Parameter and Order Estimation in A Class of Multivariate Stochastic Systems

By

Hajime AKASHI*, Kazuo NOSE** and Kamal MOUSTAFA***

(Received December 27, 1975)

Abstract

The problem of estimating the parameters and order of a class of multivariate systems is treated. The considered systems are described by a stochastic time invariant linear difference equation. We will introduce the so called canonical form III as a possible unique representation of the system. We will show that by using this canonical form, the computational effort compared with other canonical forms can be reduced. Further, we will show that the pole-zero cancellation, which is one of the methods used in identifying the order of single input-single output systems, can be extended to the multivariate systems in canonical form III.

1. Introduction

In this paper, we consider the problem of parameter and order estimations associated with the stochastic dynamical multivariate discrete system represented in eq. (1. 1)

$$\begin{aligned} y(k) &= T_1(D)u(k)+T_2(D)w(k) \\ T_1(0)&=0,\ T_2(0)=I \end{aligned} \tag{1. 1}$$

where D is the unit delay operator, $y(\cdot)$ is an m-dimensional observable covariance stationary random process, $u(\cdot)$ is a p-dimensional observable stationary random process with a positive definite covariance matrix, and $w(\cdot)$ is an m-dimensional zero mean white Gaussian noise Process with a covariance matrix ρ. The input u and the noise w are assumed to be mutually independent, and $w(k)$ is independent of $y(k-j)$ for all $j>0$. The transfer matrices $T_1(D)$ and $T_2(D)$ are of appropriate dimension whose elements are rational functions of the delay operator D.

We are interested in the estimation of $T_1(D)$ and $T_2(D)$, and in the determi-

* Department of Precision Mechanics.
** Centeral Laboratory, Kobe Steel Ltd., Kobe, Japan.
*** Department of Precision Mechanics, Exchange graduate student from Egypt working for doctor's degree from 1973 to 1977.

nation of the order of the considered system, using only the observation history of y and u processes. We will represent eq. (1. 1) by the linear difference equation (1. 2). The arguments D will be suppressed if there is no cause for confusion.

$$A(D)y(k)=G(D)u(k)+B(D)w(k) \tag{1. 2}$$

where

$$A(D)=A_0+A_1D+\cdots\cdots+A_{m1}D^{m1},\ A_0=I$$
$$G(D)=G_1D+\cdots\cdots\cdots\cdots+G_{m2}D^{m2}$$
$$B(D)=B_0+B_1D+\cdots\cdots+B_{m3}D^{m3},\ B_0=I$$
$$T_1=A^{-1}G,\ \ T_2=A^{-1}B$$

and m_1, m_2, and m_3 are defined to be the degrees of A, G, and B, respectively. The matrices A, G, and B are assumed to obey assumptions $A1$) and $A2$) of section 2.

Most of the principal methods of parameter estimation deal with the difference equation in what is called 'canonical form I', which is characterized by the lower triangularity of A, with the degree of $(A)_{ij} \leqq$ the degree of $(A)_{ii}$. Almost all the algorithms using this canonical form are rather complex[1], because all the parameters of the system are estimated simultaneously. The number of parameters is $[m^2(m_1+m_3)+mpm_2]$, which becomes very large even for small m because of the quadratic factor m^2.

In this paper, we will introduce what is called 'canonical form III', which is characterized by the diagonality of A. Using this canonical form, we will decompose the system into m subsystems each involving the estimation of a single difference equation. Kashyap and Robert[1] have introduced the so called 'canonical form II' which has the property that B is diagonal, and used it also in decomposing the system into m estimation problems. However, canonical form II involves more parameters than our proposed canonical form, because the latter is a special case of canonical form I, which has the important feature of the minimum number of unknown parameters with respect to any other canonical form[1].

While canonical forms I and II do not help much in simplifying the determination of the system order, our proposed canonical form is very useful in estimating the order of a certain class of multivariate systems. The diagonality of A makes it possible to obtain an explicit expression for the Smith form $[A, G]_s$, of the two matrices A and G by a simple method, from which the order of the system can be determined. This may be regarded as an extension of the pole-zero cancellation technique which is one of the methods used in identifying the order of univariate systems[2,3], to the multivariate case.

The structure of this paper is as follows. In section 2, we will show that eq. (1. 1) can be represented by eq. (1. 2) in canonical form III, and discuss the estimability of the parameters. Then, we will show how to transform canonical form III into canonical form II. In section 3, we will discuss the estimation algorithm using the least square technique; and in section 4, we will consider the order estimation problem. Finally, we will present a numerical example in section 5, to show how to obtain canonical form II from canonical form III, and that the latter has less parameters than the former.

2. Canonical Form III, and Uniqueness of Representation

In this section, we will discuss the uniqueness of the representation of eq. (1. 1) by eq. (1. 2) in canonical form III. In the sequel, we will assume the following unless otherwise stated.

A1) All the zeroes of the determinants of A and B lie outside the unit circle, to assure the stationarity and the invertibility of the process y[1,4].

A2) The Smith form of $[A, G, B]$ is $[I, 0]$, or equivalently A, G, and B have the greatest common left divisor (g. c. l. d.) as a unimodular matrix.

We have shown in the Appendix that eq. (1. 1) can be represented by the difference equation (1. 2) in canonical form III. Although this canonical form is a special case of canonical form I whose uniqueness has been proved by Hannan[5], we will give in Theorem 1 an alternate simpler proof for the uniqueness of canonical form III. This uniqueness of representation is necessary in order to obtain consistent estimates for the matrices A, G, and B[4].

Theorem 1

Define the sets

$\underline{M}_{m\times m}$: the set of all $m\times m$ unimodular matrices.

$\underline{N}_{m\times m}$: the set of all $m\times m$ non-singular† polynomial diagonal matrices.

Consider the difference equation (1. 2) characterized by the 3 tuple (A, G, B) with $A \in \underline{N}_{m\times m}$, and $P \in \underline{M}_{m\times m}$, where P is a g. c. l. d. of A, G, and B. It uniquely represents the process y in eq. (1. 1).

Proof

Lte (A, G, B) and $(\hat{A}, \hat{G}, \hat{B})$ be two representations for eq. (1. 1) with A, $\hat{A} \in \underline{N}_{m\times m}$ and P, $\hat{P} \in \underline{M}_{m\times m}$, where $\hat{P}$ is a g. c. l. d. of $\hat{A}$, $\hat{G}$, and $\hat{B}$. We proved in the Appendix that there exist two matrices H and K such that, H, $K \in \underline{N}_{m\times m}$ and

$$HA=K\hat{A},\quad HG=K\hat{G},\quad HB=K\hat{B} \tag{2. 1}$$

Since P, $\hat{P} \in \underline{M}_{m\times m}$, the Smith forms of (A, G, B) and $(\hat{A}, \hat{G}, \hat{B})$ are[6,7,8]

$$U[A:G:B]V=[I_m:0:0] \tag{2. 2}$$

$$\hat{U}[\hat{A}:\hat{G}:\hat{B}]\hat{V}=[I_m:0:0] \tag{2. 3}$$

† A polynomial matrix $A(D)$ is non-singular if there is only a finite number of values of D such that det. $[A(D)]=0$.

where

$$U,\ \hat{U},\ \epsilon\ \underline{M}_{m\times m},\ V,\ \hat{V}\ \epsilon\ \underline{M}_{(2m+p)\times(2m+p)}$$

Combining eqs. (2. 2) and (2. 3) leads to

$$[HU^{-1}:0:0]V^{*}=[K\hat{U}^{-1}:0:0] \tag{2. 4}$$

where $V^{*}\ \epsilon\ \underline{M}_{(2m+p)\times(2m+p)}$ is defined as

$$V^{*}\ \underline{\Delta}\ V^{-1}\hat{V}$$

If we partition V^{*} as

$$V^{*}=\begin{array}{c} \left[\begin{array}{c:c} V_{11} & V_{12} \\ \hdashline \underbrace{V_{21}}_{m} & \underbrace{V_{22}}_{m+p} \end{array}\right] \end{array}\begin{array}{l} \}m \\ \}m+p \end{array} \tag{2. 5}$$

then eq. (2. 4) gives us

$$HU^{-1}V_{11}=K\hat{U}^{-1},\quad V_{12}=0 \tag{2. 6}$$

The determinants of both sides of eq. (2. 5) are then given by[9]

$$|V^{*}|=|V_{11}|\cdot|V_{22}|=\text{constant}$$

which is satisfied if and only if $V_{11}\ \epsilon\ \underline{M}_{m\times m}$, and $V_{22}\ \epsilon\ \underline{M}_{(m+p)\times(m+p)}$. Hence, eq. (2. 6) gives us

$$H=KU^{\circ} \tag{2. 7}$$

where

$$U^{\circ}\ \epsilon\ \{\underline{M}_{m\times m}\cap\underline{N}_{m\times m}\}$$

because $H,\ K\ \epsilon\ \underline{N}_{m\times m}$.

It is readily apparent from eq. (2. 7) that U° has constant elements. Since $A(0)=I$, the first equation of eqs. (2. 1) and eq. (2. 7) yield

$$U^{\circ}=I$$

and hence

$$H=K$$

Therefore, as clear from eq. (2. 1)

$$A=\hat{A},\quad G=\hat{G},\quad B=\hat{B}$$

and Theorem 1 is proved.

Having established the uniqueness of the representation of eq. (1. 2) in canonical form III with $P\ \epsilon\ \underline{M}_{m\times m}$, we will discuss in Lemma 1 how to obtain canonical form II, which still represents eq. (1. 1) uniquely (provided that a certain set is non-empty) from canonical form III. The procedure will be demonstrated in the numerical example of section 5. We will also briefly discuss the relation between the three canonical forms I, II, and III.

Lemma 1

Denote a g. c. l. d. of $A\ \epsilon\ \underline{N}_{m\times m}$, $G\ (=AT_1)$, and $B\ (=AT_2)$ by $P\ \epsilon\ \underline{M}_{m\times m}$, and assume the set

$$\underline{Q}\ \underline{\Delta}\ \{QA\}\cap\{B^{\circ}T_2^{-1}|B^{\circ}\ \varepsilon\ \underline{N}_{m\times m}\}$$

be non-empty, and let

$$A^* \in \underline{O}$$

then, $P^* \in \underline{M}_{m\times m}$, if and only if $Q \in \underline{M}_{m\times m}$, where P^* is a g. c. l. d. of A^*, G^* $(=A^*T_1)$, and $B^*(=A^*T_2)$. Furthermore, when $P^* \in \underline{M}_{m\times m}$, (A^*, G^*, B^*) is a unique representation of eq. (1. 1).

Proof

We have

$$\begin{aligned} T_1 &= A^{-1}G = A^{*-1}G^* \\ T_2 &= A^{-1}B = A^{*-1}B^* \end{aligned} \qquad (2.\ 8)$$

It is not difficult to prove that there exist two non-singular polynomial matrices L and M such that

$$LA = MA^*, \quad LG = MG^*, \quad LB = MB^* \qquad (2.\ 9)$$

Since $A^* \in \underline{O}$, we have

$$A^* = QA \qquad (2.\ 10)$$

where we assume, $Q \in \underline{M}_{m\times m}$. Subistituting eq. (2. 10) into the first equation of eqs. (2. 9), we get

$$L = MQ \qquad (2.\ 11)$$

because A is non-singular. Using eq. (2. 11), eq. (2. 9) may be written

$$A = Q^{-1}A^*, \quad G = Q^{-1}G^*, \quad B = Q^{-1}B^* \qquad (2.\ 12)$$

The g. c. l. d. of A, G, and B is given by[8]

$$P = AR_1 + GR_2 + BR_3 \qquad (2.\ 13)$$

where R_1, R_2, and R_3 are polynomial matrices. Let N be any common left divisor of A^*, G^*, and B^*, and then, by using eq. (2. 12), we may write eq. (2. 13)

$$QP = N[A_1^*R_1 + G_1^*R_2 + B_1^*R_3]$$

Therefore, QP is a right multiple of every common left divisor of A^*, G^*, and B^*, which means that QP is a g. c. l. d. of A^*, G^*, and B^*[10], i. e.,

$$P^* = Q_1QP$$

where $Q_1 \in \underline{M}_{m\times m}$. Then, $P^* \in \underline{M}_{m\times m}$, because Q, $P \in \underline{M}_{m\times m}$. Moreover, Kashyap and Robert[1] have proved that, when $P^* \in \underline{M}_{m\times m}$, the difference equation (1. 2) in canonical form II is a unique representation of eq. (1. 1). This completes the sufficiency proof, and the necessity proof is obvious. Thus, Lemma 1 is proved.

It has been noted previously that canonical form III is a special case of canonical form I. In other words, canonical form III implies canonical form I, but not the other way around. Kashyap and Robert[1] have shown how to transform canonical form II into canonical form I, and vice versa. However, the authors have shown that the reverse is not alawys true as can be shown by a simple counter example[11]. The relation between canonical form II and III has not yet been fully explored. However, Lemma 1 shows the possibility of the existence of systems that can be expressed in both canonical forms. An example of such systems is given in section 5.

3 Estimation Algorithm

In this section, we will discuss the estimation algorithm using the well known least square technique. We will decompose the system (1. 2) in canonical form III, into m subsystems, and estimate the parameters in every subsystem separately by minimizing the loss function to be defined by eq. (3. 2). Since the process $w(\cdot)$ has been assumed white Gaussian, the least square estimate will be consistent[12]. However, the estimate is not necessarily asymptotically efficient, i. e., its variance may not approach the Cramer-Rao lower bound.

Since A is diagonal, we can easily write the scalar output of the jth subsystem of eqn. (1. 2) as

$$y_j(k)=z_j^T(k-1)\theta_j+w_j(k), \quad j=1, 2, \cdots\cdots, m \tag{3. 1}$$

where $z_j^T(k-1)=[y_j(k-1), \cdots\cdots, y_j(k-m_1), u^T(k-1), \cdots\cdots, u^T(k-m_2), w^T(k-1), \cdots\cdots, w^T(k-m_3)]$ is an n_j-dimensional vector, $w_j(\cdot)$ is the jth component of the vecotr $w(\cdot)$, θ_j is an n_j-dimensional unknown parameter vector composed of all the unknown coefficients in the jth raws of A, G, and B, and $\sum_{j=1}^{m} n_j=n$.

The loss function to be minimized is

$$J_j=\frac{1}{N}\sum_{k=1}^{N} \hat{w}_j^2(k) \text{ for sufficiently large } N \tag{3. 2}$$

where $\hat{w}_j(\cdot)$ is the estimated noise recovered from eq. (3. 1) by using $\hat{\theta}_j$ instead of θ_j, and N is the total number of measurements.

Equation (3. 1) can be written

$$Y_j=\Omega_j\theta_j+W_j \tag{3. 3}$$

where

$$Y_j^T \triangleq [y_j(k), \cdots\cdots, y_j(k+N-1)]$$

$$W_j^T \triangleq [w_j(k), \cdots\cdots, w_j(k+N-1)]$$

and $\quad \Omega_j^T \triangleq [z_j(k-1), \cdots\cdots, z_j(k+N-2)]$

The estimate $\hat{\theta}_j$ is given by

$$\hat{\theta}_j=[\Omega_j^T\Omega_j]^{-1}\Omega_j^T Y_j \tag{3. 4}$$

Because we do not know the sample values of $w_j(\cdot)$ in the matrix Ω_j of eq. (3. 4), we replace them by their estimates $\hat{w}(\cdot)$. Since intermediate estimates of $\hat{\theta}_j$ are needed for the generation of $\hat{w}(\cdot)$, we can use an updating scheme discussed in detail by Eykhoff[12], and Talmon and Van Den Boom[13].

As an alternative method, the parameters estimate can be obtained by applying the 2 stage least square method[1,11]. Using this method, $\hat{w}_j(\cdot)$ is precomputed by fitting a high order auto regressive model to the y_j process in eq. (3. 1).

It should be noted that the use of canonical form III, compared with that of canonical form II, has reduced the number of parameters n_j in every subsystem, which results in a reduction of the computation time.

The next section will be devoted to discussing the problem of identifying the order of the considered system, utilizing canonical form III.

4 Order Identification

In this section, we will show that the proposed canonical form is very useful in identifying the order of a certain class of multivariate systems. The diagonality of A has enabled us to obtain an explicit expression for the Smith form of A and G. By searching for the greatest common divisiors (g. c. d. 's) of the elements of every row of A and G, the Smith form of $[A, G]$ is shown to have all its diagonal elements equal to 1, except the last element, which is shown to be the product of those g. c. d. 's. This method is rather simple, compared with other methods, for obtaining the Smith form[8], should A be non-diagonal. Then, we will present the order identification algorithm, using this Smith form, which can be considered as an application of the pole-zero cancellation effect to the multivariate case.

4. 1 Smith Form and Minimal Order

We will define the order n of the considered system as, $n=\delta|A|$[7], where $\delta(\cdot)$ means the degree of $(\cdot)$.

Lemma 2

The system

$$A(D)y(k)=G(D)u(k)$$

is minimal, i. e., has least order $n°$, if and only if the g. c. l. d. of A and G is unimodular. A proof for Lemma 2 can be found in Wolovich[6],

Theorem 2 gives an explicit expression for the Smith form of $[A, G]$, together with the necessary and sufficient conditions for the system to be minimal, when A satisfies the following assumption. (Later, we will study the possibility of relaxing this assumption.)

A3) The diagonal elements of A are mutually prime.

Theorem 2

Define

$$g_i° \triangleq \text{g. c. d. } \{g_{i1}, g_{i2}, \cdots\cdots, g_{ip}\} \quad i=1, \cdots\cdots, m \tag{4. 1}$$

$$\Delta_i \triangleq \text{g. c. d. } \{a_i, g_i°\} \quad i=1, \cdots\cdots, m \tag{4. 2}$$

where g_{ij} is the $ijth$ element of G, and a_i is the ith diagonal element of A. Assuming assumption A3) holds, then

i) The Smith form of A and G is $[R:0]$, where

$$R \triangleq \begin{bmatrix} I_{m-1} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \prod_{i=1}^{m} \Delta_i \end{bmatrix}$$

ii) The system is minimal, i. e., R is unimodular, if and only if, Δ_i, $i=1$, ……, m, are m constants.

Proof

Let the minors of A of order $j \leqq m-1$, be given by

$$a^j(k_1, k_2, \cdots\cdots, k_{m-j}) \triangleq \prod_{\substack{l=1 \\ (l \neq k_1,\ l \neq k_2, \cdots\cdots,\ l \neq k_{m-j})}}^{m} a_l \tag{4. 3}$$

where, k_1, ……, $k_{m-j}=1$, ……, m. Define the g. c. d. of all such minors as

$$h^j \triangleq \text{g. c. d. } \{ \prod_{(l \neq 1,\ l \neq 2, \cdots\cdots l \neq m-j)} a_l, \cdots\cdots, \prod_{(l \neq j,\ l \neq j+1, \cdots\cdots,\ l \neq m)} a_l \}$$

By assumption (A3), the R. H. S. of the above equation is 1, i. e.,

$$h^j = 1 \qquad j \leqq m-1 \tag{4. 4}$$

It is obvious that h^j given by eq. (4. 4) is also the g. c. d. of all minors of order j of [A:G]. Consider now the minors of order m of [A:G], which are constructed by replacing the ith column of A by the jth column of G, and denote such a minor by $\bar{a}^m(i, j)$,

$$\bar{a}^m(i, j) \triangleq \left| \begin{bmatrix} a_l & g_{1j} & & \\ & \ddots & 0 & \\ & & g_{ij} & \\ & 0 & \ddots & \\ & & g_{mj} & a_m \end{bmatrix} \right| = g_{ij} \cdot \prod_{\substack{l=1 \\ l \neq i}}^{m} a_l \tag{4. 5}$$

Define

$$h^m(i) \triangleq \text{g. c. d. } \{\bar{a}^m(i, 1), \cdots\cdots, \bar{a}^m(i, r)\} = \prod_{l \neq i} a_l g_i^{\circ} \tag{4. 6}$$

If we define $h^m(0) \triangleq |A| = \prod_{l=1}^{m} a_l$, then the g. c. d. of $h^m(i)$, $i=0$, 1, ……, m, is given by

$$h^m \triangleq \text{g. c. d. } \{ \prod_{l=1}^{m} a_l,\ g_1^{\circ} \prod_{\substack{l=1 \\ l \neq 1}}^{m} a_l, \cdots\cdots, g_m^{\circ} \prod_{\substack{l=1 \\ l \neq m}}^{m} a_l \}$$

Using eq. (4. 2), the above equation may be written

$$h^m = \prod_{i=1}^{m} \Delta_i \tag{4. 7}$$

In fact, h^m is also the g. c. d. of all minors of order m of [A:G], because it is clear that all the elements of any raw i have Δ_i as their g. c. d., Using eqs. (4. 4) and (4. 7), the invariant polynomials of [A:G] can then be written[9)]

$$i_1 = \frac{h^m}{h^{m-1}} = \prod_{i=1}^{m} \Delta_i, \quad i_j = \frac{h^{m-j+1}}{h^{m-j}} = 1, \quad j=2, \cdots\cdots, m \tag{4. 8}$$

where $h^0 \triangleq 1$. Therefore, the Smith form of the matrix [A:G] is

$$[A:G]_S=\left[\begin{array}{c:c:c} I_{m-1} & 0 & \\ \cdots\cdots & \cdots\cdots & 0 \\ 0 & \prod_{i=1}^{m} \Delta_i & \end{array}\right]$$

This completes the proof of part i). Now let Δ_i, $i=1, \cdots\cdots, m$, be m constants, then $|R|$ =constant and R is unimodular, and the necessity proof is obvious. This completes the proof of Theorm 2.

We will study now the possibility of relaxing assumption A3). Suppose that the Δ_i 's have a g. c. d. given by

$$\Delta_0=\text{g. c. d. } \{\Delta_1, \Delta_2, \cdots\cdots, \Delta_m\} \tag{4. 9}$$

In such a case, we have

$$a_i=a_i^\circ \Delta_0 \text{ and } \Delta_i=\Delta_i^\circ \Delta_0,\ i=1, 2, \cdots\cdots, m \tag{4. 10}$$

Then, assumption A3) is replaced by the following assumption:

A3′) The diagonal elements of A have Δ_0 as a g. c. d,, and the (a_i/Δ_0) 's are mutually prime.

Under this relaxed assumption, eqs. (4. 4) and (4. 7) become

$$h^j=(\Delta_0)^j \qquad j\leqq m-1 \tag{4. 11}$$

$$h^m=(\Delta_0)^m \prod_{i=1}^{m} \Delta_i \tag{4. 12}$$

respectively, and the invariant polynomials of $[A:G]$ are

$$i_1=\Delta_0 \prod_{i=1}^{m} \Delta_i^\circ,\ i_2=i_3=\cdots=i_m=\Delta_0$$

which implies that R now takes the form

$$R=\Delta_0\left[\begin{array}{c:c} I_{m-1} & 0 \\ \cdots\cdots & \cdots\cdots \\ 0 & \prod_{i=1}^{m} \Delta_i^\circ \end{array}\right] \tag{4. 13}$$

It should be noted that part ii) of Theorem 2 is not changed by relaxing assumption A3).

4. 2 Order Identification Algorithm

In this subsection, we will treat the problem of estimating the order of the considered class of systems.

The jth subsystem (eq. (3. 1)) can be written as

$$a_j(n_j^a)y_j(k)=\sum_{i=1}^{p} g_{ji}(n_{ji}^g)u_i(k)+\sum_{\substack{i=1\\ i\neq j}}^{m} b_{ji}(n_{ji}^b)w_i(k)$$

$$+b_{jj}(n_{jj}^b)w_j(k) \tag{4. 14}$$

where n_j^a, n_{ji}^g, and n_{ji}^b are the degrees of a_j, g_{ji}, and b_{ji}, respectively. Recall that the loss function of the jth sub-system is

$$J_j=\frac{1}{N}\sum_{k=1}^{N} \hat{w}_j^2(k)=E\{\hat{w}_j^2(k)\} \tag{4. 15}$$

for sufficiently large N and ergodic signals. The estimated noise $\hat{w}_j(k)$ is given by

$$\hat{w}_j(k)=\frac{1}{\hat{b}_{jj}}[\hat{a}_j(\hat{n}_j^a)y_j(k)-\sum_{i=1}^{p}\hat{g}_{ji}(\hat{n}_{ji}^g)u_i(k)-\sum_{\substack{i=1\\ i\neq j}}^{p}b_{ji}(\hat{n}_{ji}^b)w_i(k)]$$

where $\hat{a}_j$ is the jth diagonal element of the estimated matrix $\hat{A}$, and $\hat{g}_{ji}$, $\hat{b}_{ji}$ are the jith elements of the estimated matrices $\hat{G}$ and $\hat{B}$, respectively. $\hat{n}_j^a$, $\hat{n}_{ji}^g$, and $\hat{n}_{ji}^b$ are the degrees of $\hat{a}_j$, $\hat{g}_{ji}$, and $\hat{b}_{ji}$ respectively, and will be chosen such that, $\hat{n}_j^a>n_j^a$, $\hat{n}_{ji}^g>n_{ji}^g$ and $\hat{n}_{ji}^b>n_{ji}^b$, i. e.,

$$n_j^* \triangleq \min\{\hat{n}_j^a-n_j^a,\ \hat{n}_{ji}^g-\hat{n}_{ji}^g,\ \hat{n}_{ji}^b-\hat{n}_{ji}^b\}>0$$

Let the matrix Ω_j of eq. (3. 4) be written as

$$\begin{aligned}\Omega_j=[&\eta_j(k-1),\ \cdots\cdots,\ \eta_j(k-\hat{n}_j^a),\ \upsilon_1(k-1),\ \cdots\cdots,\\ &\upsilon_1(k-\hat{n}_{j1}^g),\ \cdots\cdots,\ \upsilon_p(k-1),\ \cdots\cdots,\ \upsilon_p(k-\hat{n}_{jp}^g),\\ &\zeta_1(k-1),\ \cdots\cdots,\ \zeta_1(k-\hat{n}_{j1}^b),\ \cdots\cdots,\ \zeta_m(k-1),\ \cdots\cdots,\\ &\zeta_m(k-\hat{n}_{jm}^b)]\end{aligned} \tag{4. 16}$$

where

$$\eta_j^T(k-l)\triangleq[y_j(k-l),\ \cdots\cdots\cdots,\ y_j(k+N-l-1)]$$

$$\upsilon_q^T(k-l)\triangleq[u_q(k-l),\ \cdots\cdots\cdots,\ u_q(k+N-l-1)],\ q=1,\ \cdots\cdots,\ p$$

$$\zeta_q^T(k-l)\triangleq[w_q(k-l),\ \cdots\cdots\cdots,\ w_q(k+N-l-1)],\ q=1,\ \cdots\cdots,\ m$$

Because of the linear constraint of eq. (3. 1) and the assumption that $n_j^*>0$, in the matrix Ω_j of eq. (4. 16), the vector $\eta_j(k-1)$ is a linear combination of the vectors $\eta_j(k-2)$, ……, $\eta_j(k-n_j^a-1)$, $\upsilon_1(k-2)$, ……, $\upsilon_1(k-n_{j1}^g-1)$, ……, $\upsilon_p(k-2)$, ……, $\upsilon_p(k-n_{jp}^g-1)$, $\zeta_1(k-2)$, ……, $\zeta_1(k-n_{j1}^b-1)$, ……, $\zeta_m(k-2)$, ……, $\zeta_m(k-n_{jm}^b-1)$. Therefore, $\Omega_j^T\Omega_j$ is singular, and the estimate $\hat{\theta}_j$ (eq. (3. 4)) is not unique. On the other hand, the loss function (eqn. (4. 15)) can be written as follows, where the arguments of a_j, g_{ji}, and b_{ji} will be dropped for convenience

$$\begin{aligned}J_j=E&\left\{\left[\sum_{i=1}^{p}\frac{\hat{a}_jg_{ji}-a_j\hat{g}_{ji}}{a_j\hat{b}_{jj}}u_i(k)\right]^2\right.\\ &\left.+\left[\sum_{\substack{i=1\\ i\neq j}}^{m}\frac{\hat{a}_jb_{ji}-a_j\hat{b}_{ji}}{a_j\hat{b}_{jj}}w_i(k)\right]^2+\left[\frac{\hat{a}_jb_{jj}}{a_j\hat{b}_{jj}}w_j(k)\right]^2\right\}\\ \geqq E&\left\{\left[\frac{\hat{a}_jb_{ji}}{a_j\hat{b}_{ji}}w_j(k)\right]^2\right\}\geqq E\{w_j^2(k)\}\end{aligned}$$

where the last inequality follows from Åströn and Söderström[14]. The absolute minimum points of J_j are thus given by

$$\frac{1}{a_j\hat{b}_{jj}}\sum_{i=1}^{p}(\hat{a}_jg_{ji}-a_j\hat{g}_{ji})u_i(k)=0$$

$$\frac{1}{a_j\hat{b}_{jj}}\sum_{\substack{i=1\\i\neq j}}^{m}(\hat{a}_j b_{ji}-a_j\hat{b}_{ji})w_i(k)=0 \tag{4.17}$$

$$\frac{\hat{a}_j b_{jj}}{a_j\hat{b}_{jj}}=1$$

At this point, we will assume the vector signal $u(k)$ to be persistently exciting of order $\bar{n} \triangleq \max_{i,j}(\hat{n}_j^a+n_{ji}^g,\ n_j^a+\hat{n}_{ji}^g)$, i. e.,

$$\bar{u} \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}u(k), \text{ and } R(\tau) \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}[u(k)-\bar{u}][u(k-\tau)-\bar{u}]^T \text{ exist and the}$$

matrix

$$\begin{bmatrix} R(0) & \vdots & & \vdots & R(\bar{n}-1) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R(\bar{n}-1) & \vdots & & \vdots & R(0) \end{bmatrix}$$

is positive definite. It is to be noted that the concept of the persistently exciting scalar signal has been studied in literature, e. g., Eykhoff[12]. Now, using the assumption that $u(k)$ and $w(k)$† are persistently exciting, it is relatively easy to show that eqs. (4. 17) can be replaced by

$$\hat{a}_j g_{ji}=a_j\hat{g}_{ji} \qquad i=1,\ \cdots\cdots,\ p$$

$$\hat{a}_j b_{ji}=a_j\hat{b}_{ji} \qquad i=1,\ \cdots\cdots,\ m$$

The above two equations have their general solution as

$$\begin{aligned} &\hat{a}_j=a_j c_j \\ &\hat{g}_{ji}=g_{ji}c_j \qquad i=1,\ \cdots\cdots,\ p \\ &\hat{b}_{ji}=b_{ji}c_j \qquad i=1,\ \cdots\cdots,\ m \end{aligned} \tag{4.18}$$

where

$$c_j \triangleq 1+c_{j,1}D+\cdots\cdots+c_{j,n_j^*}D^{n_j^*}$$

and the coefficients $c_{j,1},\ \cdots\cdots,\ c_{j,n_j^*}$ are arbitrary. Equations (4. 18) demonstrate the pole-zero cancellation that has been indicated by Gustavsson[2], and used to estimate the order of univariate systems by Van Den Boom and Van Den Enden[8].

Since we assume $n_j^*>0$, the system $(\hat{A}y(k)=\hat{G}u(k))$ is not minimal, and by Theorem 2, $\hat{\Delta}_j\neq$ constant for all j, where

$$\hat{\Delta}_j \triangleq \text{g. c. d. } \{\hat{a}_j,\ \hat{g}_{j1},\ \cdots\cdots,\ \hat{g}_{jp}\}=c_j\Delta_j \tag{4.19}$$

We can approximately determine $\hat{\Delta}_j$ by first estimating the parameter vector $\hat{\theta}_j$ and then plotting the pole-zero configuration for every transfer function $\hat{g}_{ji}/\hat{a}_j$, $i=1,\ \cdots\cdots,\ p$. We will notice poles and zeroes cancelling each other, which means that there are common factors between $\hat{g}_{ji}$ and $\hat{a}_j$ for $i=1,\ \cdots\cdots,\ p$. The g. c. d. of those common factors is, by eq. (4. 19). equal to $\hat{\Delta}_j$, the minimal order n° is then

† Since $w(k)$ is assumed white with a unity covariance matrix, it is always persistently exciting.

$$n^{\circ}=\delta|\hat{A}|-\delta\left\|\begin{bmatrix} I_{m-1} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \prod_{i=1}^{m}\hat{\Delta}_i \end{bmatrix}\right\|$$

$$=\delta|\hat{A}|-\delta\{\prod_{j=1}^{m}\hat{\Delta}_j\}$$

It should be noted that the method described above for determining the (minimal) order of the considered class of systems, can be regarded as an extension of the pole-zero cancellation, used in identifying the order of univariate systems[3]. In the univariate case, we are merely interested in finding the common factor between the numerator and the denominator of the single transfer function of the system. But in the multivariate case, we determine such common factors for all of the p transfer functions in the jth subsystem, and our interest is focused on the g. c. d. of those common factors. Also by virtue of the above discussion, the parameters in every subsystem can be estimated simultaneously along with its order, after the necessary pole-zero cancellations have been done in accordance with eqs. (4. 18). This can be done since we assume $n_j^* > 0$, $j=1, \cdots\cdots, m$.

5 Numerical Example

In this section, we will show by a numerical example, how to obtain canonical form II from canonical form III, and that the latter has less parameters than the former. This leads to computational labour saving in estimating the parameters of the system.

Consider the process in canonical form III given below, with $m=2$ and $p=1$,

$$Ay(k)=Gu(k)+Bw(k)$$

The numerical values of the polynomial matrices A, G, and B are

$$A=\begin{pmatrix} 1+D+0.25D^2 & 0 \\ 0 & 1+0.8D+0.16D^2 \end{pmatrix}$$

$$G=\begin{pmatrix} D+0.8D^2 \\ D^2 \end{pmatrix} \qquad B=\begin{pmatrix} 1 & 0 \\ -D & 1+0.3D \end{pmatrix}$$

If we choose the unimodular matrix Q as

$$Q=\begin{pmatrix} 1 & 0 \\ D & 1 \end{pmatrix}$$

then, the 3 tuple (A^*, G^*, B^*) in canonical form II is

$$A^*=QA=\begin{pmatrix} 1+D+0.25D^2 & 0 \\ D+D^2+0.25D^3 & 1+0.8D+0.16D^2 \end{pmatrix}$$

$$G^*=QG=\begin{pmatrix} D+0.8D^2 \\ 2D^2+0.8D^3 \end{pmatrix} \qquad B^*=QB=\begin{pmatrix} 1 & 0 \\ 0 & 1+0.3D \end{pmatrix}$$

It is obvious that the Smith form of both $[A, G, B]$ and $[A^*, G^*, B^*]$ is $[I_m:0]$, and that the latter representation has more parameters than the former.

6 Conclusion

The parameter estimation problem in multivaiate stochastic linear time invariant systems was considered. By introducing canonical form III, the system is decomposed into m subsystems, each involving the estimation of a single difference equation. This, cnmbined with the fact that canonical form III as a special case of canonical form I has the minimum number of parameters, leads to a reduction in computational labour, compared with systems which can be represented in canonical form II introduced by Kashyap and Robert[1]. We gave an alternative proof for the uniqueness of the representation in canonical form III, and briefly discussed the estimability of the parameters and the relation between the three canonical forms I, II, and III. The method used in estimation is the well known generalized least square estimate, in which both the system and noise parameters are estimated.

We treated also the problem of order identification of the considered class of systems. As a result of representing the system in the proposed canonical form (canonical form III), we extended the pole-zero cancellation method to identify the order of the considered systems. We have shown that the order of the system can be obtained by treating the subsystems separately. We pointed out that by extending the pole-zero cancellation to the multivariate case, the parameters and order can be estimated simultaneously when the model order is assumed greater than the actual order.

Appendix

We will first show that eq. (1.1) can be represented by eq. (1.2) in canonical form III. Let

$[T_1(D)]_{ij}=c^1_{ij}(D)/d^1_{ij}(D)$, $[T_2(D)]_{ij}=c^2_{ij}(D)/d^2_{ij}(D)$ with $c^1_{ij}(D)$ and $d^1_{ij}(D)$, $c^2_{ij}(D)$ and $d^2_{ij}(D)$ are relatively prime polynomials for all i and j. Define

$$a_{ii}\overset{\Delta}{=}\alpha_i \text{ L. C. M. } (d^1_{i1}, \cdots\cdots, d^1_{ip}, d^2_{i1}, \cdots\cdots, d^2_{im}),\ i=1, \cdots\cdots, m$$

$$A(D)\overset{\Delta}{=}\text{diagonal } (a_{11}, \cdots\cdots, a_{mm})$$

where L. C. M. denotes the least common multiple, and the scalars α_i are chosen such that $a_{ii}(D)$ are monic polynomials for $i=1, \cdots\cdots, m$, then

$$G(D)=A(D)T_1(D) \text{ and } B(D)=A(D)T_2(D)$$

are polynomial matrices with $G(0)=0$ and $B(0)=I$. If the g. c. l. d. of A, G, and B is unimodular, the representation (A, G, B) is in canonical form III.

Next, we will prove eq. (2. 1). Since (A, G, B) and $(\hat{A}, \hat{G}, \hat{B})$ are two representations for the process y in eq. (1. 1). it is not difficult to show that

$$G=W\hat{G} \tag{A. 1}$$

$$B = W\hat{B} \tag{A. 2}$$

where

$$W \triangleq A\hat{A}^{-1}\dagger \tag{A. 3}$$

is a diagonal matrix because both the representations (A, G, B) and $(\hat{A}, \hat{G}, \hat{B})$ are in canonical form III. The matrix W can be readily decomposed as

$$W = H^{-1}K \tag{A. 4}$$

where H and K are non-singular polynomial diagonal matrices. Now, substituting eq. (A. 4) into eqs. (A. 3), (A. 1), and (A. 2), we obtain

$$HA = K\hat{A}, \qquad HG = K\hat{G}, \qquad HB = K\hat{B}$$

which is eq. (2. 1).

References

1) R. L. Kashyap and Robert E. Nasburg: IEEE Trans. Automat. Contr., 19, 784 (1974).
2) I. Gustavsson: Report 6903, Lund Institute of Technology, Sweden (1969).
3) A. J. W. Van Den Boom and A. W. M. Van Den Enden: Automatica, 10, 245 (1974).
4) P. E. Caines and J. Rissanen: IEEE Trans. Inform. Theory, 20, 102 (1974).
5) E. J. Hannan: Econometrica, 39, 751 (1971).
6) W. A. Wolovich: "Linear Multivariable Systems", Springer-Verlag, New York (1974).
7) H. H. Rosenbrock: "State-Space and Multivariable Theory", John Wiley & Sons, Inc., New york (1970).
8) N. Suda and S. Kodama: Systems and Control, 19, 190 (1975).
9) F. R. Gantmacher: "The Theory of Matrices", Chelsa Publishing Company, New York (1960).
10) W. A. Wolovich: Proceedings of the 2nd IFAC Symposuim, Duesseldorf, October, 1. 2. 3 (1971).
11) H. Akashi, K. Nose and K. A. Moustafa: IEEE Trans. Automat. Contr., 21, 141 (1976).
12) P. Eykhoff: "System Identification", John Wiley & Sons, London (1974).
13) J. L. Talmon and A. J. W. Van Den Boom: Proceedings of the 3rd IFAC symposuim · Identification and system parameter estimation, The Hague/Delft, 711 (1973).
14) K. J. Åström and T. Söderström: IEEE Trans. Automat. Contr., 19, 769 (1974).

† Note that the matrix W is not necessarily a polynomial matrix.