Automatica, Vol. 25, No. 1, pp. 147–151, 1989
Printed in Great Britain.

0005-1098/89 $3.00 + 0.00
Pergamon Press plc

Brief Paper

Consistent Order Estimation for Linear Stochastic Feedback Control Systems (CARMA Model)*

LEI GUO,† HAN-FU CHEN† and JI-FENG ZHANG†

Key Words—Stochastic systems; discrete time systems; feedback control; adaptive control; convergence analysis; (order estimation).

Abstract—A new criterion CIC is introduced to estimate orders (p_0, q_0, r_0) of the linear stochastic feedback control system with correlated noise described by a CARMA model. It is proved that the estimate is strongly consistent when the upper bounds for p_0, q_0 and r_0 are available, but neither the stability condition nor the ergodicity of the input and output are imposed on the system.

1. *Introduction*

THE ORDER ESTIMATE for an ARMA process is one of the important problems in time series analysis. The estimates (p_n, r_n) for unknown orders (p_0, r_0) are usually given by minimizing some criterion, for example, AIC(p, r) (Akaike, 1969), BIC(p, r) (Akaike, 1977) and ΦIC(p, r) (Hannan and Quinn, 1979). But all these results cannot be applied to the feedback control system, described by the so-called CARMA model, which essentially differs from the ARMA model by additional control terms which are crucial for all real control systems and depend upon the past input and output.

In a recent paper (Chen and Guo, 1987b), having introduced a new criterion to be minimized, the authors obtained consistent estimates for orders of the multidimensional feedback control system with uncorrelated noise. In this paper by introducing a new criterion CIC, we have generalized these results to the correlated noise case, i.e. we have obtained strongly consistent estimates (p_n, q_n, r_n) for orders (p_0, q_0, r_0) of the CARMA process. At the same time, the criterion and conditions used in Chen and Guo (1987b) are simplified.

2. *Statement of problem*

Let the l-input, m-output stochastic control system be described by the following CARMA model

$$\begin{aligned} &A(z)y_n = B(z)u_n + C(z)w_n, \quad n \geq 0;\\ &y_n = w_n = 0, \quad u_n = 0, \quad n < 0 \end{aligned} \tag{1}$$

where w_n is an m-dimensional driven noise, $A(z)$, $B(z)$ and $C(z)$ are matrix polynomials in the shift-back operator z

$$A(z) = \mathrm{I} + A_1 z + \cdots + A_{p_0} z^{p_0}, \quad p_0 \geq 0, \tag{2a}$$

$$B(z) = B_1 z + \cdots + B_{q_0} z^{q_0}, \quad q_0 \geq 0, \tag{2b}$$

$$C(z) = \mathrm{I} + C_1 z + \cdots + C_{r_0} z^{r_0}, \quad r_0 \geq 0 \tag{2c}$$

with unknown orders p_0, q_0 and r_0 and unknown matrix coefficients A_i, B_j and C_k $(1 \leq i \leq p_0, 1 \leq j \leq q_0, 1 \leq k \leq r_0)$.

In the sequel, we denote by $\lambda_{\min}(X)$ the minimum eigenvalue of a matrix X, and by the norm $\|X\|$ we mean the maximum singular value of X.

We make the following assumptions.

(H_1). The driven noise is a martingale difference sequence with respect to a non-decreasing family of σ-algebras $\{\mathscr{F}_n\}$ and such that

$$\sup_n E[\|w_{n+1}\|^\beta \mid \mathscr{F}_n] < \infty, \quad \text{a.s. for some } \beta \geq 2.$$

(H_2). For any $n \geq 1$, u_n is $\mathscr{F}_n$-measurable.

(H_3). The transfer matrix $C^{-1}(z) - \frac{1}{2}\mathrm{I}$ is strictly positive real, i.e.

$$C^{-1}(e^{i\theta}) + C^{-\tau}(e^{-i\theta}) - \mathrm{I} > 0, \quad \forall \theta \in [0, 2\pi].$$

(H_4). The true orders (p_0, q_0, r_0) belong to a known finite set M:

$$M \triangleq \{(p, q, r): 0 \leq p \leq p^*, 0 \leq q \leq q^*, 0 \leq r \leq r^*\}.$$

(H_5). A sequence of positive numbers $\{a_n\}$ can be found such that

$$\frac{(\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}}{a_n} \xrightarrow[n\to\infty]{} 0, \quad \text{a.s. for some } \alpha > 1, \tag{3}$$

and

$$\frac{a_n}{\lambda_{\min}^{(p,q,r)}(n)} \xrightarrow[n\to\infty]{} 0, \text{ a.s.} \quad \forall (p, q, r) \in M^* \tag{4}$$

where $\delta(.)$ is the Dirac function

$$\delta(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases},$$

and M^* denotes the set consisting of three points:

$$M^* = \{(p_0, q^*, r^*), (p^*, q_0, r^*), (p^*, q^*, r_0)\} \tag{5}$$

and where $\lambda_{\min}^{(p,q,r)}(n)$ denotes the minimum eigenvalue of

$$\sum_{i=0}^{n-1} \varphi_i^0(p, q, r)\varphi_i^{0\tau}(p, q, r) - \frac{1}{d}\mathrm{I}, \quad (d = mp^* + lq^* + mr^*), \tag{6}$$

with

$$\varphi_n^0(p, q, r) \triangleq \{y_n^\tau \dots y_{n-p+1}^\tau \quad u_n^\tau \dots u_{n-q+1}^\tau \quad w_n^\tau \dots w_{n-r+1}^\tau\}^\tau, \tag{7}$$

and

$$\bar{s}_n^0 = \sum_{i=0}^{n-1} \|\varphi_i^0(p^*, q^*, r^*)\|^2 + 1. \tag{8}$$

System (1) under Assumptions (H_1)–(H_5) is, generally speaking, neither stationary nor ergodic because (i) the system input u_n may be an arbitrary $\mathscr{F}_n$-measurable function; (ii) the matrix polynomial $A(z)$ may be unstable, i.e. zeros of $\det A(z)$ may lie outside the closed unit disk and (iii) the

* Received 24 April 1987; revised 10 November 1987; received in final form 23 June 1988. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Y. Sunahara under the direction of Editor P. C. Parks.

† Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China.

system (1) is causal, i.e. the process vanishes for negative time n. Therefore, the usual treatments and criteria developed for estimating orders of stationary ARMA processes (Hannan and Rissanen, 1982; Hannan and Kavalieris, 1984) are not applicable in the present situation.

In Section 3 for a large class of adaptive systems we shall specify $\{a_n\}$ and show that Assumption (H_5) is indeed satisfied.

We introduce the regression vector

$$\bar{\varphi}_n = [y_n^\tau \ldots y_{n-p^*+1}^\tau \quad u_n^\tau \ldots u_{n-q^*+1}^\tau \quad \hat{w}_n^\tau \ldots \hat{w}_{n-r^*+1}^\tau]^\tau, \tag{9}$$

corresponding to the system of largest possible orders, where the estimate $\hat{w}_n$ for w_n is recursively defined as follows:

$$\hat{w}_n = y_n - \bar{\theta}_n^\tau \bar{\varphi}_{n-1}, \quad n \geqslant 0; \quad \hat{w}_n = 0, \quad n < 0, \tag{10}$$

$$\begin{aligned}\bar{\theta}_{n+1} &= \bar{\theta} + \bar{a}_n \bar{P}_n \bar{\varphi}_n (y_{n+1}^\tau - \bar{\varphi}_n^\tau \bar{\theta}_n),\\ \bar{P}_{n+1} &= \bar{P}_n - \bar{a}_n \bar{P}_n \bar{\varphi}_n \bar{\varphi}_n^\tau \bar{P}_n, \quad \bar{a}_n = (1 + \bar{\varphi}_n^\tau \bar{P}_n \bar{\varphi}_n)^{-1},\end{aligned} \tag{11}$$

with initial value $\bar{\theta}_0$ arbitrarily chosen and $\bar{P}_0 = dI$, where d is given in (6).

For any $(p, q, r) \in M$ set

$$\theta(p, q, r) = [-A_1 \ldots -A_p \quad B_1 \ldots B_q \quad C_1 \ldots C_r]^\tau \tag{12}$$

where by definition

$$A_i = 0, \; B_j = 0, \; C_k = 0 \quad \text{for} \quad i > p_0, \; j > q_0, \; k > r_0. \tag{13}$$

The extended least squares estimate

$$\theta_n(p, q, r) = [-A_{ln} \ldots -A_{pn} \quad B_{ln} \ldots B_{qn} \quad C_{ln} \ldots C_{rn}]^\tau, \tag{14}$$

for $\theta(p, q, r)$ at time n is given by

$$\theta_n(p, q, r) = \left(\sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi_i^\tau(p, q, r) + \frac{1}{d}\mathrm{I}\right)^{-1} \sum_{i=0}^{n-1} \varphi_i(p, q, r) y_{i+1}^\tau, \tag{15}$$

where

$$\varphi_n(p, q, r) \triangleq [y_n^\tau \ldots y_{n-p+1}^\tau u_n^\tau \ldots u_{n-q+1}^\tau \hat{w}_n^\tau \ldots \hat{w}_{n-r+1}^\tau]^\tau. \tag{16}$$

We introduce a new information criterion CIC(p, q, r), where the first "C" emphasizes that the criterion is designed for feedback control systems:

$$\mathrm{CIC}(p, q, r)_n = \sigma_n(p, q, r) + (p + q + r)a_n, \tag{17}$$

where the subscript n denotes the data size, and where a_n is given in Assumption (H_5) and $\sigma_n(p, q, r)$ is a residual given by

$$\sigma_n(p, q, r) = \sum_{i=0}^{n-1} \|y_{i+1} - \theta_n^\tau(p, q, r)\varphi_i(p, q, r)\|^2. \tag{18}$$

Finally, the estimate (p_n, q_n, r_n) for (p_0, q_0, r_0) is given by minimizing CIC$(p, q, r)_n$,

$$(p_n, q_n, r_n) = \underset{(p,q,r)\in M}{\arg\min}\, \mathrm{CIC}(p, q, r)_n. \tag{19}$$

3. *Strong consistency of order estimates*

Theorem 1. Under conditions (H_1)–(H_5) the order estimate (p_n, q_n, r_n) for (p_0, q_0, r_0) given by (19) is strongly consistent:

$$(p_n, q_n, r_n) \xrightarrow[n\to\infty]{} (p_0, q_0, r_0), \quad \text{a.s.}$$

We first prove three lemmas. Define

$$\xi_n = \hat{w}_n - w_n = y_n - w_n - \bar{\theta}_n^\tau \bar{\varphi}_{n-1}, \tag{20}$$

where $\hat{w}_n$ is defined by (10).

Lemma 1. If conditions (H_1)–(H_4) hold then

$$\sum_{i=0}^{n-1} \|\xi_{i+1}\|^2 = O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{c\delta(\beta-2)}), \quad \text{a.s.} \quad \forall c > 1,$$

where $\bar{s}_n^0$ is defined by (8).

Proof. The estimation established in the proof of Theorem 1 of Chen and Guo (1986a) holds true, and from (29) and (30) of that paper it follows that

$$\sum_{i=0}^{n-1} \|\xi_{i+1}\|^2 = O((\log \bar{s}_n)(\log\log \bar{s}_n)^{c\delta(\beta-2)}), \quad \text{a.s.} \quad \forall c > 1, \tag{21}$$

where $\bar{s}_n$ is defined by

$$\bar{s}_n = 1 + \sum_{i=0}^{n-1} \|\bar{\varphi}_i\|^2$$

with $\bar{\varphi}_i$ given by (9).

Using (9), (7), (8), (20) and (21) it is easy to show that

$$\bar{s}_n = O(\bar{s}_n^0), \quad \text{a.s.} \tag{22}$$

From this and (21) the desired result follows. □

Lemma 2. Let Assumption (H_1) be satisfied, and let the random vector φ_n be $\mathscr{F}_n$-measurable, $\forall n$. Then as $n \to \infty$,

$$\left\|\left(\sum_{i=0}^{n-1} \varphi_i\varphi_i^\tau + \varepsilon \mathrm{I}\right)^{-1/2} \sum_{i=0}^{n-1} \varphi_i w_{i+1}^\tau\right\|^2 = O((\log s_n)(\log\log s_n)^{c\delta(\beta-2)}), \quad \text{a.s.} \quad \forall c > 1,$$

where $\varepsilon > 0$, and

$$s_n \triangleq 1 + \sum_{i=0}^{n-1} \|\varphi_i\|^2.$$

Proof. Set

$$S_n = \sum_{i=0}^{n-1} \varphi_i w_{i+1}^\tau, \quad R_n = \left(\sum_{i=0}^{n-1} \varphi_i\varphi_i^\tau + \varepsilon \mathrm{I}\right)^{-1}. \tag{23}$$

By the matrix inversion formula, it is clear that

$$R_{n+1} = R_n - b_n R_n \varphi_n \varphi_n^\tau R_n, \quad b_n = (1 + \varphi_n^\tau R_n \varphi_n)^{-1}.$$

Hence

$$\begin{aligned}\operatorname{tr} S_{n+1}^\tau R_{n+1} S_{n+1} &= \operatorname{tr} S_1^\tau R_1 S_1 + \sum_{i=1}^{n} (2b_i w_{i+1}^\tau S_i^\tau R_i \varphi_i + b_i \varphi_i^\tau R_i \varphi_i \|w_{i+1}\|^2\\ &\quad - b_i \|S_i^\tau R_i \varphi_i\|^2).\end{aligned} \tag{24}$$

By the estimate for the martingale difference sequence (Chen and Guo, 1986a; Lai and Wei, 1982) we know that

$$\operatorname{tr} S_{n+1}^\tau R_{n+1} S_{n+1} = O(1) + \sum_{i=0}^{n} b_i \varphi_i^\tau R_i \varphi_i \|w_{i+1}\|^2. \tag{25}$$

But in (29) and (30) of Chen and Guo (1986a) we have shown that

$$\sum_{i=0}^{n} b_i \varphi_i^\tau R_i \varphi_i \|w_{i+1}\|^2 = O((\log s_{n+1})(\log\log s_{n+1})^{c\sigma(\beta-2)}), \quad \text{a.s.} \quad \forall c > 1. \tag{26}$$

Thus, combining (25) with (26) we conclude that

$$\|R_{n+1}^{1/2} S_{n+1}\|^2 \leqslant \operatorname{tr} S_{n+1}^\tau R_{n+1} S_{n+1} = O((\log s_{n+1})(\log\log s_{n+1})^{c\delta(\beta-2)}), \quad \text{a.s.} \quad \forall c > 1.$$

This proves the lemma. □

Lemma 3. Under conditions (H_1)–(H_5), CIC$(p, q, r)_n$ defined by (17) satisfies

$$\mathrm{CIC}(p, q, r)_n - \mathrm{CIC}(p_0, q_0, r_0)_n \geqslant \begin{cases} a_n(p + q + r - p_0 - q_0 - r_0 + o(1)), \text{ a.s.} & \\ \quad \text{if } (s, t, \lambda) = (p, q, r), & (27)\\ \lambda_{\min}^{(s,t)}(n)(\frac{1}{4}\alpha_0 + o(1)), \text{ a.s.} & \\ \quad \text{if } (s, t, \lambda) \neq (p, q, r), & (28)\end{cases}$$

for any $(p, q, r) \in M$, where $\alpha_0 = \min\{\|A_{p_0}\|^2, \|B_{q_0}\|^2, \|C_{r_0}\|^2\} > 0$,

$$(s, t, \lambda) \triangleq (p \vee p_0, q \vee q_0, r \vee r_0), \tag{29}$$

and $a \vee b$ denotes $\max(a, b)$.

Proof. By Lemma 1 we have

$$\sum_{i=0}^{n-1} \|\varphi_i(p, q, r) - \varphi_i^0(p, q, r)\|^2 = O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{c\delta(\beta-2)}), \quad \text{a.s.} \quad \forall c > 1, \tag{30}$$

for any $(p, q, r) \in M$, where $\varphi_i(p, q, r)$ is defined by (16). By (18) we have for $(s, t, \lambda) = (p, q, r)$

$$\begin{aligned}\sigma_n(p, q, r) &= \operatorname{tr} \tilde{\theta}_n^\tau(p, q, r) \sum_{i=0}^{n-1} \varphi_i(p, q, r) \times \varphi_i^\tau(p, q, r)\tilde{\theta}_n(p, q, r)\\ &\quad + 2 \operatorname{tr} \tilde{\theta}_n^\tau(p, q, r) \sum_{i=0}^{n-1} \varphi_i(p, q, r) \times [-\theta^\tau(p, q, r)\varphi_i^\xi(p, q, r) + w_{i+1}]^\tau\\ &\quad + \sum_{i=0}^{n-1} \|\theta^\tau(p, q, r)\varphi_i^\xi(p, q, r) - w_{i+1}\|^2, \end{aligned}\tag{31}$$

where

$$\tilde{\theta}_n(p, q, r) = \theta(p, q, r) - \theta_n(p, q, r),$$
$$\varphi_n^\xi(p, q, r) = \varphi_n(p, q, r) - \varphi_n^0(p, q, r).$$

By (30) and Schwarz inequality it follows that for any $(p, q, r) \in M$,

$$\begin{aligned}&\left\| \left(\sum_{j=0}^{n-1} \varphi_j(p, q, r)\varphi_j^\tau(p, q, r) + \frac{1}{d}I \right)^{-1/2} \times \sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi_i^{\xi\tau}(p, q, r) \right\|^2\\ &\leq \operatorname{tr}\left[\sum_{i=0}^{n-1} \left(\sum_{j=0}^{n-1} \varphi_j(p, q, r)_j^\tau(p, q, r) + \frac{1}{d}I \right)^{-1} \times \varphi_i(p, q, r)\varphi_i^\tau(p, q, r) \right] \sum_{i=0}^{n-1} \|\varphi_i^\xi(p, q, r)\|^2\\ &= O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{c\delta(\beta-2)}), \quad \forall c > 1, \text{ a.s.}\end{aligned}\tag{32}$$

Further, by Lemma 1 and an estimation for martingales (see e.g. Chen and Guo (1986a)) it is easy to see that

$$\sum_{i=0}^{n-1} \|\theta^\tau(p, q, r)\varphi_i^\xi(p, q, r) - w_{i+1}\|^2 = \sum_{i=0}^{n-1} \|w_{i+1}\|^2 + O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{c\delta(\beta-2)}). \tag{33}$$

Hence, by Lemma 2 and by using (22), (31)–(33), it is not difficult to conclude that for all $(s, t, \lambda) = (p, q, r)$

$$\sigma_n(p, q, r) = O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{c\delta(\beta-2)}) + \sum_{i=0}^{n-1} \|w_{i+1}\|^2, \quad \forall c > 1. \tag{34}$$

From this and (17) we get

$$\mathrm{CIC}(p, q, r)_n - \mathrm{CIC}(p_0, q_0, r_0)_n = a_n\left(p + q + r - p_0 - q_0 - r_0 + O\left(\frac{(\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{c\delta(\beta-2)}}{a_n}\right)\right), \quad \text{a.s.} \tag{35}$$

for any $c > 1$ and $(s, t, \lambda) = (p, q, r)$.

Taking $c \in (1, \alpha)$ and using Assumption (H_5), we obtain (27) from (35).

Now we proceed to prove (28).

For any fixed $(p, q, r) \in M$, set as

$$\begin{aligned}\tilde{\theta}_n'(p, q, r) = [&A_1 - A_{1n}' \dots A_s - A_{sn}'\\ &B_1 - B_{1n}' \dots B_t - B_{tn}'\\ &C_1 - C_{1n}' \dots C_\lambda - C_{\lambda n}']^\tau,\end{aligned}\tag{36}$$

where A_i, B_j, C_k, $l \leq i \leq s$, $l \leq j \leq t$, $l \leq k \leq \lambda$ are defined by (2a)–(2c) and (13), and

$$A_{in}' = \begin{cases} A_{in}, & i \leq p \\ 0, & i > p \end{cases}; \quad B_{jn}' = \begin{cases} B_{jn}, & j \leq q \\ 0, & j > q \end{cases}; \quad C_{kn}' = \begin{cases} C_{kn}, & k \leq r \\ 0, & k > r \end{cases}. \tag{37}$$

By a similar method to (31), for any $(p, q, r) \in M$, we have

$$\begin{aligned}\sigma_n(p, q, r) &= \operatorname{tr} \tilde{\theta}_n'^\tau(p, q, r) \sum_{i=0}^{n-1} \varphi_i(s, t, \lambda) \times \varphi_i^\tau(s, t, \lambda)\tilde{\theta}_n'(p, q, r)\\ &\quad + 2 \operatorname{tr} \tilde{\theta}_n'^\tau(p, q, r) \sum_{i=0}^{n-1} \varphi_i(s, t, \lambda) \times (-\theta^\tau(s, t, \lambda)\varphi_i^\xi(s, t, \lambda) + w_{i+1})^\tau\\ &\quad + \sum_{i=0}^{n-1} \|w_{i+1} - \theta^\tau(s, t, \lambda)\varphi_i^\xi(s, t, \lambda)\|^2.\end{aligned}\tag{38}$$

By (6) and (7), in the case where $(s, t, \lambda) \neq (p, q, r)$, we have

$$\lambda_{\min}^{(s,t,\lambda)}(n) \geq \min\{\lambda_{\min}^{(p_0,q^*,r^*)}(n), \lambda_{\min}^{(p^*,q_0,r^*)}(n), \lambda_{\min}^{(p^*,q^*,r_0)}(n)\}, \tag{39}$$

which tends to infinity as $n \to \infty$ by Assumption (H_5).

By an argument completely similar to that used in the proof of Theorem 2 of Chen and Guo (1986a) it can be shown that for sufficiently large n

$$\lambda_{\min}\left(\sum_{i=0}^{n-1} \varphi_i(s, t, \lambda)\varphi_i^\tau(s, t, \lambda) \right) \geq \tfrac{1}{3}\lambda_{\min}\left(\sum_{i=0}^{n-1} \varphi_i^0(s, t, \lambda)\varphi_i^{0\tau}(s, t) \right) \geq \tfrac{1}{4}\lambda_{\min}^{(s,t,\lambda)}(n), \tag{40}$$

for $(s, t, \lambda) \neq (p, q, r)$.

Also, in the case where $(s, t, \lambda) \neq (p, q, r)$, it can be seen from (36) that

$$\|\tilde{\theta}_n'(p, q, r)\|^2 \geq \min\{\|A_{p_0}\|^2, \|B_{q_0}\|^2, \|C_{r_0}\|^2\} = \alpha_0 > 0. \tag{41}$$

Hence, by (40) and (41), for the first term on the right-hand side of (38), we have

$$\operatorname{tr} \tilde{\theta}_n'^\tau(p, q, r) \sum_{i=0}^{n-1} \varphi_i(s, t, \lambda)\varphi_i^\tau(s, t, \lambda)\tilde{\theta}_n'(p, q, r) \geq \tfrac{1}{4}\alpha_0\lambda_{\min}^{(s,t,\lambda)}(n). \tag{42}$$

By Lemma 2 and (32), we estimate the second term on the right-hand side of (38) as follows

$$\begin{aligned}&2\left| \operatorname{tr} \tilde{\theta}_n'^\tau(p, q, r) \sum_{i=0}^{n-1} \varphi_i(s, t, \lambda) \times (-\theta^\tau(s, t, \lambda)\varphi_i^\xi(s, t, \lambda) + w_{i+1})^\tau \right|\\ &\leq \left\| \tilde{\theta}_n'^\tau(p, q, r)\left(\sum_{i=0}^{n-1} \varphi_i(s, t, \lambda)\varphi_i^\tau(s, t, \lambda) \right)^{1/2} \right\| \cdot O(\sqrt{(\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}}).\end{aligned}\tag{43}$$

By a similar consideration to (33), we have

$$\sum_{i=0}^{n-1} \|w_{i+1} - \theta^\tau(s, t, \lambda)\varphi_i^\xi(s, t, \lambda)\|^2 = \sum_{i=0}^{n} \|w_{i+1}\|^2 + O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}). \tag{44}$$

Combining (38), (42), (43) and (44) it follows that

$$\begin{aligned}&\sigma_n(p, q, r)\\ &\geq \frac{\alpha_0}{4}\lambda_{\min}^{(s,t,\lambda)}(n)\left(1 + O\left(\left(\frac{(\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}}{\lambda_{\min}^{(s,t,\lambda)}(n)}\right)^{1/2}\right)\right)\\ &\quad + \sum_{i=0}^{n-1} \|w_{i+1}\|^2,\end{aligned}\tag{45}$$

for $(s, t, \lambda) \neq (p, q, r)$.

On the other hand, from (34), we have

$$\sigma_n(p_0, q_0, r_0) = O((\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}) + \sum_{i=0}^{n-1} \|w_{i+1}\|^2,$$

which, together with (45) and (17), implies

$$\begin{aligned}
\mathrm{CIC}(p, q, r)_n &- \mathrm{CIC}(p_0, q_0, r_0)_n \\
&\geq \lambda_{\min}^{(s,t,\lambda)}(n)\Bigg(\frac{\alpha_0}{4} + O\Bigg(\Bigg(\frac{(\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}}{\lambda_{\min}^{(s,t,\lambda)}(n)}\Bigg)^{1/2}\Bigg) \\
&+ O\Bigg(\frac{(\log \bar{s}_n^0)(\log\log \bar{s}_n^0)^{\alpha\delta(\beta-2)}}{\lambda_{\min}^{(s,t,\lambda)}(n)}\Bigg) \\
&+ O\Bigg(\frac{a_n}{\lambda_{\min}^{(s,t,\lambda)}(n)}\Bigg)\Bigg), \quad \text{for } (s, t, \lambda) \neq (p, q, r). \qquad (46)
\end{aligned}$$

Then, by (3), (4) and (39), the desired result (28) follows from (46). This completes the proof of Lemma 3.

Proof of Theorem 1. We need only to show that any limit point of (p_n, q_n, r_n) coincides with (p_0, q_0, r_0). Let $(p', q', r') \in M$ be the limit of a subsequence $(p_{n_k}, q_{n_k}, r_{n_k})$. It suffices to prove the impossibility of the following situations:

$$(1) \quad p' < p_0 \quad \text{or} \quad q' < q_0 \quad \text{or} \quad r' < r_0; \qquad (47)$$

$$(2) \quad p' + q' + r' > p_0 + q_0 + r_0. \qquad (48)$$

If (47) was true, then by the definition (19), (28) with $(p, q, r) = (p', q', r')$, (39) and (H_5), we could see that for all sufficiently large k,

$$\begin{aligned}
0 &\geq \mathrm{CIC}(p_{n_k}, q_{n_k}, r_{n_k})_{n_k} - \mathrm{CIC}(p_0, q_0, r_0)_{n_k} \\
&= \mathrm{CIC}(p', q', r')_{n_k} - \mathrm{CIC}(p_0, q_0, r_0)_{n_k} \\
&\geq \lambda_{\min}^{(s,t,\lambda)}(n_k)\Big(\frac{\alpha_0}{4} + o(1)\Big) \xrightarrow[n\to\infty]{} \infty. \qquad (49)
\end{aligned}$$

But this is impossible.

If (48) was true, then by (19), (27), (H_5) and the impossibility of (47), for a sufficiently large k we would have

$$\begin{aligned}
0 &\geq \mathrm{CIC}(p_{n_k}, q_{n_k}, r_{n_k})_{n_k} - \mathrm{CIC}(p_0, q_0, r_0)_{n_k} \\
&= \mathrm{CIC}(p', q', r')_{n_k} - \mathrm{CIC}(p_0, q_0, r_0)_{n_k} \\
&\geq a_{n_k}(p' + q' + r' - p_0 - q_0 - r_0 + o(1)) \xrightarrow[k\to\infty]{} \infty.
\end{aligned}$$

But this also is impossible.

Thus the proof has been completed.

4. *Application to adaptive control systems*

In this section we specify the sequence $\{a_n\}$ used in (17) and show that the selected $\{a_n\}$ satisfies Assumption (H_5) for a large class of important adaptive control systems.

In adaptive control, the attenuating excitation technique is very successful in getting the minimality of control performance and consistency of parameter estimate simultaneously (Chen and Guo, 1986a, b). We shall describe this method.

Let $\{v_n\}$ be a sequence of l-dimensional mutually independent random vectors independent of $\{w_n\}$ with properties,

$$Ew_n = 0, \quad Ev_nv_n^\tau = \frac{1}{n^\varepsilon}\mathrm{I}, \quad \|v_n\|^2 \leq \frac{\sigma^2}{n^\varepsilon}, \quad \varepsilon \in \Big[0, \frac{1}{2(t+1)}\Big) \qquad (50)$$

where $t \triangleq (m+1)p^* + q^* + r^* - 1$ and $\sigma^2 > 0$ is a constant.

Without loss of generality, we assume that

$$\mathscr{F}_n = \sigma\{w_i, v_i, i \leq n\}, \quad \mathscr{F}'_n = \sigma\{w_i, v_{i-1}, i \leq n\}.$$

Let u_n^0 be an l-dimensional $\mathscr{F}'_n$-measurable desired control. Obviously, any feedback (adaptive) control is of this kind. The attenuating excitation technique suggests that one takes the actual control for the system as

$$u_n = u_n^0 + v_n. \qquad (51)$$

The control defined by (51) is termed "attenuating excitation control" (Chen and Guo, 1986b).

We need the following assumptions.

(H_6). There is a positive definite matrix R, such that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} w_iw_i^\tau = R > 0, \quad \text{a.s.}$$

(H_7). $A(z)$, $B(z)$ and $C(z)$ have no common left factors and A_{p_0}, B_{q_0} and C_{r_0} are of row full rank.

Theorem 2. Suppose that the "attenuating excitation control" (51) is applied to system (1) and that Assumptions (H_1), (H_3), (H_4), (H_6) and (H_7) are satisfied. If there is a non-negative number δ, $\delta \in [0, (1 - 2\varepsilon(t+1))/(2t+3)]$ such that

$$\frac{1}{n}\sum_{i=0}^{n} (\|y_i\|^2 + \|u_i^0\|^2) = O(n^\delta), \quad \text{a.s.}, \qquad (52)$$

then

$$(p_n, q_n, r_n) \xrightarrow[n\to\infty]{} (p_0, q_0, r_0), \quad \text{a.s.},$$

where (p_n, q_n, r_n) is given by (17)–(19) with a_n being any sequence satisfying

$$\frac{(\log n)(\log\log n)^{\alpha\delta(\beta-2)}}{a_n} \xrightarrow[n\to\infty]{} 0, \quad \text{for some } a > 1, \qquad (53)$$

and

$$\frac{a_n}{n^{1-(t+1)(\varepsilon+\delta)}} \xrightarrow[n\to\infty]{} 0, \qquad (54)$$

where ε and t appear in (50). In addition, assume that $A(z)$ is stable. Then, condition (52) can be weakened as

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

Proof. Noticing that $1 - (t+1)(\varepsilon + \delta) > 0$, we know that there exists a sequence $\{a_n\}$ satisfying (53), (54). By Assumption (H_6) and (52), it follows that

$$\bar{s}_n^0 = O(n^{1+\delta}).$$

Hence by (53) we see that (3) is satisfied. Consequently, by Theorem 1 and (54) for proving Theorem 2 it suffices to show that

$$\liminf_{n\to\infty} n^{-\alpha_1}\lambda_{\min}^{(p,q,r)}(n) \neq 0, \quad \text{a.s.}, \qquad (55)$$

for any $(p, q, r) \in M^*$, where $\alpha_1 = 1 - (t+1)(\varepsilon + \delta)$.

If (55) was not true, then along the lines of the argument used in Chen and Guo (1986a, 1987a), we could find a $(p+r)m + lq$-dimensional vector η satisfying $\|\eta\| = 1$ and

$$\eta = (\alpha^{(0)\tau} \ldots \alpha^{(p-1)\tau} \quad \beta^{(0)\tau} \ldots \beta^{(q-1)\tau}\gamma^{(0)\tau} \ldots \gamma^{(r-1)\tau})^\tau$$

$$\begin{aligned}
0 = &\sum_{i=0}^{p-1} \alpha^{(i)\tau}z^i(\mathrm{Adj}\, A(z))[B(z), C(z)] \\
&+ \sum_{i=0}^{q-1} \beta^{(i)\tau}z^i[\det A(z)\mathrm{I}, 0] \\
&+ \sum_{i=0}^{r-1} \gamma^{(i)\tau}z^i[0, (\det A(z))\mathrm{I}_m]. \qquad (56)
\end{aligned}$$

We now show that

$$\deg\Big(\sum_{i=0}^{p-1} \alpha^{(i)\tau}z^i\Big) \leq p_0 - 1. \qquad (57)$$

If $(p, q, r) = (p_0, q^*, r^*)$ then (57) is trivial.

If $(p, q, r) = (p^*, q_0, r^*)$, then by (56) and (H_7) we have

$$\begin{aligned}
\deg&\Big(\sum_{i=0}^{p-1} \alpha^{(i)\tau}z^i\Big) + (m-1)p_0 + q_0 \\
&= \deg\Big(\sum_{i=0}^{q_0-1} \beta^{(i)\tau}z^i(\det A(z))\Big) \leq q_0 - 1 + mp_0,
\end{aligned}$$

which is tantamount to (57). Similarly, (57) can be verified in the case where $(p, q, r) = (p^*, q^*, r_0)$. Thus (57) has been proved.

By Assumption (H_7) there exist matrix polynomials $M(z)$, $N(z)$ and $L(z)$ such that

$$A(z)M(z)+[B(z),\,C(z)][N(z)^{\tau},\,L(z)^{\tau}]^{\tau}=I.$$

From this and (56) it follows that

$$\begin{aligned}\sum_{i=0}^{p-1}\alpha^{(i)\tau}z^{i}\operatorname{Adj}A(z)&=\sum_{i=0}^{p-1}\alpha^{(i)\tau}z^{i}(\operatorname{Adj}A(z))(A(z)M(z)\\&\quad+[B(z),\,C(z)][N(z)^{\tau},\,L(z)^{\tau}]^{\tau})\\&=(\det A(z))\Bigg(\sum_{i=0}^{p-1}\alpha^{(i)\tau}z^{i}M(z)\\&\quad-\sum_{i=0}^{q-1}\beta^{(i)\tau}z^{i}N(z)-\sum_{i=0}^{r-1}\gamma^{(i)\tau}z^{i}L(z)\Bigg).\end{aligned}\tag{58}$$

By (57) we have

$$\deg\left(\sum_{i=0}^{p-1}\alpha^{(i)\tau}z^{i}\operatorname{Adj}A(z)\right)$$
$$\leq p_0-1+(m-1)p_0=mp_0-1<mp_0=\deg(\det A(z)).\tag{59}$$

Consequently, from (58) and (59) we conclude that $\alpha^{(i)}=0$, $0\leq i\leq p-1$. Then by (56) it follows that $\beta^{(i)}=0$, $0\leq i\leq q-1$, $\gamma^{(j)}=0$, $0\leq j\leq r-1$. Therefore $\eta=0$, and this is impossible.

5. *Conclusion*

Up until our most recent paper (Chen and Guo, 1987b) the order estimation problem was solved exclusively for the ARMA model containing no control term. This paper gives a consistent estimate for orders of the feedback control systems with correlated noise, while in Chen and Guo (1987b) only the uncorrelated noise case is considered. Applying results obtained to the adaptive control system leads to consistent estimates for both orders and unknown coefficients of the system. The estimate presented in the paper is nonrecursive and requires availability of upper bounds for unknown orders. This probably requires further research.

Acknowledgements—The project was supported by National Natural Science Foundation of China and by the TWAS Research Grant No. 87-43.

References

Akaike, H. (1969). Fitting autoregressive models for prediction. *Ann. Inst. Statist. Math.*, **21,** 243–247.

Akaike, H. (1977). On entropy maximisation principle. In P. R. Krishaiah (Ed.), *Applications of Statistics*, pp. 27–41. North-Holland, Amsterdam.

Chen, H. F. and L. Guo (1986a). Convergence rate of least-squares identification and adaptive control for stochastic systems. *Int. J. Control,* **44,** 1459–1476.

Chen, H. F. and L. Guo (1986b). Strong consistency of parameter estimates in optimal stochastic adaptive tracking systems. *Scientia Sinica (Series A)*, **XXIX,** 1145–1156.

Chen, H. F. and L. Guo (1987a). Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost. *SIAM J. Control Opt.*, **25,** 845–867.

Chen, H. F. and L. Guo (1987b). Consistent estimation of the order of stochastic control systems. *IEEE Trans. Aut. Control,* **AC-32,** 531–535.

Hannan, E. J. and L. Kavalieris (1984). A method for autoregressive moving average estimation. *Biometrika,* **72,** 273–280.

Hannan, E. J. and B. G. Quinn (1979). The determination of the order of an autoregression. *J. R. Statist. Soc. B.*, **41,** 190–195.

Hannan, E. J. and J. Rissanen (1982). Recursive estimation of mixed autoregressive moving average order. *Biometrika,* **69,** 81–94.

Lai, T. L. and C. Z. Wei (1982). Least squares estimation in stochastic regression models with application to identification and control of dynamic systems. *Annals Statist.*, **10,** 154–166.