison between Sufficient Conditions

In this section we give some sufficient conditions guaranteeing $\Phi(n,0)\xrightarrow[n\to\infty]{}0$. We first list three different conditions usually used for proving consistency.

a) There is a positive definite matrix R such that[3]

$$\frac{1}{n}\sum_{i=1}^{n}\varphi_i\varphi_i^{\tau}\xrightarrow[n\to\infty]{}R.$$

b) $r_n\xrightarrow[n\to\infty]{}\infty$ and the ratio of the maximum to minimum eigenvalues of

$$\sum_{i=1}^{n}\varphi_i\varphi_i^{\tau}+\frac{1}{d}I$$

is bounded:

$$\lambda_{\max}^n/\lambda_{\min}^n \leqslant \gamma, \qquad \forall n \geqslant 0$$

where d denotes the dimension of φ_i and γ may depend on ω[10].

c) There exist α,β and T which may depend on ω and $T>0, 0<\alpha<\infty$ and $0<\beta<\infty$ such that[2,4,5]

$$\sum_{i=m(t)}^{m(t+\alpha)} \frac{\varphi_i\varphi_i^\tau}{r_i} \geqslant \beta I, \qquad \forall t \geqslant T,$$

where

$$m(t) = \max[n; t_n \leqslant t], \qquad t \geqslant 0,$$

$$t_n = \sum_{i=0}^{n-1} \beta_i, \ t_0 = 0, \ \beta_i = \|\varphi_i\|^2/r_i.$$

In [4] it is proved that Condition a) implies Conditiion b) which in turn implies Condition c). The next theorem establishes the relationship between Condition c) and (11).

Theorem 2. *For any fixed ω Condition c) implies*

$$\Phi(n,0) \xrightarrow[n\to\infty]{} 0.$$

Proof. We notice at once that if Condition c) holds, then $r_n \xrightarrow[n\to\infty]{} \infty$, since otherwise t_n goes to a finite limit as $n\to\infty$ and $m(t)$ equals ∞ for sufficiently large t. But this will make Condition c) impossible.

It is obvious that $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$ is equivalent to $x_n \xrightarrow[n\to\infty]{} 0$ for any x_0 where x_n denotes the solution of the difference equation

$$x_{n+1} = \left(I - \frac{\varphi_n\varphi_n^\tau}{r_n}\right) x_n. \tag{17}$$

It is easy to see that $\|x_n\|$ converges to a finite limit as $n\to\infty$ and

$$\sum_{i=0}^{\infty} \frac{\|\varphi_i^\tau x_i\|^2}{r_i} < \infty, \tag{18}$$

since by (17)

$$\|x_{n+1}\|^2 \leqslant \|x_n\|^2 - \frac{\|\varphi_n^\tau x_n\|^2}{r_n}.$$

We now rewrite (17) as

$$x_{n+1} = x_0 - J_{n+1}, \tag{19}$$

where

$$J_{n+1} = \sum_{i=0}^{n} \frac{\varphi_i\varphi_i^\tau}{r_i} x_i. \tag{20}$$

Denote by $J(t)$ the linear interpolating function of $\{J_n\}$ with interpolation length equal to $\{\beta_n\}$, i. e.

$$\begin{cases} J(t) = \dfrac{t_{n+1}-t}{\beta_n} J_n + \dfrac{t-t_n}{\beta_n} J_{n+1}, \ t\in[n, n+1), \\ J(t) = J_n. \end{cases} \tag{21}$$

Similarly, $x(t)$ denotes the linear interpolation of $\{x_n\}$. Then

$$x(t) = x_0 - J(t). \tag{22}$$

It is worth noting that

$$x(t_n) = x_n. \tag{23}$$

Define a family $\{x_n(t)\}$ of continuous functions by

$$x_n(t) = x(t+n). \tag{24}$$

Since $\|x_n\|$ converges, the family $\{x_n(t)\}$ is uniformly bounded. We prove that it is equi-continuous.

By (8), from $r_n \to \infty$ it follows that $t_n \to \infty$ and hence $m(t) \xrightarrow[n\to\infty]{} \infty$. For any $t \in [t_n, t_{n+1}]$, from (21) it is known that

$$\|J(t) - J(t_n)\| = \frac{t - t_n}{\beta_n}\|J_{n+1} - J_n\| \leqslant \left\|\frac{\varphi_n\varphi_n^\tau}{r_n}x_n\right\| \leqslant \frac{\|\varphi_n^\tau x_n\|}{r_n^{1/2}}.$$

Then for any $t \geqslant 0$ and $\Delta > 0$ we have

$$\begin{aligned}\|x_n(t+\Delta) - x_n(t)\| &\leqslant \|J(t+n+\Delta) - J(t_{m(t+n+\Delta)})\| \\ &\quad + \|J(t+n) - J(t_{m(t+n)})\| + \|J(t_{m(t+n+\Delta)}) - J(t_{m(t+n)})\| \\ &\leqslant \frac{\|\varphi^\tau_{m(t+n+\Delta)}x_{m(t+n+\Delta)}\|}{r^{1/2}_{m(t+n+\Delta)}} + \frac{\|\varphi^\tau_{m(t+n)}x_{m(t+n)}\|}{r^{1/2}_{m(t+n)}} \\ &\quad + \left(\sum_{j=m(t+n)}^{m(t+n+\Delta)-1}\beta_j\right)^{\frac{1}{2}}\left(\sum_{j=m(t+n)}^{m(t+n+\Delta)-1}\frac{\|\varphi_j^\tau x_j\|^2}{r_j}\right)^{\frac{1}{2}} \xrightarrow[n\to\infty]{} 0.\end{aligned} \tag{25}$$

The last convergence follows from (18) and $m(t) \xrightarrow[t\to\infty]{} \infty$.

Completely similar result can be obtained for $\Delta < 0$. Then by the Arzela-Ascoli theorem, from $\{x_n(t)\}$ we can select a subsequence $\{x_{n_k}(t)\}$ which uniformly converges to a continuous function $x^0(t)$ in any finite interval. Again by (25) we see that the limit function $x^0(t)$, in fact, is a constant vector $x^0(t) \equiv x^0$.

We now show $x^0 = 0$. For any integer $i \in [0, m(t+n_k+\alpha) - m(t+n_k)]$ we have for $t \geqslant 1$

$$\begin{aligned} t - 1 &\leqslant t_{m(t+n_k)+1} - n_k - 1 \leqslant t_{m(t+n_k)+1} - \beta_{m(t+n_k)} - n_k \\ &= t_{m(t+n_k)} - n_k \leqslant t_{m(t+n_k)+i} - n_k \\ &\leqslant t_{m(t+n_k)} + \sum_{j=m(t+n_k)}^{m(t+n_k)+i-1}\beta_j - n_k \\ &\leqslant t_{m(t+n_k)} + \sum_{j=m(t+n_k)}^{m(t+n_k+\alpha)}\beta_j - n_k \leqslant t + \alpha + 2,\end{aligned}$$

i.e.

$$t_{m(t+n_k)+i} - n_k \in [t-1, t+\alpha+2].$$

Then the convergence

$$x_{n_k}(t_{m(t+n_k)+i} - n_k) \xrightarrow[k\to\infty]{} x^0$$

or

$$x_{m(t+n_k)+i} \xrightarrow[k\to\infty]{} x^0$$

is uniform in the integers i belonging to $[0, m(t+n_k+\alpha)-m(t+n_k)]$.

Hence we conclude that

$$\left\|\sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \frac{\varphi_i\varphi_i^\tau}{r_i}(x_i-x^0)\right\| \leqslant \sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \beta_i\|x_i-x^0\|$$
$$\leqslant \max_{0\leqslant i\leqslant m(t+n_k+\alpha)-m(t+n_k)} \|x_{m(t+n_k)+i}-x^0\|\cdot(2+\alpha)\xrightarrow[k\to\infty]{} 0.$$

From here, Condition c) and (18) we have

$$\beta\|x^0\|^2 \leqslant x^{0\tau} \sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \frac{\varphi_i\varphi_i^\tau}{r_i} x^0$$
$$= x^{0\tau} \sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \frac{\varphi_i\varphi_i^\tau}{r_i} x_i + x^{0\tau} \sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \frac{\varphi_i\varphi_i^\tau}{r_i}(x^0-x_i)$$
$$\leqslant \|x^0\|\sqrt{2+\alpha}\left(\sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \frac{\|\varphi_i x_i\|^2}{r_i}\right)^{\frac{1}{2}}$$
$$+ \|x^0\|\cdot\left\|\sum_{i=m(t+n_k)}^{m(t+n_k+\alpha)} \frac{\varphi_i\varphi_i^\tau}{r_i}(x_i-x^0)\right\|\xrightarrow[k\to\infty]{} 0.$$

This means that $x^0=0$ and

$$x(t+n_k)\xrightarrow[k\to\infty]{} 0.$$

Then we can select a subsequence $\{x_{m_k}\}$ from $\{x_n\}$ so that $x_{m_k}\xrightarrow[k\to\infty]{} 0$. This together with the fact that $\|x_n\|$ tends to a finite limit shows $x_n \to 0$. By the arbitrariness of x_0 we conclude

$$\Phi(n, 0)\xrightarrow[n\to\infty]{} 0.$$

Thus, we have proved that if one of Conditions a), b), c) holds, then θ_n given by (5) is consistent.

5. Other Types of System Noise

We now discuss the case where ε_i is driven by a martingale difference sequence

$$\varepsilon_n = w_n + C_1 w_{n-1} + \cdots + C_r w_{n-r} \tag{25}$$

with unknown matrices $C_k, k=1,\cdots,r$, which we also need to estimate.

For the present case we change the notations introduced in (2)—(5) to the following:

$$\theta^\tau = [-A_1\cdots -A_p B_1\cdots B_q C_1\cdots C_r], \tag{27}$$
$$\varphi_n^\tau = [y_n^\tau\cdots y_{n-p+1}^\tau u_n^\tau\cdots u_{n-q+1}^\tau y_n^\tau - \varphi_{n-1}^\tau\theta_{n-1}\cdots y_{n-r+1}^\tau - \varphi_{n-r}^\tau\theta_{n-r}]. \tag{28}$$

Then Algorithm (5) will give estimates for all of A_i, B_j and C_k, $i=1,\cdots,p, j=1,\cdots, q, k=1,\cdots,r$.

As in Example 1 of Section 2 we assume that $\{\mathscr{F}_n\}$ is a nondecreasing family of sub-σ-algebras of $\mathscr{F}$, w_n and φ_n are $\mathscr{F}_n$-measurable and that

$$E(w_n|\mathscr{F}_{n-1})=0,\ E(\|w_n\|^2|\mathscr{F}_{n-1})\leqslant c_0 r_{n-1}^{\varepsilon}, \tag{29}$$

where c_0 and ε may depend on ω and $c_0>0$, $\varepsilon\in[0,1)$. (29)

Set

$$\xi_n = y_n - w_n - \theta_{n-1}^{\tau}\varphi_{n-1}, \tag{30}$$

$$\varphi_n^0=[y_n^{\tau}\cdots y_{n-p+1}^{\tau}u_n^{\tau}\cdots u_{n+q+1}w_n^{\tau}\cdots w_{n-r+1}^{\tau}], \tag{31}$$

$$\varphi_n^{\xi\tau}=[0\cdots 0\quad 0\cdots 0\xi_n^{\tau}\cdots\xi_{n-r+1}^{\tau}]. \tag{32}$$

Obviously, we have

$$\varphi_n=\varphi_n^0+\varphi_n^{\xi}. \tag{33}$$

Theorem 3. *If $r=0$, or $r>0$ but $C(z)-\frac{1}{2}I$ is strictly positive real, then $\Phi(n,0)\xrightarrow[n\to\infty]{}0$ implies $\theta_n\to\theta$, where the transfer matrix $C(z)$ is defined by*

$$C(z)=I+C_1z+\cdots+C_rz^r.$$

Proof. From Lemma 1 in[5] we have

$$\sum_{n=0}^{\infty}\frac{\|\varphi_n^{\xi}\|^2}{r_n}<\infty,\qquad \text{a. s.} \tag{34}$$

For ε_n expressed by (26) we have

$$y_{n+1}=\theta^{\tau}\varphi_n^0+w_{n+1}.$$

Then

$$\begin{aligned}\theta_{n+1}&=\theta_n+\frac{\varphi_n}{r_n}(\varphi_n^{0\tau}\theta+w_{n+1}^{\tau}-\varphi_n^{\tau}\theta_n)\\&=\theta_n+\frac{\varphi_n}{r_n}(\varphi_n^{\tau}\theta-\varphi_n^{\xi\tau}\theta+w_{n+1}^{\tau}-\varphi_n^{\tau}\theta_n)\\&=\theta_n+\frac{\varphi_n}{r_n}(\varphi_n^{\tau}\tilde\theta_n-\varphi_n^{\xi\tau}\theta+w_{n+1}^{\tau}),\end{aligned}$$

and

$$\tilde\theta_{n+1}=\left(I-\frac{\varphi_n\varphi_n^{\tau}}{r_n}\right)\tilde\theta_n+\frac{\varphi_n\varphi_n^{\xi\tau}}{r_n}\theta-\frac{\varphi_n}{r_n}w_{n+1}^{\tau}.$$

From here it follows that

$$\tilde\theta_{n+1}=\Phi(n+1,0)\tilde\theta_0+\sum_{j=0}^{n}\Phi(n+1,j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\theta-\sum_{j=0}^{n}\Phi(n+1,j+1)\frac{\varphi_j}{r_j}w_{j+1}^{\tau}. \tag{35}$$

It is easy to see that if φ_n is defined by (28), then (10) still holds true with $d=mp+lq+mr$. Thus, by (34) and $\Phi(n,0)\xrightarrow[n\to\infty]{}0$ we have

$$\begin{aligned}&\left\|\sum_{j=0}^{n}\Phi(n+1,j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\right\|\\&\leqslant\left(\sum_{j=N+1}^{n}\frac{\|\Phi(n+1,j+1)\varphi_j\|^2}{r_j}\right)^{\frac12}\left(\sum_{j=N+1}^{n}\frac{\|\varphi_j^{\xi}\|^2}{r_j}\right)^{\frac12}\end{aligned}$$

$$+\left\|\sum_{j=0}^{N}\Phi(n+1,\ j+1)\frac{\varphi_j\varphi_j^{\xi}}{r_j}\right\|$$

$$\leqslant\sqrt{d}\left(\sum_{j=N+1}^{n}\frac{\|\varphi_j^{\xi}\|^2}{r_j}\right)^{\frac{1}{2}}+\left\|\sum_{j=0}^{N}\Phi(n+1,j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\right\|\to 0 \tag{36}$$

as $n\to\infty$ and then $N\to\infty$.

From Example 1 in Section 1 we know that there are $c>0$ and $\delta>0$ such that

$$\left\|\sum_{i=n}^{\infty}\frac{\varphi_i}{r_i}w_{i+1}^{\tau}\right\|\leqslant cr_n^{-\delta},\qquad \forall n.$$

Then in the sufficiency part of Theorem 1, if ε_i is replaced by w_i,(15) is still valid, i.e.

$$\sum_{j=0}^{n}\Phi(n+1,\ j+1)\frac{\varphi_j}{r_j}w_{j+1}^{\tau}\xrightarrow[n\to\infty]{}0.$$

Hence from (35), (36) we assert that $\tilde{\theta}_{n+1}\xrightarrow[n\to\infty]{}0$.

Remark. It is worth noting that all conclusions including Theorem 2 in Section 4 remain valid if φ_n is given by (28).

By definition φ_n depends upon past estimates $\theta_i, i\leqslant n-1$, and so does $\Phi(n,0)$. We now give a condition which is independent of the estimate and is equivalent to $\Phi(n,0)\xrightarrow[n\to\infty]{}0$.

Consider

$$\Phi^0(n+1,i)=\left(I-\frac{\varphi_n^0\varphi_n^{0\tau}}{r_n^0}\right)\Phi^0(n,i),\ \Phi^0(i,i)=I \tag{37}$$

where φ_n^0 is defined by (31) and

$$r_n^0=1+\sum_{i=1}^{n}\|\varphi_i^0\|^2. \tag{38}$$

Theorem 4. *If* $C(z)-\frac{1}{2}I$ *is strictly positive real, then* $\Phi(n,0)\xrightarrow[n\to\infty]{}0$ *iff* $\Phi^0(n,\ 0)\xrightarrow[n\to\infty]{}0$.

Proof. Suppose $\Phi(n,\ 0)\xrightarrow[n\to\infty]{}0$. We first show $r_n\xrightarrow[n\to\infty]{}\infty$.

Without loss of generality assume $\|\varphi_0\|\neq 1$. From the following chain of equalities

$$\begin{aligned}\det\Phi(n+1,\ 0)&=\prod_{i=0}^{n}\det\Phi(i+1,\ i)\\&=\prod_{i=0}^{n}\det\left(I-\frac{\varphi_i\varphi_i^{\tau}}{r_i}\right)=(1-\|\varphi_0\|^2)\prod_{i=1}^{n}\left(1-\frac{\|\varphi_i\|^2}{r_i}\right)\\&=(1-\|\varphi_0\|^2)\prod_{i=1}^{n}r_{i-1}/r_i\\&=(1-\|\varphi_0\|^2)/r_n\end{aligned}$$

we conclude $r_n \to \infty$ by $\Phi(n,0) \to 0$.

From (33), (34) and the Kronecker lemma it is easy to see

$$\frac{r_n^0}{r_n} = \frac{r_n - 2\sum_{i=1}^{n} \varphi_i^\tau \varphi_i^{\xi\tau} + \sum_{i=1}^{n} \|\varphi_i^\xi\|^2}{r_n} \xrightarrow[n\to\infty]{} 1. \tag{39}$$

Then

$$\sum_{i=0}^{\infty} \frac{\|\varphi_i^\xi\|^2}{r_i^0} < \infty. \tag{40}$$

Noticing

$$\Phi^0(n+1,0) = \left(I - \frac{\varphi_n\varphi_n^\tau}{r_n}\right)\Phi^0(n,0) + \left(\frac{\varphi_n\varphi_n^\tau}{r_n} - \frac{\varphi_n^0\varphi_n^{0\tau}}{r_n^0}\right)\Phi^0(n,0)$$

we obtain

$$\begin{aligned}
\Phi^0(n+1,0) &= \Phi(n+1,0) + \sum_{j=0}^{n} \Phi(n+1,j+1)\left(\frac{\varphi_j\varphi_j^\tau}{r_j} - \frac{\varphi_j^0\varphi_j^{0\tau}}{r_j^0}\right)\Phi^0(j,0) \\
&= \Phi(n+1,0) + \sum_{j=0}^{n} \Phi(n+1,j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\Phi^0(j,0) \\
&\quad + \sum_{j=0}^{n} \Phi(n+1,j+1)\frac{\varphi_j^\xi\varphi_j^{0\tau}}{r_j^0}\Phi^0(j,0) \\
&\quad + \sum_{j=0}^{n} \frac{\Phi(n+1,j+1)\varphi_j}{r_j^{1/2}}\left(\sqrt{\frac{r_j^0}{r_j}} - \sqrt{\frac{r_j}{r_j^0}}\right)\frac{\varphi_j^{0\tau}\Phi^0(j,0)}{(r_j^0)^{1/2}}.
\end{aligned} \tag{41}$$

By (34), Lemma 1, $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$ and the fact that $\|\Phi^0(j,0)\| \leqslant 1$, it is clear that

$$\begin{aligned}
&\left\|\sum_{j=0}^{n} \Phi(n+1,j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\Phi^0(j,0)\right\| \\
&\quad \leqslant \left\|\sum_{j=0}^{N} \Phi(n+1,j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\Phi^0(j,0)\right\| \\
&\quad + \left(\sum_{j=N+1}^{n} \frac{\|\Phi(n+1,j+1)\varphi_j\|^2}{r_j}\right)^{\frac{1}{2}}\left(\sum_{j=N+1}^{n} \frac{\|\varphi_j^\xi\|^2}{r_j}\right)^{\frac{1}{2}} \xrightarrow[\substack{n\to\infty \\ N\to\infty}]{} 0.
\end{aligned} \tag{42}$$

Since

$$\begin{aligned}
d \geqslant \operatorname{tr}\Phi^{0\tau}(N,0)\Phi^0(N,0) &\geqslant \sum_{j=N}^{\infty} \operatorname{tr}[\Phi^{0\tau}(j,0)\Phi^0(j,0) - \Phi^{0\tau}(j+1,0)\Phi^0(j+1,0)] \\
&\geqslant \sum_{j=N}^{\infty} \frac{\|\Phi^{0\tau}(j,0)\varphi_j^0\|^2}{r_j^0},
\end{aligned} \tag{43}$$

by (40),(43) and the fact that $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$ and $\|\Phi(n,j)\| \leqslant 1$ for $n \geqslant j$ we find that

$$\left\|\sum_{j=0}^{n}\Phi(n+1,j+1)\frac{\varphi_j^{\xi}\varphi_j^{0\tau}}{r_j^0}\Phi^0(j,0)\right\| \leqslant \left\|\sum_{j=0}^{N}\Phi(n+1,j+1)\frac{\varphi_j^{\xi}\varphi_j^{0\tau}}{r_j^0}\Phi^0(j,0)\right\|$$

$$+\left(\sum_{j=N+1}^{n}\frac{\|\varphi_j^{\xi}\|^2}{r_j^0}\right)^{\frac{1}{2}}\left(\sum_{j=N+1}^{n}\frac{\|\Phi^{0\tau}(j,0)\varphi_j^0\|^2}{r_j^0}\right)^{\frac{1}{2}} \xrightarrow[\substack{n\to\infty\\ N\to\infty}]{} 0. \tag{44}$$

By (39) it is immediate that for all $j \geqslant N$,

$$\left|\sqrt{\frac{r_j^0}{r_j}}-\sqrt{\frac{r_j}{r_j^0}}\right| \leqslant 2,$$

if N is large enough; then

$$\left\|\sum_{j=0}^{n}\frac{\Phi(n+1,j+1)\varphi_j}{r_j^{1/2}}\left(\sqrt{\frac{r_j^0}{r_j}}-\sqrt{\frac{r_j}{r_j^0}}\right)\frac{\varphi_j^{0\tau}\Phi^0(j,0)}{r_j^{0\,1/2}}\right\|$$

$$\leqslant \left\|\sum_{j=0}^{N}\frac{\Phi(n+1,j+1)\varphi_j}{r_j^{1/2}}\left(\sqrt{\frac{r_j^0}{r_j}}-\sqrt{\frac{r_j}{r_j^0}}\right)\frac{\varphi_j^{0\tau}\Phi^0(j,0)}{(r_j^0)^{1/2}}\right\|$$

$$+2\left(\sum_{j=N+1}^{n}\frac{\|\Phi(n+1,j+1)\varphi_j\|^2}{r_j}\right)^{\frac{1}{2}}\left(\sum_{j=N+1}^{n}\frac{\|\varphi_j^{0\tau}\Phi^0(j,0)\|^2}{(r_j^0)}\right)^{\frac{1}{2}} \xrightarrow[\substack{n\to\infty\\ N\to\infty}]{} 0. \tag{45}$$

Combining (42),(44) and (45), from (41) we conclude that $\Phi^0(n,0)\xrightarrow[n\to\infty]{}0$, if $\Phi(n,0)\xrightarrow[n\to\infty]{}0$.

Conversely, if $\Phi^0(n,0)\xrightarrow[n\to\infty]{}0$, then a similar argument leads to $r_n^0\xrightarrow[n\to\infty]{}\infty$, and this implies $r_n\xrightarrow[n\to\infty]{}\infty$, since if the limit of r_n as $n\to\infty$ were finite then by (34) the series $\sum_{j=1}^{\infty}\|\varphi_j^{\xi}\|^2$ and $\sum_{i=1}^{\infty}\varphi_i^{\tau}\varphi_i^{\xi}$ would converge and this contradicts $r_n^0\xrightarrow[n\to\infty]{}\infty$ because

$$r_n^0 = r_n - 2\sum_{i=1}^{n}\varphi_i^{\tau}\varphi_i^{\xi} + \sum_{i=1}^{n}\|\varphi_i^{\xi}\|^2. \tag{46}$$

Thus, (39) still holds, and by the expression

$$\Phi(n+1,0) = \Phi^0(n+1,0) + \sum_{j=0}^{n}\Phi^0(n+1,j+1)\left(\frac{\varphi_j^0\varphi_j^{0\tau}}{r_j^0}-\frac{\varphi_j\varphi_j^{\tau}}{r_j}\right)\Phi(j,0),$$

an argument similar to that discussed above proves $\Phi(n,0)\xrightarrow[n\to\infty]{}0$.

In this paper we have only considered the discrete-time system. Similar results for the continuous-time system will be published elsewhere.

References

[1] Caines, P. E. and Lafortune, S., Adaptive control with recursive identification for stochastic linear system, 1982, *IEEE Trans, Autom. Control* (to appear).

[2] Chen, H. F., Strong consistency of recursive identification under correlated noise, *J. of Systems Science and Mathematical Sciences*, **1**: 1(1981), 34—52.

[3] Chen, H. F., Strong consistency and convergence rate of least squares identification, *Scientia Sinica*

(Series A), **25**: 7(1982), 771—784.
[4] Chen, H. F., Recursive system identification and adaptive control by use of the modified least squares algorithm, *SIAM J. Control and Optimization*, **22**: 5(1984).
[5] Chen, H. F. and Caines, P. E., Strong consistency of the stochastic gradient algorithm of adaptive control, 1983, *IEEE Trans. Autom. Control* (to appear).
[6] Chow, Y. S., Local convergence of martingales and the law of large numbers, *The Annals of Mathematical Statistics*, **36**(1965), 552—558.
[7] Goodwin, G. G., Ramadge P. T. and Caines, P. E., Discrete-time stochastic adaptive control, *SIAM J. Control and Optimization*, **19**: 6(1981), 829—843
[8] Hardy, G. H., J. E. Littlewood and G. Polya, Inequalities, Cambridge University Press, 1934.
[9] Ljung, L., Consistency of the least squares identification method, *IEEE Trans. Autom. Control*, AC-21 (1976), 779—781.
[10] Moore, J. B., On strong consistency of least squares identification algorithm, *Automatica*, **14** (1978), 505—509.
[11] Sin, K. S. and Goodwin, G. C., Stochastic adaptive control using a modified least squares algorithm, *Automatica*, **18** (1982), 315—321.
[12] Solo, V., The convergence of AML, *IEEE Trans. Autom. Control*, AC-**24**: 6(1979), 958—962.

离散时间随机系统参数估计的强一致性

陈翰馥 郭雷
(中国科学院系统科学研究所)

摘 要

对噪声具有实际重要性的一类线性随机控制系统，本文对由随机梯度算法给出的参数估计得到了一个使它为强一致的充分必要条件。同时对 ARMA 噪声情形，证明这个条件比熟知的持续激励条件为弱。