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 \((p_{0}, r_{0})\) for unknown orders \((p_{0}, r_{0})\) are usually given by minimizing some criterion, for example, AIC \((p, r)\) (Akaike, P. C., 1974), BIC \((p, r)\) (Akaike, 1977), and \(\Phi I C(p, r)\) (Tellman 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 \((p_{0}, q_{0}, r_{0})\) 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

\[
A(z)y_{n} = B(z)u_{n} + C(z)\omega_{n}, \quad n \geq 0;
\]

\[y_{n} = w_{n} = 0, \quad u_{n} = 0, \quad n < 0\]

(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 \(z\)

\[
A(z) = 1 + A_{1}z + \cdots + A_{p_{0}}z^{p_{0}}, \quad p_{0} \geq 0,\]

\[
B(z) = B_{1}z + \cdots + B_{q_{0}}z^{q_{0}}, \quad q_{0} \geq 0,\]

\[
C(z) = C_{1}z + \cdots + C_{r_{0}}z^{r_{0}}, \quad r_{0} \geq 0\]

(2a)

\[A(z) = 1 + A_{1}z + \cdots + A_{p_{0}}z^{p_{0}}, \quad p_{0} \geq 0,\]

\[B(z) = B_{1}z + \cdots + B_{q_{0}}z^{q_{0}}, \quad q_{0} \geq 0,\]

\[C(z) = C_{1}z + \cdots + C_{r_{0}}z^{r_{0}}, \quad r_{0} \geq 0\]

(2c)

with unknown orders \(p_{0}, q_{0}, r_{0}\) 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.

(H1). The driven noise is a martingale difference sequence with respect to a non-decreasing family of \(\sigma\)-algebras \((\mathcal{F}_{n})\) and such that

\[\sup_{n} E[\|w_{n+1}\|^{p} | \mathcal{F}_{n}] < \infty, \quad \text{a.s. for some } \beta \geq 1.\]

(H2). For any \(n \geq 1\), \(u_{n}\) is \(\mathcal{F}_{n}\)-measurable.

(H3). The transfer matrix \(\Phi^{-\theta}(s) - I\) is strictly positive real, i.e.

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

(H4). The true orders \((p_{0}, q_{0}, r_{0})\) belong to a known finite set \(M\):

\[M = \{(p, q, r) : 0 < p < p^{*}, 0 < q < q^{*}, 0 < r < r^{*}\}\]

(H5). A sequence of positive numbers \(\{a_{n}\}\) can be found such that

\[
\frac{\log \log \rho_{n}^{(p, q, r)}}{\log \log \log \rho_{n}^{(p, q, r)}} \rightarrow 0 \quad \text{as } \text{for some } \alpha > 1, \quad n \rightarrow \infty
\]

and

\[
\frac{\rho_{n}^{(p, q, r)}}{\lambda_{\min}(\rho_{n}^{(p, q, r)})} \rightarrow 0 \quad \text{a.s. } \forall (p, q, r) \in M^{*}
\]

where \(\delta_{0}\) is the Dirac function

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

and \(M^{*}\) denotes the set consisting of three points:

\[M^{*} = \{(p_{0}, q^{*}, r^{*}), (p^{*}, q_{0}, r^{*}), (p^{*}, q^{*}, r_{0})\}\]

(3)

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

\[\sum_{i=0}^{n-1} \rho_{i}^{(p, q, r)} \rho_{i}^{(p, q, r)} - 1 \quad d = mp^{*} + qt^{*} + mr^{*}, \quad (d = \rho_{n}^{(p, q, r)} - 1)
\]

with

\[
\rho_{i}^{(p, q, r)} = \delta_{0}(Y_{i}^{*}, \ldots, Y_{i-p-1}^{*}, u_{i-1}^{*}, \ldots, u_{i-q-1}^{*}, w_{i-q}^{*}, \ldots, w_{i-r}^{*})
\]

and

\[
\rho_{n}^{(p, q, r)} = \sum_{i=0}^{n-1} \|\rho_{i}^{(p, q, r)}\|^{2} + 1.
\]

System (1) under Assumptions (H1)–(H4) is, generally speaking, neither stationary nor ergodic because (i) the system input \(u_{n}\) may be an arbitrary \(\mathcal{F}_{n}\)-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
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 $(a_k)$ and show that Assumption (10) is indeed satisfied.

We introduce the regression vector

$$\hat{\theta}_n = [y_n, \ldots, y_{n-p+1}, u_n, \ldots, u_{n-q-1}, \hat{\Theta}_n, \ldots, \hat{\Theta}_{n-r+s+1}],$$

(9)

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

$$\hat{\Theta}_n = y_n - \hat{\Theta}_n \hat{\Phi}_{n-1}, \quad n = 0, \quad \hat{\Theta}_0 = 0, \quad n < 0,$$

(10)

$$\hat{\Theta}_{n+1} = \hat{\Theta}_n + \hat{\Phi}_n \hat{\Phi}_n \hat{\Phi}_n, \quad n = 1 + \hat{\Phi}_n \hat{\Phi}_n \hat{\Phi}_n,$$

(11)

with initial value $\hat{\Theta}_0$ arbitrarily chosen and $\hat{\Phi}_n = dI_n$, where $d$ is given in (6).

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

$$\theta(p, q, r) = [-A_1, \ldots, -A_p, B_1, \ldots, B_q, C_1, \ldots, C_r]^T,$$

(12)

where by definition

$$A_i = 0, \quad B_i = 0, \quad C_k = 0 \quad \text{for} \quad i > p, j > q, k > r.$$  

(13)

The extended least squares estimate

$$\hat{\theta}_n(p, q, r) = [-A_p, \ldots, -A_n, B_n, \ldots, B_{n-p}, C_n, \ldots, C_{n-r}]^T,$$

(14)

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

$$\hat{\theta}_n(p, q, r) = \left( \sum_{q=0}^{n-1} \varphi(q, p, r) \varphi(q, p, r) \right)^{-1} \left( \sum_{q=0}^{n-1} \varphi(q, p, r) y(q, p, r) \right),$$

(15)

where

$$\varphi(q, p, r) = [y_{n-q-1}, \ldots, y_{n-p-1}, u_{n-q-1}, \ldots, u_{n-r-1}, \hat{\Theta}_n, \ldots, \hat{\Theta}_{n-r+s}].$$

(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:

$$\text{CIC}(p, q, r)_n = a_n(p, q, r) + (p + q + r) a_n,$$

(17)

where the subscript $n$ denotes the data size, and where $a_n$ is given in Assumption (10) and $a_n(p, q, r)$ is a residual given by

$$a_n(p, q, r) = \sum_{r=0}^{n-1} \| x_{r+1} - \theta(p, q, r) \varphi(q, p, r) \|^2.$$

(18)

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

$$(p_n, q_n, r_n) = \arg \min_{(p, q, r) \in M} \text{CIC}(p, q, r)_n.$$  

(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:

$$(p_n, q_n, r_n) \rightarrow (p_0, q_0, r_0), \quad \text{a.s.}$$

(20)

We first prove three lemmas. Define

$$\delta_n = \delta_n - \delta_n = y_n - \delta_n - \delta_n - \delta_n,$$

(21)

where $\delta_n$ is defined in (10).

**Lemma 1.** If conditions (H_1)-(H_5) hold then

$$\sum_{r=0}^{n-1} \| \delta_{r+1} \|^2 = O((\log n)^{2}((\log \log n)^{2})^{2} - 2), \quad \text{a.s.} \ \forall c > 0,$$

(22)

where $\delta_n$ 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

$$\sum_{r=0}^{n-1} \| \delta_{r+1} \|^2 = O((\log n)^{2}((\log \log n)^{2})^{2} - 2), \quad \text{a.s.} \ \forall c > 0,$$

(23)

where $\delta_n$ is defined by

$$\delta_n = 1 + \sum_{r=0}^{n-1} \| \delta_{r+1} \|^2$$

(24)

with $\delta_n$ given by (9), (20) and (21) it is easy to show that

$$\delta_n = O(n^{2}), \quad \text{a.s.}$$

(25)

From this and (21) the desired result follows.

**Lemma 2.** Let Assumption (H_1) be satisfied and let the random vector $\varphi_n$ be $F_n$-measurable. Then as $n \rightarrow \infty$,

$$\left( \sum_{n=0}^{\infty} \| \varphi_n \|^2 + \varepsilon \right) \sum_{n=0}^{\infty} \| \varphi_n \|^2$$

(26)

$$\rightarrow O((\log \log n)^{2}((\log \log n)^{2})^{2} - 2), \quad \text{a.s.} \ \forall c > 0,$$

(27)

where $\varepsilon > 0$, and

$$\varepsilon = \frac{1}{\sum_{n=0}^{\infty} \| \varphi_n \|^2}$$

(28)

**Proof.** Set

$$S_n = \sum_{n=0}^{\infty} \varphi_{n} \varphi_{n+1}, \quad R_n = \left( \sum_{n=0}^{\infty} \varphi_{n} \varphi_{n+1} + \varepsilon \right)^{-1}$$

(29)

By the matrix inversion formula, it is clear that

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

(30)

Hence

$$\text{tr} S_{n}^T R_{n-1} S_{n} + \text{tr} S_{n}^T R_{n} S_{n}$$

$$= b_n \varphi_n \varphi_n^T R_n \varphi_n \left( \sum_{n=0}^{\infty} \| \varphi_n \|^2 + \varepsilon \right)^{-1}$$

(31)

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

$$\text{tr} S_{n}^T R_{n-1} S_{n-1} = O(1) + \sum_{n=0}^{\infty} b_n \varphi_n \varphi_n^T R_n \varphi_n \left( \sum_{n=0}^{\infty} \| \varphi_n \|^2 + \varepsilon \right)^{-1}$$

(32)

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

$$\sum_{n=0}^{\infty} b_n \varphi_n \varphi_n^T R_n \varphi_n \left( \sum_{n=0}^{\infty} \| \varphi_n \|^2 + \varepsilon \right)^{-1}$$

(33)

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

$$\| R_n \|^2 + \text{tr} S_{n}^T R_{n-1} S_{n}$$

$$= O((\log \log n)^{2}((\log \log n)^{2})^{2} - 2), \quad \text{a.s.} \ \forall c > 0.$$  

(34)

This proves the lemma.

**Lemma 3.** Under conditions (H_1)-(H_3), CIC($p, q, r$) defined by (17) satisfies

$$\text{CIC}(p, q, r)_n - \text{CIC}(p_0, q_0, r_0)_n$$

(35)

$$= \begin{cases} a_n(p + q + r - p_0 - q_0 - r_0 + o(1)), \quad \text{a.s.} \quad \text{if} \quad (s, t, \lambda) = (p, q, r), \\ a_n(o(1)), \quad \text{a.s.} \quad \text{if} \quad (s, t, \lambda) \neq (p, q, r) \end{cases}$$

(36)

$$= \begin{cases} a_n(p + q + r - p_0 - q_0 - r_0 + o(1)), \quad \text{a.s.} \quad \text{if} \quad (s, t, \lambda) = (p, q, r), \\ a_n(o(1)), \quad \text{a.s.} \quad \text{if} \quad (s, t, \lambda) \neq (p, q, r) \end{cases}$$

(37)
for any \((p, q, r) \in M\), where \(\alpha_0 = \min\{\|A_{ij}\|^2, \|B_{jk}\|^2, \|C_{kl}\|^2\} > 0\),
\[\Theta(p, q, r) = (p \vee p_0, q \vee q_0, r \vee r_0),\]  
and \(a \vee b\) denotes \(\max(a, b)\).

Proof. By Lemma 1 we have
\[
\sum_{i=0}^{n-1} \|\varphi_i(p, q, r) - \varphi^*_i(p, q, r)\|^2 = O(\log C_n^2\log \log C_n)^{\alpha_0(\beta - 2)}, \text{ a.s. } \forall \epsilon > 1, \tag{30}
\]
for any \((p, q, r) \in M\), where \(\varphi_i(p, q, r)\) is defined by (16). By (18) we have for \((s, t, \lambda) = (p, q, r)\),
\[
\sigma_{n}(p, q, r) = \sum_{i=0}^{n-1} \varphi_i(p, q, r) \times \varphi^*_i(p, q, r) + 2 \text{tr} \vartheta^*_i(p, q, r) \sum_{i=0}^{n-1} \varphi_i(p, q, r) \times \vartheta_i(p, q, r) + \frac{1}{d} \sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2, \tag{31}
\]
where \(\vartheta_i(p, q, r) = \varphi_i(p, q, r) - \Theta(p, q, r) - \vartheta_i(p, q, r)\).

By (30) and Schwarz inequality it follows that for any \((p, q, r) \in M\),
\[
\left\|\sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi^*_i(p, q, r) + \frac{1}{d} \sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2 \right\|^{\frac{1}{2}} \leq \left\|\sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi^*_i(p, q, r) + \frac{1}{d} \sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2 \right\|^{\frac{1}{2}}, \tag{32}
\]
Further, by Lemma 1 and an estimation for martingales (see e.g. Chen and Guo (1986a)) it is easy to see that
\[
\sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2 = \sum_{i=0}^{n-1} \|w_i\|^2 + O(\log C_n^2\log \log C_n)^{\alpha_0(\beta - 2)}. \tag{33}
\]
Hence, by Lemma 2 and by using (22), (31)–(33), it is not difficult to conclude that for all \((s, t, \lambda) = (p, q, r)\),
\[
\sigma_{n}(p, q, r) = O\left(\left(\log C_n^2\log \log C_n^{\alpha_0(\beta - 2)}\right)^{\frac{1}{4}}\right), \tag{34}
\]
From this and (17) we get
\[
CIC(p, q, r), \text{CIC}(p_0, q_0, r_0) = \alpha_0(p + q + r - p_0 - q_0 - r_0) + O\left(\frac{\log C_n^2}{\log \log C_n^{\alpha_0(\beta - 2)}\log \log C_n^{\alpha_0(\beta - 2)}}\right), \tag{35}
\]
for any \(c > 1\) and \((s, t, \lambda) = (p, q, r)\).

Taking \(c \geq 1\) and using Assumption (H3), we obtain (27) from (35).

Now we proceed to prove (28).

For any fixed \((p, q, r) \in M\) set
\[
\hat{\Theta}_n(p, q, r) = [A_1 - A^*_n, \ldots, A_m - A_m^*, B_1 - B_1^*, \ldots, B_m - B_m^*, C_1 - C_1^*, \ldots, C_m - C_m^*], \tag{36}
\]
where \(A_i, B_i, C_i, i \leq i \leq p, i \leq j \leq t, i \leq k \leq \lambda\) are defined by (28a)–(28c) and (13), and
\[
A^*_n = \left\{ \begin{array}{ll} A_m, & i > p, \\ A_m - A^*_n, & i = \hat{p}, \\ 0, & i > p, \\ i < q, & i = q, \\ C_m, & k > r, \\ 0, & k < r. \\ \end{array} \right. \tag{37}
\]

By a similar method to (31), for any \((p, q, r) \in M\), we have
\[
\sigma_{n}(p, q, r) = \text{tr} \hat{\Theta}_n(p, q, r) \sum_{i=0}^{n-1} \varphi_i(p, q, r) \times \varphi^*_i(p, q, r) + 2 \text{tr} \hat{\vartheta}_n(p, q, r) \sum_{i=0}^{n-1} \varphi_i(p, q, r) \times \vartheta_i(p, q, r) + \frac{1}{d} \sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2, \tag{38}
\]
where \(\vartheta_i(p, q, r) = \varphi_i(p, q, r) - \Theta(p, q, r) - \vartheta_i(p, q, r)\).

By (6) and (7), in the case where \((s, t, \lambda) \neq (p, q, r)\), we have
\[
\lambda_{\min}(\sum_{i=0}^{n-1} \varphi_i(s, t, \lambda)\varphi^*_i(s, t, \lambda)) \geq \lambda_{\min}(\sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi^*_i(p, q, r)), \tag{40}
\]
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
\[
\|\hat{\Theta}_n(p, q, r)\| \geq \min(\|A_m\|^2, \|B_m\|^2, \|C_m\|^2) = \alpha_0 > 0. \tag{41}
\]

Hence, by (40) and (41), for the first term on the right-hand side of (38), we have
\[
\text{tr} \hat{\Theta}_n(p, q, r) \sum_{i=0}^{n-1} \varphi_i(s, t, \lambda)\varphi^*_i(s, t, \lambda) \geq \lambda_{\min}(\sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi^*_i(p, q, r)), \tag{42}
\]
By Lemma 2 and (32), we estimate the second term on the right-hand side of (38) as follows
\[
2 \text{tr} \hat{\vartheta}_n(p, q, r) \sum_{i=0}^{n-1} \varphi_i(s, t, \lambda) \vartheta_i(p, q, r) + \frac{1}{d} \sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2 \leq O(\log C_n^2\log \log C_n^{\alpha_0(\beta - 2)}), \tag{43}
\]
By a similar consideration to (33), we have
\[
\sum_{i=0}^{n-1} \|\vartheta_i(p, q, r)\|^2 \geq \sum_{i=0}^{n-1} \|w_i\|^2 + O(\log C_n^2\log \log C_n^{\alpha_0(\beta - 2)}). \tag{44}
\]
Combining (38), (42), (43) and (44) it follows that
\[
\alpha_0(p, q, r) \geq \frac{\alpha_0}{4} \lambda_{\min}(\sum_{i=0}^{n-1} \varphi_i(p, q, r)\varphi^*_i(p, q, r))^{1/2} \tag{50}
\]
for \((s, t, \lambda) \neq (p, q, r)\).
On the other hand, from (34), we have
\[ a_{n}(p_{0}, q_{0}, r_{0}) = O((\log p_{0}^{2})(\log \log p_{0}^{2})^{\alpha n^{1/2}}) + \sum_{i=0}^{n-1} \|w_{i}\|^{2}, \]
which, together with (45) and (17), implies
\[ \text{CIC}(p, q, r_{n}) - \text{CIC}(p_{0}, q_{0}, r_{0})_{n} \]
\[ \geq \lambda_{\min}^{(2,1)} \left( \frac{2}{4} \right) \left( \frac{\left( \text{log} \log p_{0}^{2} \right)^{\alpha n^{1/2}}}{\lambda_{\min}^{(2,1)}(p_{0})} \right)^{1/2} \]
\[ + O\left( \left( \frac{\text{log} \log p_{0}^{2}}{\lambda_{\min}^{(2,1)}(p_{0})} \right)^{\alpha n^{1/2}} \right) \]
\[ + O\left( \frac{a_{n}}{\lambda_{\min}^{(2,1)}(p_{0})} \right), \] for \( (s, t, \lambda) \neq (p, q, r). \) (46)

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 \((p_{n}, q_{n}, r_{n})\).

If it suffices to prove the impossibility of the following situations:

(1) \( p' < p_{0} \) or \( q' < q_{0} \) or \( r' < r_{0} \);

(2) \( p' + q' + r' > p_{0} + q_{0} + r_{0} \).

If (47) was true, then by the definition (19), (28) with \((p, q, r) = (p', q', r')\), (39) and (H2), we could see that for all sufficiently large \( k \),

\[ 0 = \text{CIC}(p_{n}, q_{n}, r_{n})_{n} - \text{CIC}(p_{0}, q_{0}, r_{0})_{n} \]
\[ = \text{CIC}(p', q', r')_{n} - \text{CIC}(p_{0}, q_{0}, r_{0})_{n} \]
\[ \geq \lambda_{\min}^{(2,1)}(r_{n}) \left( \frac{2}{4} + o(1) \right) \rightarrow_{n \rightarrow \infty} 0. \] (49)

But this is impossible.

If (48) was true, then by (19), (27), (H2) and the impossibility of (47), for a sufficiently large \( k \) we would have

\[ 0 = \text{CIC}(p_{n}, q_{n}, r_{n})_{n} - \text{CIC}(p_{0}, q_{0}, r_{0})_{n} \]
\[ = \text{CIC}(p', q', r')_{n} - \text{CIC}(p_{0}, q_{0}, r_{0})_{n} \]
\[ \geq a_{n}(p' + q' + r' - p_{0} - q_{0} - r_{0} + o(1)) \rightarrow_{n \rightarrow \infty} 0. \]

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 (H2) 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 variables independent of \((w_{n})\) with properties,

\[ Ew_{n} = 0, \hskip0.5cm Ew_{n}v_{n}^{*} = \frac{1}{n} I, \hskip0.5cm \|v_{n}\|^{2} = \frac{2}{n}, \hskip0.5cm \varepsilon \in [0, \frac{1}{2(l+1)}) \] (50)

where \( l \delta \leq (m+1)p + q + r - 1 \) and \( \sigma^{2} \geq 0 \) is a constant.

Without loss of generality, assume that

\[ \mathcal{F}_{n} = \sigma(w_{i}, v_{i}, i \leq n), \hskip0.5cm \mathcal{F}_{n}^{*} = \sigma(w_{i}, v_{i-1}, i \leq n). \]

Let \( u_{n}^{*} \) be an \( l \)-dimensional \( \mathcal{F}_{n}^{*} \)-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

\[ u_{n} = u_{n}^{*} + v_{n}. \] (51)

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

We need the following assumptions.

(H2). There is a positive definite matrix \( R \), such that

\[ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n} \|w_{i}\|^{2} = R > 0, \hskip0.5cm \text{a.s.} \] (H2). \( A(z), B(z) \) and \( C(z) \) have no common left factors and \( A_{nr}, B_{nr}, C_{nr} \) are of row full rank.

**Theorem 2.** Suppose that the “attenuating excitation control” (51) is applied to system (1) and that Assumptions (H1), (H2), (H3), (H4) are satisfied. If there is a non-negative number \( \delta, \delta \in [0, (1-2(e+1))/(2+3)] \) such that

\[ \sum_{n=0}^{\infty} \|y_{n}\|^{2} + \|u_{n}\|^{2} = O(n^{\delta}), \hskip0.5cm \text{a.s.} \] (52)

then

\[ (p_{n}, q_{n}, r_{n}) \rightarrow_{n \rightarrow \infty} (p_{0}, q_{0}, r_{0}), \hskip0.5cm \text{a.s.} \]

where \((p_{n}, q_{n}, r_{n})\) is given by (17)–(19) with \( a_{n} \) being any sequence satisfying

\[ \frac{(\text{log} n)(\text{log} \log n)^{5\alpha n^{1/2}}}{a_{n}} \rightarrow_{n \rightarrow \infty} 0, \hskip0.5cm \text{for some } a > 1, \] (53)

and

\[ \frac{a_{n}}{n^{1/2(-e-1+\varepsilon)}} \rightarrow_{n \rightarrow \infty} 0. \] (54)

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

\[ \sum_{n=0}^{\infty} \|w_{n}\|^{2} = O(n^{\delta}). \] (55)

**Proof.** Noticing that \( 1 - (e+1)(e+\delta) > 0 \), we know that there exists a sequence \((a_{n})\) satisfying (53), (54). By Assumption (H2) and (52), it follows that

\[ a_{n} = O(n^{1/2}). \] (56)

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

\[ \liminf_{n \rightarrow \infty} n^{-a_{n}}[(p_{n}, q_{n}, r_{n})] = 0, \hskip0.5cm \text{a.s.} \] (55)

for any \((p, q, r) \in M^{*}\), where \( a_{n} = 1 - (e+1)(e+\delta) \).

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

\[ \eta = (\eta^{(1)}, \ldots, \eta^{(p+e)}(p+e-1), \ldots, \eta^{(p+e)}) \]

\[ = \sum_{i=0}^{p+e} \beta_{i}^{(p+e)}z(\text{Adj} A(z)[B(z), C(z)]) \]

\[ + \sum_{i=0}^{p+e} \beta_{i}^{(p+e)}z(\text{det} A(z), 0) \]

\[ + \sum_{i=0}^{p+e} \gamma^{(p+e)}z(0, (\text{det} A(z)), A(z)) \] (56)

We now show that

\[ \deg \left( \sum_{i=0}^{p+e} \alpha_{i}^{(p+e)}x^{i} \right) = p_{0} - 1. \] (57)

If \((p, q, r) = (p_{0}, q_{0}, r_{0})\) then (57) is trivial.

If \((p, q, r) = (p^{*}, q_{*}, r_{*})\), then by (56) and (H2) we have

\[ \deg \left( \sum_{i=0}^{p+e} \beta_{i}^{(p+e)}x^{i} \right) = (m+1)p_{0} + q_{0} \]

\[ = \deg \left( \sum_{i=0}^{p+e} \gamma^{(p+e)}z(\text{det} A(z)) \right) \leq q_{0} - 1 + mp_{0}, \]

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.), there exist matrix polynomials $M(z)$, $N(z)$ and $L(z)$ such that

$$A(z)M(z) + \{B(z), C(z)\}N(z) + L(z) = 1.$$ 

From this and (56) it follows that

$$\sum_{i=0}^{p-1} \alpha^{(i)}z^i \text{Adj} A(z) = \sum_{i=0}^{q-1} \alpha^{(i)}z^i (\text{Adj} A(z)) (A(z) M(z)$$

$$+ [B(z), C(z)]N(z) + L(z))$$

$$= (\det A(z)) \left( \sum_{i=0}^{p-1} \alpha^{(i)}z^i M(z) - \sum_{i=0}^{q-1} \beta^{(i)}z^i N(z) - \sum_{i=0}^{r-1} \gamma^{(i)}z^i L(z) \right).$$

By (57), we have

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

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


