Parameter Identification of Systems with Noise in Input and Output

By

Hajime AKASHI* and Kamal A. F. MOUSTAFA**

(Received March 31, 1975)

Abstract

In the problem of system parameters identification, most treatments made previously have assumed that the input to the system was free from noise. However, there would be many instances, where it would be more practical to assume the input to be accompanied with additive noise. Such a case is considered in the present paper, and the asymptotic unbiased estimate is obtained under certain conditions. The extended matrix approach with the ordinary least square method is used for the estimation of the parameters of the systems and the noise filter. Order identification is also discussed for this system with input noise. An application of the obtained solution to a numerical example shows that it gives a satisfactory result, both in parameter identification and in order identification.

1. Introduction

In most of the studies on the system parameters identification, the input to the system is considered free from noise. However, it would be natural to assume that in many actual problems, the system input is also contaminated with measurement noise. If the input is accompanied with noise, then it is found that the input and the equation error become correlated. Therefore, the assumption[1] made generally in the estimation of system parameters, that the input and the equation error are uncorrelated, no longer holds true.

In the present paper, the general situation in which the input and the output are both contaminated with noise will be discussed by the ordinary least square (OLS) approach. It is shown that, due to the correlation between the input and the equation error, the estimation of the parameters will be made

* Department of Precision Mechanics.
** Department of Precision Mechanics, Exchange graduate student from Egypt working for doctor's degree from 1973 to 1977.

only with a bias, unless an additional requirement is fulfilled. In section 4, one such requirement is shown to be that the first parameter of the numerator of the system pulse transfer function (b_0 by notation) should be known without error.

In section 5, the above requirement, (which may not always be satisfied in actual problems) is replaced by another, more practical, condition that the noise covariance should be known *a priori*. With the knowledge of the noise covariance, it is shown that the estimate of the bias can be calculated if a certain scalar quantity is non-zero. This will be discussed in the lemma given in section 5.

Then in section 6, the determination of the order of the system will be discussed for the system considered here, in which the input noise is present. The order of both the system and the noise filter, can be determined, by evaluating the loss function for a number of trial orders.

Finally, a numerical example is considered to illustrate the result obtained in the present paper. It is found that the parameters of both the system and the noise filter can be estimated with a fair accuracy, and the order can be determined clearly by the behaviour of the loss function.

2. Problem Formulation

Let $x(k)$ and $r(k)$ be the exact output and input signals at the discrete time k respectively, and the system equation be given by

$$x(k)=[B(z^{-1})/A(z^{-1})]r(k) \tag{2.1}$$

where $B(z^{-1})$ and $A(z^{-1})$ are defined by

$$B(z^{-1})=b_0+\sum_{i=1}^{q} b_i z^{-i}, \qquad b_0\neq 0 \tag{2.2a}$$

and

$$A(z^{-1})=1+\sum_{i=1}^{q} a_i z^{-i} \tag{2.2b}$$

b_i's and a_i's in the above equations are the system parameters. In the present paper, both the output and the input are assumed to accompany additive noise, as

$$\begin{aligned} y(k)&=x(k)+n(k)\\ u(k)&=r(k)+v(k) \end{aligned} \tag{2.3}$$

where $n(k)$ and $v(k)$ are the noises with respective characteristics and uncorrelated with $r(k)$. The generalized model of this system is shown by the diagram in Fig. 1. Here, a vector notation will be used for the input and output data for

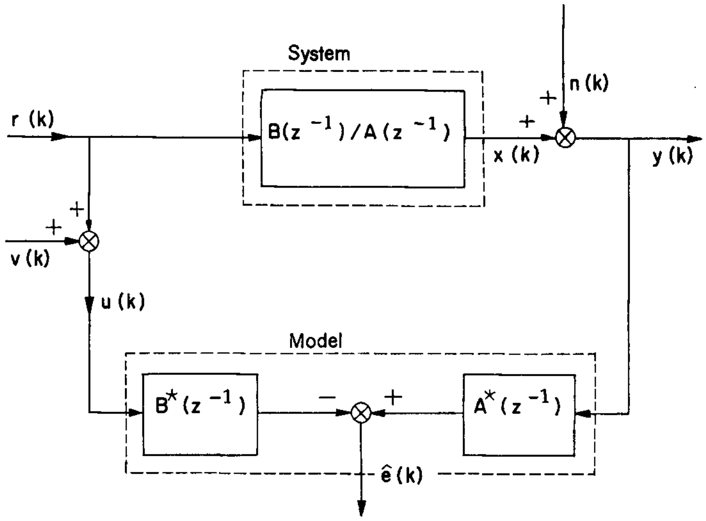

Fig. 1. The system and its generalized model.

the number of samples K

$$\underset{\sim}{u}^T \triangleq [u(k), u(k-1), \cdots, u(k+K-1)]$$

$$\underset{\sim}{y}^T \triangleq [y(k), y(k-1), \cdots, y(k+K-1)]$$

Similar definitions will be given to the noise vectors $\underset{\sim}{n}$ and $\underset{\sim}{v}$. If $\Omega_q(u, y)$ is defined as

$$\Omega_q(u, y) \triangleq \begin{bmatrix} u(k), & \cdots\cdots, & u(k-q), & y(k-1), & \cdots\cdots, & y(k-q) \\ \vdots & & \vdots & \vdots & & \vdots \\ u(k+K-1), & \cdots, & u(k+K-q-1), & y(k+K-2), & \cdots, & y(k+K-q-1) \end{bmatrix}$$

$$\triangleq [U : Y] \tag{2.4}$$

and $\underset{\sim}{\theta}$ is the vector of the parameters to be identified,

$$\underset{\sim}{\theta}^T \triangleq [b_0, b_1, \cdots, b_q, -a_1, \cdots, -a_q] \tag{2.5}$$

then eq. (2.1) can be written

$$\underset{\sim}{y} = [\Omega_q(u, y)]\underset{\sim}{\theta} + \underset{\sim}{e} \tag{2.6}$$

where the vector $\underset{\sim}{e}$ is given by

$$\underset{\sim}{e} = \underset{\sim}{n} - \Omega_q(v, n)\underset{\sim}{\theta} \tag{2.7}$$

In the above, $\underset{\sim}{e}$ is the vector with K elements of

$$e(k) = A(z^{-1})n(k) - B(z^{-1})v(k) \tag{2.8}$$

and $\Omega_q(v, n)$ is the same as eq. (2.4), with the argument replaced by v and n.

Now, the problem of our concern is to find the estimate of the parameter $\underline{\theta}$ of the system (2.1) and (2.2), given the data of input and output with noise as given by (2.3). That the input also accompanies the noise $\underline{v}$ is the new assumption made in the present paper. Because of this, the ordinary least square method, OLS, will lead only to a biased estimate of the parameters even though the equation error, $e(k)$, is white†. In the following sections it will be shown that this is the case, unless a certain restrictive condition is satisfied.

3. Biased Least Square Estimate

Let $\hat{\underline{e}}$ be the estimate of the equation error $\underline{e}$ obtained by the estimated parameter $\hat{\underline{\theta}}$

$$\hat{\underline{e}}=\underline{y}-[\Omega_q(u, y)]\hat{\underline{\theta}} \tag{3.1}$$

The error function V will be difined by

$$V\triangleq\frac{1}{K}[\hat{\underline{e}}^T\hat{\underline{e}}] \tag{3.2}$$

Minimizing the loss function with respect to the estimate $\hat{\underline{\theta}}$ (OLS)[4], we have

$$\hat{\underline{\theta}}=[\Omega_q^T(u, y)\Omega_q(u, y)]^{-1}\Omega_q^T(u, y)\underline{y} \tag{3.3}$$

Using eqs. (2.6) and (2.7) in the above, we have

$$\hat{\underline{\theta}}=\underline{\theta}+[\Omega_q^T(u, y)\Omega_q(u, y)]^{-1}\Omega_q^T(u, y)[\underline{n}-\Omega_q(v, n)\underline{\theta}] \tag{3.4}$$

The second term in the above equation is the bias of the estimate of the parameter vector. Taking the probability limit of eq. (3.4) as $K\to\infty$, we have

$$\begin{aligned}\operatorname*{plim}_{K\to\infty}\hat{\underline{\theta}}=\underline{\theta}-&\left[\operatorname*{plim}_{K\to\infty}\frac{\Omega_q^T(u, y)\Omega_q(u, y)}{K}\right]^{-1}\\ &\times\left[\operatorname*{plim}_{K\to\infty}\frac{\Omega_q^T(u, y)\Omega_q(v, n)}{K}\underline{\theta}-\operatorname*{plim}_{K\to\infty}\frac{\Omega_q^T(u, y)\underline{n}}{K}\right]\end{aligned} \tag{3.5}$$

For simplicity, we will drop the notation of $K\to\infty$ of eq. (3.5) in later descriptions. The above equation may be written as

$$\operatorname{plim}\hat{\underline{\theta}}=\theta-Q^{-1}\left[\begin{pmatrix} Q_{vv} & Q_{vn} \\ Q_{vn}^T & Q_{nn}\end{pmatrix}\theta-\begin{pmatrix}\phi_{nv}(0)\\ \vdots \\ \phi_{nv}(q) \\ \phi_{nn}(1) \\ \vdots \\ \phi_{nn}(q)\end{pmatrix}\right] \tag{3.6}$$

† As is well known, if $e(k)$ is white and $v(k)\equiv 0$, the OLS method will lead to an asymptotically unbiased estimate of the parameters[3,4].

where

$$Q \triangleq \operatorname{plim}\left[\frac{\Omega_q^T(u, y)\Omega_q(u, y)}{K}\right] \tag{3.7}$$

which is a nonsingular matrix for $K>2q+1$[6)]

$$Q_{vv} \triangleq \begin{bmatrix} \phi_{vv}(0), & \cdots\cdots\cdots, & \phi_{vv}(q) \\ \vdots & \ddots & \vdots \\ \phi_{vv}(q), & \cdots\cdots\cdots, & \phi_{vv}(0) \end{bmatrix} \tag{3.8}$$

$$Q_{nn} \triangleq \begin{bmatrix} \phi_{nn}(0), & \cdots, & \phi_{nn}(q-1) \\ \vdots & \ddots & \vdots \\ \phi_{nn}(q-1), & \cdots, & \phi_{nn}(0) \end{bmatrix} \tag{3.9}$$

$$Q_{vn} \triangleq \begin{bmatrix} \phi_{nv}(1), & \cdots\cdots\cdots, & \phi_{nv}(q) \\ \hdashline \phi_{nv}(0), & \cdots, & \phi_{nv}(q-1) \\ \vdots & \ddots & \vdots \\ \phi_{nv}(q-1), & \cdots, & \phi_{nv}(0) \end{bmatrix} \tag{3.10}$$

$$\phi_{\xi\xi}(j) \triangleq \operatorname{plim}\frac{1}{K}\sum_{i=k}^{k+K-1}\xi(i)\xi(i-j),$$

$$\phi_{\xi\eta}(j) \triangleq \operatorname{plim}\frac{1}{K}\sum_{i=k}^{k+K-1}\xi(i)\eta(i-j) \tag{3.11}$$

Suppose we assume that the noises n and v are generated from the zero mean stationary white sequences $w_1(k)$ and $w_2(k)$ by the equations

$$w_1(k) = A(z^{-1})n(k) \tag{3.12a}$$

and

$$w_2(k) = B(z^{-1})v(k) \tag{3.12b}$$

and that $\underline{w}_1$ and $\underline{w}_2$ have the following property:

$$\begin{aligned} E[\underline{w}_1\underline{w}_1^T] &= \sigma_1^2 I \\ E[\underline{w}_2\underline{w}_2^T] &= \sigma_2^2 I \\ E[\underline{w}_1\underline{w}_2^T] &= \sigma_3^2 I \end{aligned} \tag{3.13}$$

where $E[\cdot]$ denotes the expectation sign, and I is the identity matrix. Let us define $w(k)$ as

$$w(k) \triangleq w_1(k) - w_2(k).$$

Then, we have

$$w(k) = n(k) + \sum_{i=1}^{q} a_i n(k-i) - \{b_0 v(k) + \sum_{i=1}^{q} b_i v(k-i)\} \tag{3.14}$$

Now, multiplying the above equation successively by $n(k-1), \cdots, n(k-q)$, $v(k)$, $\cdots, v(k-q)$, taking the expectation, and arranging the resulting equations in matrix form, we get

$$\begin{bmatrix} Q_{vv} & Q_{vn} \\ Q_{vn}^T & Q_{nn} \end{bmatrix} \underline{\theta} = \begin{bmatrix} \phi_{nv}(0) \\ \vdots \\ \phi_{nv}(q) \\ \cdots \\ \phi_{nn}(1) \\ \vdots \\ \phi_{nn}(q) \end{bmatrix} - \begin{bmatrix} g \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \tag{3.15}$$

In deriving the above, we used the following property:

$$E[w(k)n(k-i)] = 0 \quad \text{for} \quad i=1, \cdots, q$$

$$E[w(k)v(k-i)] = \begin{cases} 0 & \text{for} \quad i=1, \cdots, q \\ g & \text{for} \quad i=0 \end{cases} \tag{3.16}$$

where g is a non-zero scalar, indicating the correlation of $w(k)$ and $v(k)$. Equations (3.16) hold true because $n(k)$ and $v(k)$ are dependent only on the past values of $w(k)$. Substituting eq. (3.15) into (3.6), we have†

$$\lim E[\hat{\underline{\theta}}] = \underline{\theta} + Q^{-1} \begin{bmatrix} g \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.17}$$

Since g is not zero as is clear from eq. (3.16), the estimate is asymptotically biased, although the equation error is assumed white.

Comparing eq. (3.12) with (2.8), we see that $e(k) = w(k)$, and is therefore white.

4. Unbiased Estimation Assuming b_0 Known

In this section, it will be shown that if the parameter b_0 of $\underline{\theta}$ is assumed known *a priori*, then it is possible to obtain an unbiased estimate of all the other parameters. The generalized model including filters for the noise, which satisfies

---

† The non-random vector $\underline{c}$ is the plim of $\underline{x}(\kappa)$ for $\kappa \to \infty$, i.e.,

$$\text{plim}\, \underline{x}(\kappa) = \underline{c}$$

if

$$\lim_{\kappa \to \infty} P\{|\underline{x}(\kappa) - \underline{c}| \geqq \varepsilon)\} = 0 \quad \text{for each} \quad \varepsilon > 0$$

This implies

$$\lim_{\kappa \to \infty} E\{\underline{x}(\kappa)\} = \underline{c}$$

the assumptions given in the sequel will be discussed.

Asymptotically unbiased estimation

Let us now consider a matrix equation obtained by dropping the first row of eq. (3.15), and rearranging it as follows:

$$\begin{bmatrix} \phi_{vv}(1) \\ \vdots \\ \phi_{vv}(q) \\ \cdots \\ \phi_{nv}(1) \\ \vdots \\ \phi_{nv}(q) \end{bmatrix} b_0 + \left[\begin{array}{c:c} \begin{matrix} \phi_{vv}(0), \cdots, \phi_{vv}(q-1) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{vv}(q-1), \cdots, \phi_{vv}(0) \end{matrix} & \begin{matrix} \phi_{nv}(0), \cdots, \phi_{nv}(q-1) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{nv}(q-1), \cdots, \phi_{nv}(0) \end{matrix} \\ \hdashline \begin{matrix} \phi_{nv}(0), \cdots, \phi_{nv}(q-1) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{nv}(q-1), \cdots, \phi_{nv}(0) \end{matrix} & Q_{nn} \end{array}\right] \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_q \\ \cdots \\ -a_1 \\ \vdots \\ -a_q \end{bmatrix} = \begin{bmatrix} \phi_{nv}(1) \\ \vdots \\ \phi_{nv}(q) \\ \cdots \\ \phi_{nn}(1) \\ \vdots \\ \phi_{nn}(q) \end{bmatrix} \tag{4.1}$$

Defining the vector $\underline{z}$ as

$$\underline{z} \triangleq \underline{y} - b_0 \underline{u} \tag{4.2}$$

and $\underline{\theta}'$ as

$$\underline{\theta}' \triangleq [b_1, \cdots, b_q, \ -a_1, \cdots, \ -a_q]^T \tag{4.3}$$

Similarly

$$\Omega_q'(u, y) \triangleq [U' \vdots Y] \tag{4.4}$$

where

$$U' \triangleq \begin{bmatrix} u(k-1), \cdots\cdots\cdots, u(k-q) \\ \vdots \qquad\qquad \vdots \\ u(k+K-2), \cdots, u(k+K-q-1) \end{bmatrix} \tag{4.5}$$

Then, eq. (2.6) is written

$$\underline{z} = \Omega_q'(u, y)\underline{\theta}' + \underline{e} \tag{4.6}$$

Defining

$$Q' \triangleq \text{plim}\left[\frac{\Omega_q'^T(u, y)\Omega_q'(u, y)}{K}\right] \tag{4.7}$$

we may write eq. (3.6) as follows:

$$\lim E[\hat{\underline{\theta}}'] = \underline{\theta}' - Q'^{-1}\left[\left[\begin{array}{c:c} \begin{matrix} \phi_{vv}(0), \cdots, \phi_{vv}(q-1) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{vv}(q-1), \cdots, \phi_{vv}(0) \end{matrix} & \begin{matrix} \phi_{nv}(0), \cdots, \phi_{nv}(q-1) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{nv}(q-1), \cdots, \phi_{nv}(0) \end{matrix} \\ \hdashline \begin{matrix} \phi_{nv}(0), \cdots, \phi_{nv}(q-1) \\ \vdots \quad \ddots \quad \vdots \\ \phi_{nv}(q-1), \cdots, \phi_{nv}(0) \end{matrix} & Q_{nn} \end{array}\right]\right.$$

$$\times \underline{\theta}' - \begin{pmatrix} \phi_{nv}(1) \\ \vdots \\ \phi_{nv}(q) \\ \cdots\cdots \\ \phi_{nn}(1) \\ \vdots \\ \phi_{nn}(q) \end{pmatrix} + b_0 \begin{pmatrix} \phi_{vv}(1) \\ \vdots \\ \phi_{vv}(q) \\ \cdots\cdots \\ \phi_{nv}(1) \\ \vdots \\ \phi_{nv}(q) \end{pmatrix} \Big] \tag{4.8}$$

But, the second term in the above equation vanishes if we substitute eq. (4.1). Thus

$$\lim E[\hat{\underline{\theta}}'] = \underline{\theta}' \tag{4.9}$$

which means that $\underline{\theta}'$ is an asymptotically unbiased estimate.

It is thus made clear that, if the equation error is white and if the parameter b_0 is known *a priori*, then we can obtain an asymptotic and an unbiased estimate of all the other parameters by OLS.

Generalized model

The white equation error arises if the noise given to the system is considered to be produced by a filter to which white noise is given as input[3,4]. It will be shown now that such an extended model will be the case discussed in the above. Fig. 2 shows the block diagram of this extended model.

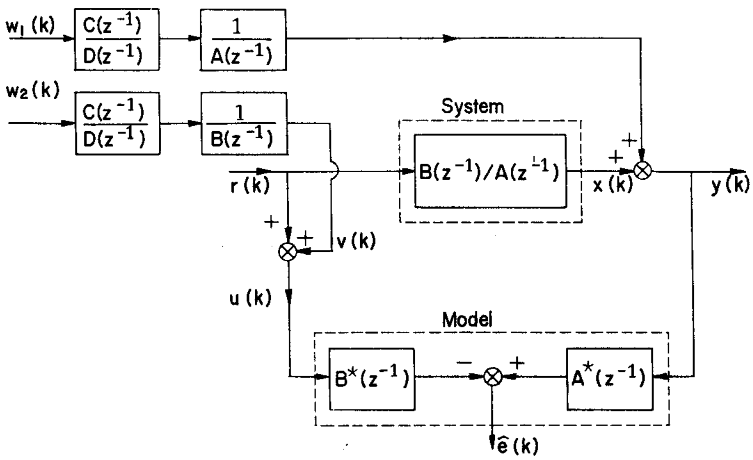

Fig. 2. System model and noise filter.

From the figure, we see that the following equations hold

$$n(k) = \frac{1}{A(z^{-1})} \frac{C(z^{-1})}{D(z^{-1})} w_1(k)$$

$$v(k)=\frac{1}{B(z^{-1})}\frac{C(z^{-1})}{D(z^{-1})}w_2(k) \tag{4.10}$$

We will consider what is called the moving average and auto-regressive model for the noise filter. Thus

$$\begin{aligned} C(z^{-1})&=1+\sum_{i=1}^{s}c_i z^{-i} \\ D(z^{-1})&=1+\sum_{i=1}^{s}d_i z^{-i} \end{aligned} \tag{4.11}$$

Here, s is the order, c_i's and b_i's are the noise filter parameters. Then, $e(k)$† is given by

$$e(k)=A(z^{-1})n(k)-B(z^{-1})v(k)=\frac{C(z^{-1})}{D(z^{-1})}w(k) \tag{4.12}$$

writing in vector form as eq. (2.6), we have

$$\underline{e}=[W \,\vdots\, \bar{E}]\begin{pmatrix} \underline{c} \\ \cdots\cdots \\ -\underline{d} \end{pmatrix}+\underline{w} \tag{4.13}$$

where

$$\bar{E}\triangleq\begin{pmatrix} e(k-1), & \cdots\cdots\cdots\cdots, & e(k-s) \\ \vdots & & \vdots \\ e(k+K-2), & \cdots, & e(k+K-s-1) \end{pmatrix}, \quad W\triangleq\begin{pmatrix} w(k-1), & \cdots\cdots\cdots\cdots, & w(k-s) \\ \vdots & & \vdots \\ w(k+K-2), & \cdots, & w(k+K-s-1) \end{pmatrix}$$

$$\underline{c}^T\triangleq[c_1,\cdots,c_s], \qquad \underline{d}^T\triangleq[d_1,\cdots,d_s] \tag{4.14}$$

Substituting eq. (4.13) into eq. (4.6), we may write

$$\underline{z}=\Omega_q(u,y,w,e)\underline{\gamma}+\underline{w} \tag{4.15}$$

where

$$\Omega_q(u,y,w,e)\triangleq[\Omega_q'(u,y) \,\vdots\, W \,\vdots\, \bar{E}]$$

and

$$\underline{\gamma}^T\triangleq[\underline{b}^T, -\underline{a}^T, \underline{c}^T, -\underline{d}^T]$$

The estimate $\hat{\underline{\gamma}}$ is obtained by minimizing

$$V'=\frac{1}{K}[\hat{\underline{w}}^T\hat{\underline{w}}] \tag{4.16}$$

where

$$\hat{\underline{w}}\triangleq\underline{z}-\Omega_q(u,y,w,e)\hat{\underline{\gamma}} \tag{4.17}$$

† For the generalized model, the equation error is defined as $w(k)$.

We have

$$\hat{\underline{\gamma}}=[\Omega_q^T(u, y, w, e)\Omega_q(u, y, w, e)]^{-1}\Omega_q^T(u, y, w, e)\underline{z} \tag{4.18}$$

This estimate is based on the assumptions mentioned earlier, namely that the equation error $w(k)$ is stationary white and b_0 is known *a priori*. It is seen from eq. (4.18) that it is asymptotically unbiased, i.e.,

$$\lim E[\hat{\underline{\gamma}}]=\underline{\gamma} \tag{4.19}$$

5. Unbiased Estimation of All the Parameters

The case in which a particular parameter, b_0, is known exactly, may not always be practical. If b_0 is assumed unknown like other parameters, it is still possible to make an unbiased estimatation with an alternative assumption, that is, the covariance matrix of the noise is known.

We first obtain the unbiased estimate of b_0 with the above assumption, and then use this estimate in the calculation of other parameters according to section 4.

Estimation of b_0

The condition that $\hat{b}_0$ minimizes the loss function (4.16) is

$$\frac{\partial V'(\hat{b}_0)}{\partial \hat{b}_0}=0$$

which leads to

$$\hat{b}_0=[\underline{u}^T\underline{u}]^{-1}[\underline{u}^T\underline{y}-\underline{u}^T\Omega_q(u, y, w, e)\hat{\underline{\gamma}}] \tag{5.1}$$

Now, eq. (4.18) becomes

$$\hat{\underline{\gamma}}=[\Omega_q^T(u, y, w, e)\Omega_q(u, y, w, e)]^{-1}\Omega_q^T(u, y, w, e)\hat{\underline{z}} \tag{5.2}$$

where

$$\hat{\underline{z}}\triangleq\underline{y}-\hat{b}_0\underline{u} \tag{5.3}$$

Substituting $\hat{\gamma}$ from eq. (5.2) into eq. (5.1), we have

$$\hat{b}_0=[\underline{u}^TQ_2\underline{u}]^{-1}[\underline{u}^TQ_2\underline{y}] \tag{5.4}$$

where

$$Q_2\triangleq I-Q_1 \tag{5.5}$$

and

$$Q_1\triangleq\Omega[\Omega^T\Omega]^{-1}\Omega^T \tag{5.6}$$

Here, the arguments of Ω were dropped for simplicity. From eq. (5.4), it is seen that $[\underline{u}^T Q_2 \underline{u}]$ has to be non-zero, in order for $\hat{b}_0$ to exist. The necessary and sufficient condition for $[\underline{u}^T Q_2 \underline{u}]$ to be non-zero will be obtained in the following lemma.

Lemma

The scalar, $h=\underline{u}^T[I-Q_1]\underline{u}=\underline{u}^T Q_2 \underline{u}$, is greater than zero if and only if

1) $K>2(q+s)$
2) The control sequence $\{u(i)\}$ is not all zero for $i=1, \cdots, K$.

Proof

In the equation

$$h=\underline{u}^T\underline{u}-\underline{u}^T Q_1 \underline{u} \tag{5.7}$$

The matrix Q_1 is idempotent[4], i. e., $Q_1{}^2=Q_1$. Therefore, the above equation may be written as

$$h=\underline{u}^T\underline{u}-\underline{u}^T Q_1{}^T Q_1 \underline{u}=\underline{t}^T\underline{t} \tag{5.8}$$

where

$$\underline{t}\triangleq\underline{u}-\underline{u}_1 \tag{5.9}$$

and

$$\underline{u}_1\triangleq Q_1\underline{u} \tag{5.10}$$

Transforming the matrix Q_1 into the diagonal form $Q_1{}^*$ by means of the non-singular matrix T, we have

$$Q_1{}^*=T^{-1}Q_1T \tag{5.11}$$

Since the matrix Q_1 is idempotent, it can be shown that its eigen values are either 0 or 1†. Therefore, $Q_1{}^*$ takes the form

$$Q_1{}^*=\mathrm{diag}\,[1, 1, \cdots, 1, 0, \cdots, 0] \tag{5.12}$$

where the number of 1's in the above equation equals the rank of the matrix Q_1, i. e.,

$$\mathrm{rank}\,[Q_1{}^*]=\mathrm{rank}\,[Q_1]=2(q+s) \qquad \text{for} \quad K\geqslant 2(q+s)^{6)} \tag{5.13}$$

Let us define the vactor $\underline{u}^*$ as

$$\underline{u}^*\triangleq T^{-1}\underline{u} \tag{5.14}$$

† See appendix.

To prove the sufficiency of the lemma, let $K>2(q+s)$, then the i th element of the vactor $[Q_1{}^*\underline{u}^*]$ is given by

$$[Q_1{}^*\underline{u}^*]_i=\begin{cases}u(i)^* & 1\leqslant i\leqslant 2(q+s)\\ 0 & 2(q+s)<i\leqslant K\end{cases} \tag{5.15}$$

where $u(i)^*$ is the i th element of $\underline{u}^*$. Since T is nonsingular, then for the control sequence not all zero we have from eq. (5.14)

$$u(i)^*\neq 0 \qquad \text{for some} \quad i=2(q+s),\ \cdots, K \tag{5.16}$$

Comparing eq. (5.15) with (5.16), we see that

$$Q_1{}^*\underline{u}^*\neq\underline{u}^* \tag{5.17}$$

Substituting for $Q_1{}^*$ from eq. (5.11), and using eqs. (5.10) and (5.14), we have

$$\underline{u}_1\neq\underline{u} \tag{5.18}$$

and therefore

$$\underline{t}\neq 0 \tag{5.19}$$

This completes the sufficiency proof as is clear from eq. (5.8).

In order to prove the necessity of the lemma, let

$$K=2(q+s)^{\dagger}$$

Then, we may write

$$\text{rank}\,[Q_1{}^*]=\text{rank}\,[Q_1]=K \tag{5.20}$$

which simply means

$$Q_1{}^*\underline{u}^*=\underline{u}^* \tag{5.21}$$

which in turn means that $h=0$, and this proves the necessity of the first condition. Let $u(i)=0$, $i=1, \cdots, K$, then, eq. (5.14) is gives as

$$\underline{u}^*=0 \tag{5.22}$$

and consequently

$$Q_1{}^*\underline{u}^*=\underline{u}^*$$

which again means that $h=0$. This shows the necessity of the second condition.

Asymptoticity of $\hat{b}_0$

Substituting the value of $\underline{y}$ from eq. (4.2) into eq. (5.1) and using (4.15), we may write

† It is clear from eq. (5.6) that K cannot be smaller than $2(q+s)$.

$$\hat{b}_0 = b_0 + [\underline{u}^T\underline{u}]^{-1}\underline{u}^T[\underline{w} + \Omega(\underline{\gamma} - \hat{\underline{\gamma}})] \tag{5.22}$$

Defining $\Delta b_0 \triangleq \hat{b}_0 - b_0$, eq. (5.2) may be written

$$\hat{\underline{\gamma}} = \underline{\gamma} - \Delta b_0 [\Omega^T\Omega]^{-1}\Omega^T\underline{u} \tag{5.23}$$

Substibuting eq. (5.23) into eq. (5.22) and rearranging, we obtain

$$\Delta b_0 = [\underline{u}^T Q_2 \underline{u}]^{-1}\underline{u}^T\underline{w} \tag{5.24}$$

Taking the probability limit as $K \to \infty$, we get the following, since $\underline{w}_1$ and $\underline{w}_2$ are both assumed uncorrelated with the input $\underline{\gamma}$

$$\text{plim}\,(\Delta b_0) = q_1^{-1}\,\text{plim}\left[\frac{\underline{v}^T\underline{w}}{K}\right] \tag{5.25}$$

where

$$q_1 \triangleq \text{plim}\,\frac{\underline{u}^T Q_2 \underline{u}}{K}$$

But $v(k)$ is related to $w_2(k)$ by

$$v(k) = \frac{C(z^{-1})}{B(z^{-1})D(z^{-1})} w_2(k) \tag{5.26}$$

or

$$v(k) = \frac{1}{b_0} w_2(k) + \epsilon(k-1) \tag{5.27}$$

where $\epsilon(k-1)$ contains terms depending only on the previous values of $w_2(k)$ and $v(k)$. Since $w_2(k)$ is assumed white, we can write

$$\begin{aligned} E[v(k)w_2(k)] &= \frac{1}{b_0}E[w_2(k)w_2(k)] + E[w_2(k)\epsilon(k-1)] \\ &= \frac{1}{b_0}E[w_2(k)w_2(k)] \end{aligned} \tag{5.28}$$

Since the noise has been assumed to be stationary, eq. (5.25) is written as†

$$\lim E[\Delta b_0] = q_1^{-1} \lim \{E[v(k)w_1(k)] - E[v(k)w_2(k)]\} \tag{5.29}$$

Considering the general case of correlating noise, we assume

$$w_1(k) = \nu w_2(k) \tag{5.30}$$

where ν is a real number. Using eqs. (5.28) and (5.30), we may write eq. (5.29) as

$$\lim E[\Delta b_0] = (\nu - 1) q_1^{-1}\left[\frac{\sigma_2^2}{b_0}\right] \tag{5.31}$$

---

† See footnote of page 189.

where $\sigma_2{}^2$, is the variaince of $w_2(k)$.

Substituting eq. (5.30) into eq. (5.22), we finally get

$$b_0 = 0.5\hat{b}_0\left[1+\sqrt{1+\frac{4q_2}{\hat{b}_0{}^2}}\right] \tag{5.32}$$

where

$$q_2 \triangleq \frac{(\nu-1)\sigma_2{}^2}{q_1}$$

and the limit sign is omitted. Summing up, we can get an asymptotically unbiased estimate for the parameter b_0 by first estimating the biased value from eq. (5.4), and then by using eq. (5.32).

6. Order Identification of the System

The equation error as estimated from eq. (4.17), based on a model of order p, is written as

$$\hat{\underline{w}} = \underline{z} - \Omega_p(u, y, w, e)\hat{\underline{\gamma}} = [\Omega_q\underline{\gamma} - \Omega_p\hat{\underline{\gamma}}] + \underline{w} \triangleq \underline{m} + \underline{w} \tag{6.1}$$

The loss function, eq. (4.16), can be written as

$$V' = \frac{1}{K}[\underline{m}^T + \underline{w}^T][\underline{m} + \underline{w}] \tag{6.2}$$

for $K \to \infty$, and since the estimator (4.18) is shown to be asymptotically unbiased, it can be shown that the following holds true[2]:

$$\text{plim}\,[\hat{a}_i] = \begin{cases} \neq a_i & \text{for} \quad p<q, \; i=1, \cdots, p \\ = a_i & \text{for} \quad p=q, \; i=1, \cdots, p \\ = a_i & \text{for} \quad p>q, \; i=1, \cdots, q \\ = 0 & \text{for} \quad p>q, \; i=q+1, \cdots, p \end{cases} \tag{6.3}$$

and similar relations also hold for the other parameters $\underline{b}$, $\underline{c}$ and $\underline{d}$. It can be shown that V' is written[2]

$$V' = \begin{cases} \dfrac{1}{K}\underline{w}^T\underline{w} & \text{for} \quad p \geqslant q \\ \dfrac{1}{K}[\underline{w}^T\underline{w} + \underline{m}^T\underline{m}] & \text{for} \quad p<q \end{cases} \tag{6.4}$$

In the same way as for the model without input noise[2], the order identification of the system with input noise can be detected by plotting eq. (6.4). For $p<q$, the variation of the function V' will be relatively large until $p=q$. For $p \geqslant q$, the plot will be more or less constant and we can see the change from the plot

of the V' function. The same argument can be used also for determining the noise model order.

7. Example

For the sake of comparison, the authors chose the same process as used by Åström[2,7], only with a non-zero value for b_0. The process dynamics is given by

$$x(k)=\frac{1+z^{-1}+0.5z^{-2}}{1-1.5z^{-1}+0.7z^{-2}}r(k)$$

and the noise filter is

$$\frac{C(z^{-1})}{D(z^{-1})}=\frac{1+0.2z^{-1}}{1-0.6z^{-1}}$$

Signals $r(k)$, $w_1(k)$ and $w_2(k)$ are taken to be zero mean white Gaussian sequences with variances 1.0, 0.2 and 0.1, respectively. The calculation was carried out for the case of $w_1(k)$ and $w_2(k)$ are mutually independent, and for the case where tney are dependent with $w_2(k)=0.5w_1(k)$.

In the correlated case, the parameters were estimated first by assuming b_0 known, and then assuming b_0 unknown along with other parameters. A total of 1,500 samples was made available for the calculation.

8. Results and Conclusion

The OLS method is used for the identification of parameters of discrete systems, in which not only the output but also the input are contaminated with

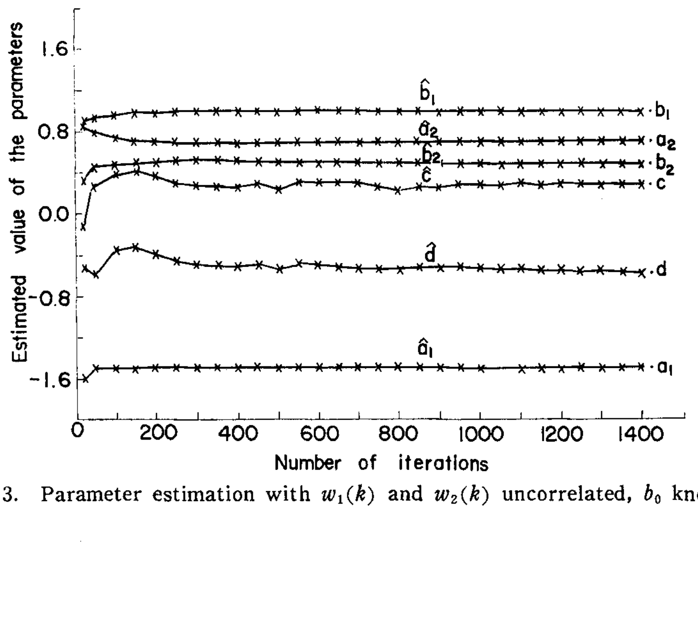

Fig. 3. Parameter estimation with $w_1(k)$ and $w_2(k)$ uncorrelated, b_0 known.

additive noise.

First, it was found that if the equation error is white and if the first parameter of the numerator of the system, b_0, is known *a priori*, then it is possible to estimate the other parameters asymptotically without bias. Fig. 3 shows the result of applying the method to the illustrative example for the case of uncorrelated input and output measurement noises. It shows satisfactory identification precess, already giving a good estimation for about 300 samples. Fig. 4 shows the same for the case of correlated input and output measurement noises, also with good performance.

Secondly, it was shown that, instead of the perfect knowledge of b_0, the

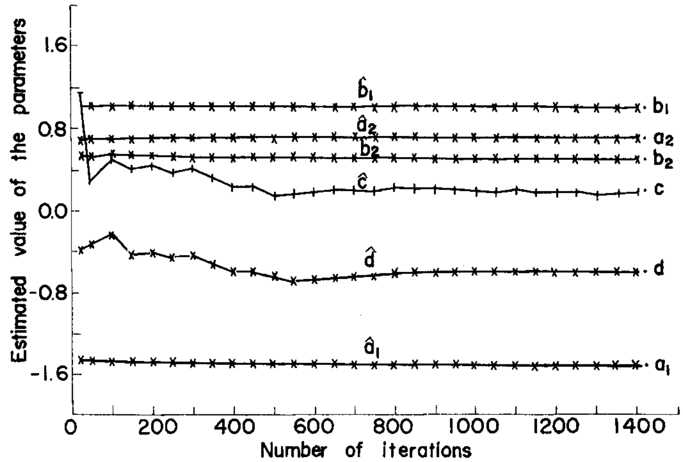

Fig. 4. Parameter estimation with $w_1(k)$ and $w_2(k)$ correlated, b_0 known.

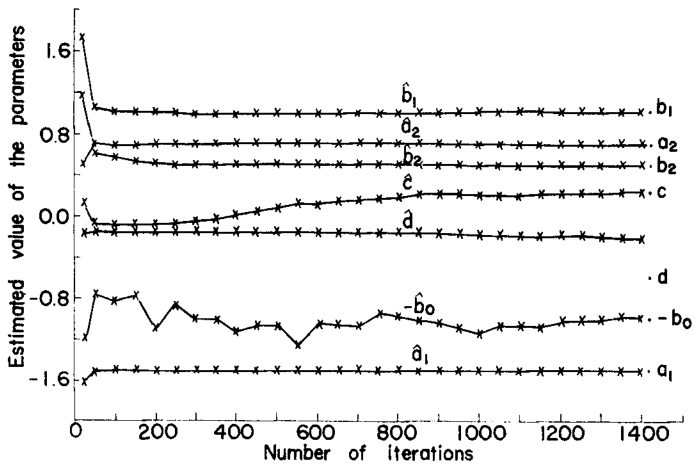

Fig. 5. Estimation of all the parameters with $w_1(k)$ and $w_2(k)$ correlated, b_0 unknown.

assumption of the knowledge of the noise covariance matrix can lead to the asymptotic estimation of all the parameters including b_0 without bias. Fig. 5 presents the result of estimation for this case. Excepting b_0, the other parameters show a reasonably fast convergence to the true values. It is seen that the estimation of b_0 takes more samples than others. The reason for such a behaviour can not be determined from this example alone, further investigation in the future is intended. Order identification for the problem considered in this paper, namely the system where input also accompanies noise, was tried for various values of noise order.

It is found that if s is assumed greater than zero, the system order is correctly identified to be 2, as illustrated in Fig. 6. If, however, s is taken to be zero, the determination of the system order is questionable.

Having determined the system order, the loss function was plotted against various values of the order s, as shown in Fig. 7. It is found that for the right order of the system, which is 2, the value of s is clearly identified as unity as seen in Fig. 7, and it is the correct order of the filter for the noise.

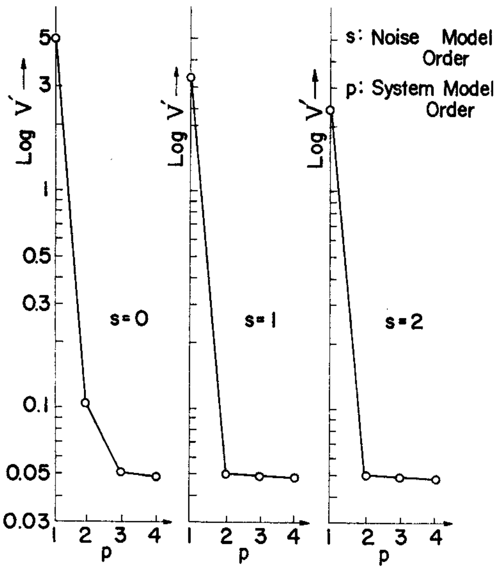

Fig. 6. Determination of the order of the system.

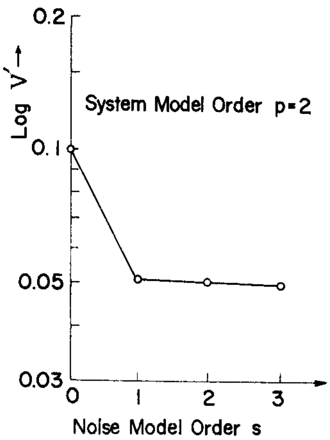

Fig. 7. Determination of the noise order.

Appendix

Let the ith eigen value of Q_1 be denoted by λ_i, and the corresponding eigen vector by ν_i, then

$$Q_1 \underline{\nu}_i = \lambda_i \underline{\nu}_i \tag{A.1}$$

Since the matrix Q_1 is idempotent, we have

$$Q_1 = Q_1^2 \tag{A.2}$$

therefore

$$Q_1 \underline{\nu}_i = Q_1 Q_1 \underline{\nu}_i = \lambda_i^2 \underline{\nu}_i \tag{A.3}$$

Comparing eq. (A.1) with eq. (A.3), we see that

$$\lambda_i = \lambda_i^2$$

The above equation is satisfied only for $\lambda_i = 0$ or 1.

References

1) R. L. Kashyap and Robert E. Nasburg: "Parameter estimation in multivariate stochastic difference equations," IEEE Trans. Automat. Contr., Vol. AC-19, pp. 784–797 (1974).
2) A. J. W. Van Den Boom and A. W. M. Van Den Enden: "The determination of the orders of process and noise dynamics", Automatica, Vol. 10, pp. 245–256 (1974).
3) J. L. Talmon and A. J. W. Van Den Boom: "On the estimation of transfer parameters of process and noise dynamics using a single stage estimator", Proceedings third IFAC Symposium of Identification, The Hague/Delft (1973).
4) P. Eykhoff: "System Idenfitication", John Wiley & Sons, London (1974).
5) J. M. Mendel: "Discrete Techniques of Parameter Estimation", Marcel Dekker, Inc., New York (1973).
6) R. C. K. Lee: "Optimal Estimation, Identification and Control", The M. I. T. Press, Cambridge, Massachusetts (1964)
7) K. J. Åström: "Lectures on the identification problem—the least squares method", Report 6806, Lund Institute of Technology, Sweden (1968).