ディジタル信号処理手法による連続系の同定に関する研究

楊, 子江
九州大学工学研究科電子工学専攻

https://doi.org/10.11501/3088155

出版情報：九州大学, 1991, 博士（工学）, 課程博士
バージョン：

Chapter 5

Identification in the Presence of Input-Output Measurement Noises Using Bias-Compensated Least-Squares Method

5.1 Introduction

In the previous chapters, the direct recursive identification algorithms using the digital filtering techniques have been discussed, in the case where only the output is corrupted by a measurement noise. And it was confirmed that the pass-band of the pre-filters should be chosen such that it matches that of the system under study as closely and that when only the output measurement is corrupted by a high measurement noise, the IV methods give excellent results.

However, in some practical situations, it may not be possible to avoid the noise when measuring the input signal. In this case, the standard identification methods may give erroneous results (Söderström 1981).

Models where both inputs and outputs are contaminated by errors are usually called errors-in-variables (EV) models. The problem of identifiability of the EV models has been discussed by Anderson and Deistler (1984, 1987), Anderson (1985), Deistler and Anderson (1989), Kalman (1983) etc. So far, only a limited number of works have discussed the identification algorithms of discrete-time systems in the presence of input noise. Söderström (1981) investigated some methods of system identification from noise-corrupted input-output data.

And it was pointed out that the joint-output JO approach using a prediction error method, while being quite computationally demanding, leads to accurate results (as expected, in view of the optimal PEM accuracy properties). The uniqueness problem of this method has been discussed by Stoica and Nehorai (1987). These pioneering works may be the best known in the literature.

In the case where both the input and output measurements are corrupted by white noises, Fernando and Nicholson (1985) proposed the Koopmans-Levin method based on the singular-value decomposition . Although this method is computationally robust, it requires *a priori* knowledge of the ratio of the variances of the input and output noises. Wada and Eguchi (1986) applied the efficient BCLS method (Sagara and Wada 1977) without *a priori* knowledge of the noise variances. A similar idea was also investigated by Feng and Zheng (1988), Zheng and Feng (1989), which is considered as a direct extension of the work of Sagara and Wada (1977).

In this work, the problem of identification of continuous systems is considered when both the discrete input and output measurements are contaminated by white noises (Sagara, Yang and Wada 1991d). It will be found that in the presence of input measurement noise, it is not appropriate to let the pass-band of the filters match that of the continuous system under study as suggested in some previous works. Our simulation results will show that in this case the pass-band of the digital low-pass filters should be chosen such that it includes the main frequencies of both the system input and output signals in some range. Since most physical systems are low-pass systems, we emphasize that the selection of the pass-band of the pre-filters should be based on the main frequencies of the input signals which excite the system modes.

When the pre-filters are designed appropriately such that the effects of the noises are sufficiently reduced, the LS method still gives acceptable results. In the case of high noises, since the LS estimate is biased, then the BCLS method is utilized to obtain consistent estimates. The BCLS algorithm compensates the bias of the LS estimate by the estimates of both the input and output noise variances and hence yields a consistent estimate. This approach seems to be more convenient than the early works. Both classes of filters (FIR filter and IIR filter) are employed. The FIR filters can be applied to the BCLS method directly, whereas the IIR filters require some approximations.

5.2 Statement of the problem

Consider the following SISO continuous system described by the ordinary differential equation

$$
\begin{aligned}
A(p)x(t) &= B(p)u(t) \\
A(p) &= \sum_{i=0}^{n} a_i p^{n-i} \quad (a_0 = 1) \\
B(p) &= \sum_{i=1}^{n} b_i p^{n-i}
\end{aligned}
\tag{5.1}
$$

Our goal is to identify the system parameters from the noisy sampled input-output data:

$$
\begin{aligned}
y(k) &= x(k) + e(k) \\
w(k) &= u(k) + v(k)
\end{aligned}
\tag{5.2}
$$

where $v(k)$ and $e(k)$ are white noises such that

$$
\begin{gathered}
\mathrm{E}[e(k)] = 0, \ \mathrm{E}[e(k)^2] = \sigma_e^2 \\
\mathrm{E}[v(k)] = 0, \ \mathrm{E}[v(k)^2] = \sigma_v^2 \\
\mathrm{E}[e(k)v(k)] = 0, \ \mathrm{E}[u(k)v(k)] = 0, \ \mathrm{E}[u(k)e(k)] = 0
\end{gathered}
\tag{5.3}
$$

Since differential operations may accentuate the noise effects, it is necessary here to introduce a digital low-pass filter which would reduce the noise effects sufficiently as described in the previous chapters. Then we can obtain a discrete-time estimation model with continuous system parameters.

5.3 Discrete-time estimation models

In this section, we describe the estimation models derived by using the two classes of filters, taking both the input and output noises into account.

5.3.1 FIR filtering approach

Using the FIR filters described in section 3.3.1, we have

$$
\sum_{i=0}^{n} a_i \xi_{Fiy}(k) = \sum_{i=1}^{n} b_i \xi_{Fiw}(k) + r_F(k) \tag{5.4}
$$

where

$$
\begin{aligned}
\xi_{Fiy}(k) &= Q_F(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}y(k) = \xi_{Fix}(k)+\xi_{Fie}(k)\\
\xi_{Fiw}(k) &= Q_F(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}w(k) = \xi_{Fiu}(k)+\xi_{Fiv}(k)\\
\xi_{Fix}(k) &= Q_F(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}x(k)\\
\xi_{Fiu}(k) &= Q_F(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}u(k)\\
\xi_{Fie}(k) &= Q_F(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}e(k) = \sum_{j=0}^{M_F+n} f_j^i z^{-j}e(k)\\
\xi_{Fiv}(k) &= Q_F(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}v(k) = \sum_{j=0}^{M_F+n} f_j^i z^{-j}v(k)\\
r_F(k) &= \sum_{i=0}^{n} a_i\xi_{Fie}(k) - \sum_{i=1}^{n} b_i\xi_{Fiv}(k) = \sum_{j=0}^{M_F+n}\alpha_j z^{-j}e(k) - \sum_{j=0}^{M_F+n}\beta_j z^{-j}v(k)\\
\alpha_j &= f_j^0 + \sum_{i=1}^{n} f_j^i a_i, \quad \beta_j = \sum_{i=1}^{n} f_j^i b_i
\end{aligned}
\tag{5.5}
$$

5.3.2 IIR filtering approach

Using the Butterworth digital IIR filter described in section 3.3.2, we have

$$
\sum_{i=0}^{n} a_i\xi_{Iiy}(k) = \sum_{i=1}^{n} b_i\xi_{Iiw}(k) + r_I(k) \tag{5.6}
$$

where

$$
\begin{aligned}
\xi_{Iiy}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}y(k) = \xi_{Iix}(k)+\xi_{Iie}(k)\\
\xi_{Iiw}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}w(k) = \xi_{Iiu}(k)+\xi_{Iiv}(k)\\
\xi_{Iix}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}x(k)\\
\xi_{Iiu}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}u(k)\\
\xi_{Iie}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}e(k)\\
\xi_{Iiv}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}v(k)\\
r_I(k) &= \sum_{i=0}^{n} a_i\xi_{Iie}(k) - \sum_{i=1}^{n} b_i\xi_{Iiv}(k)
\end{aligned}
\tag{5.7}
$$

5.4 LS method and its bias

When the digital low-pass filters have been designed, we have the discrete-time estimation model of equation (5.4) for the FIR filtering approach, or the model of equation (5.6) for the IIR filtering approach. Both can be written in vector form:

$$\begin{aligned}
\xi_{0y}(k) &= \mathbf{z}^T(k)\theta + r(k) \\
\mathbf{z}^T(k) &= [-\xi_{1y}(k),\cdots,-\xi_{ny}(k),\xi_{1w}(k),\cdots,\xi_{nw}(k)] \\
&= \left[-\mathbf{z_y}^T(k),\mathbf{z_w}^T(k)\right] = \left[-\left\{\mathbf{z_x}^T(k)+\mathbf{z_e}^T(k)\right\},\left\{\mathbf{z_u}^T(k)+\mathbf{z_v}^T(k)\right\}\right] \\
\theta^T &= [a_1,\cdots,a_n,b_1,\cdots,b_n] = \left[\mathbf{a}^T,\mathbf{b}^T\right]
\end{aligned} \tag{5.8}$$

where

$$\begin{aligned}
\mathbf{z_e}^T(k) &= [\xi_{1e}(k),\cdots,\xi_{ne}(k)] \\
\mathbf{z_x}^T(k) &= [\xi_{1x}(k),\cdots,\xi_{nx}(k)] \\
\mathbf{z_v}^T(k) &= [\xi_{1v}(k),\cdots,\xi_{nv}(k)] \\
\mathbf{z_u}^T(k) &= [\xi_{1u}(k),\cdots,\xi_{nu}(k)]
\end{aligned} \tag{5.9}$$

and

$$\begin{aligned}
&\mathbf{z}^T(k) = \mathbf{z_F}^T(k) \\
&\quad \xi_{iy}(k) = \xi_{Fiy}(k), \xi_{ix}(k) = \xi_{Fix}(k), \xi_{ie}(k) = \xi_{Fie}(k) \\
&\quad \xi_{iw}(k) = \xi_{Fiw}(k), \xi_{iu}(k) = \xi_{Fiu}(k), \xi_{iv}(k) = \xi_{Fiv}(k), r(k) = r_F(k) \quad \text{(FIR filter)} \\
&\mathbf{z}^T(k) = \mathbf{z_I}^T(k) \\
&\quad \xi_{iy}(k) = \xi_{Iiy}(k), \xi_{ix}(k) = \xi_{Iix}(k), \xi_{ie}(k) = \xi_{Iie}(k) \\
&\quad \xi_{iw}(k) = \xi_{Iiw}(k), \xi_{iu}(k) = \xi_{Iiu}(k), \xi_{iv}(k) = \xi_{Iiv}(k), r(k) = r_I(k) \quad \text{(IIR filter)}
\end{aligned} \tag{5.10}$$

We can estimate the continuous system parameters by the following LS method:

$$\widehat{\theta} = \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\mathbf{z}^T(k)\right]^{-1} \cdot \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\xi_{0y}(k)\right] \tag{5.11}$$

Investigating the limiting behaviour of the LS estimator when the number of data tends to infinity leads to

$$\operatorname*{plim}_{N\to\infty} \widehat{\theta} = \theta + \operatorname*{plim}_{N\to\infty} NP(N)\left[\operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)r(k)\right] \tag{5.12}$$

where

$$P(N) = \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\mathbf{z}^T(k)\right]^{-1} \tag{5.13}$$

With straightforward calculations, we have

$$\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{k=k_S+1}^{k_S+N}\mathbf{z}(k)r(k) = \mathrm{E}\begin{bmatrix} -\xi_{1y}(k) \\ \vdots \\ -\xi_{ny}(k) \\ \xi_{1w}(k) \\ \vdots \\ \xi_{nw}(k) \end{bmatrix}\left\{\sum_{i=0}^{n} a_i\xi_{ie}(k) - \sum_{i=1}^{n} b_i\xi_{iv}(k)\right\}$$

$$= -\mathrm{E}\begin{bmatrix} \xi_{1y}(k)\sum_{i=0}^{n} a_i\xi_{ie}(k) \\ \vdots \\ \xi_{ny}(k)\sum_{i=0}^{n} a_i\xi_{ie}(k) \\ \xi_{1w}(k)\sum_{i=1}^{n} b_i\xi_{iv}(k) \\ \vdots \\ \xi_{nw}(k)\sum_{i=1}^{n} b_i\xi_{iv}(k) \end{bmatrix} \neq 0 \tag{5.14}$$

Therefore the LS estimator is asymptotically biased in general due to the effects of the input and output measurement noises.

Now we will try to express the result of equation (5.14) with the input-output noise variances σ_v^2 and σ_e^2. The FIR filtering approach is first investigated. Based on the above discussions, we have the following results through straightforward calculations:

$$\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{k=k_S+1}^{k_S+N}\mathbf{z}(k)r(k) = \mathrm{E}[\mathbf{z_F}(k)r_F(k)] = \mathrm{E}\begin{bmatrix} -\xi_{F1e}(k)\sum_{i=0}^{n} a_i\xi_{Fie}(k) \\ \vdots \\ -\xi_{Fne}(k)\sum_{i=0}^{n} a_i\xi_{Fie}(k) \\ -\xi_{F1v}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \\ \vdots \\ -\xi_{Fnv}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \end{bmatrix}$$

$$
= -\mathrm{E}\begin{bmatrix} \xi_{F1e}(k)\left[\xi_{F0e}(k)+\sum_{i=1}^{n} a_i\xi_{Fie}(k)\right] \\ \vdots \\ \xi_{Fne}(k)\left[\xi_{F0e}(k)+\sum_{i=1}^{n} a_i\xi_{Fie}(k)\right] \\ \xi_{F1v}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \\ \vdots \\ \xi_{Fnv}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \end{bmatrix}
$$

$$
= -\mathrm{E}\begin{bmatrix} \xi_{F1e}(k)\xi_{F0e}(k) \\ \vdots \\ \xi_{Fne}(k)\xi_{F0e}(k) \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \mathrm{E}\begin{bmatrix} \xi_{F1e}(k)\sum_{i=1}^{n} a_i\xi_{Fie}(k) \\ \vdots \\ \xi_{Fne}(k)\sum_{i=1}^{n} a_i\xi_{Fie}(k) \\ \xi_{F1v}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \\ \vdots \\ \xi_{Fnv}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \end{bmatrix}
$$

$$
= -\sigma_e^2\begin{bmatrix} \sum_{j=0}^{M_F+n} f_j^1 f_j^0 \\ \vdots \\ \sum_{j=0}^{M_F+n} f_j^n f_j^0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \mathrm{E}\begin{bmatrix} \xi_{F1e}(k) \\ \vdots \\ \xi_{Fne}(k) \\ \xi_{F1v}(k) \\ \vdots \\ \xi_{Fnv}(k) \end{bmatrix} [\xi_{F1e}(k),\cdots,\xi_{Fne}(k),\ \xi_{F1v}(k),\cdots,\xi_{Fnv}(k)] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}
$$

$$
= -\mathbf{H}_0\sigma_e^2 - \mathbf{H}_F\mathbf{D}\ \boldsymbol{\theta} \tag{5.15}
$$

where

$$\mathbf{F} = \begin{bmatrix} f_0^1 & f_1^1 & \cdots & f_{M_F+n}^1 \\ f_0^2 & f_1^2 & \cdots & f_{M_F+n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ f_0^n & f_1^n & \cdots & f_{M_F+n}^n \end{bmatrix}, \quad \mathbf{H}_F = \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{F}\mathbf{F}^T \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \sigma_e^2 \mathbf{I}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \sigma_v^2 \mathbf{I}_{n\times n} \end{bmatrix}, \quad \mathbf{H}_0 = \begin{bmatrix} \sum_{j=0}^{M_F+n} f_j^1 f_j^0 \\ \vdots \\ \sum_{j=0}^{M_F+n} f_j^n f_j^0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.16}$$

and $\mathbf{I}_{n\times n}$ is an $n \times n$ identity matrix, $\mathbf{0}_{n\times n}$ is an $n \times n$ zero matrix.

For the IIR filters, unfortunately, it is uneasy to obtain the above results directly, since, usually calculations of the correlations of the outputs of the IIR filters are not straightforward. However, we can have similar results if $Q_I(z^{-1})$ is approximated by an FIR filter $Q_{FI}(z^{-1})$ with a sufficiently large length M_F:

$$Q_{FI}(z^{-1}) = \sum_{m=0}^{M_F} q_m z^{-m} \approx Q_I(z^{-1}) = \frac{\left(\frac{T}{2}\right)^{m-n} (1+z^{-1})^{m-n}}{\sum_{i=0}^{m} r_i z^{-i}} \tag{5.17}$$

Without loss of generality, the coefficients q_m and the order M_F of $Q_{FI}(z^{-1})$ are denoted as the same of those of $Q_F(z^{-1})$ defined in equation (3.22) for convenience of notation. If the length M_F is sufficiently large and hence the frequency response of $Q_{FI}(z^{-1})$ is compatiable to that of $Q_I(z^{-1})$, then the approximation error is neglectable compared to the stochastic noise sources.

Hence we have the following approximations:

$$
\begin{aligned}
\xi_{Iie}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}e(k) \\
&\approx Q_{FI}(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}e(k) \\
&= \sum_{j=0}^{M_F+n} f_j^i z^{-j}e(k) \\
\xi_{Iiv}(k) &= Q_I(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}v(k) \\
&\approx Q_{FI}(z^{-1})\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}v(k) \\
&= \sum_{j=0}^{M_F+n} f_j^i z^{-j}v(k) \\
r_I(k) &= \sum_{i=0}^{n} a_i\xi_{Iie}(k) - \sum_{i=1}^{n} b_i\xi_{Iiv}(k) \approx \sum_{j=0}^{M_F+n}\alpha_j z^{-j}e(k) - \sum_{j=0}^{M_F+n}\beta_j z^{-j}v(k) \\
&\alpha_j = f_j^0 + \sum_{i=1}^{n} f_j^i a_i, \quad \beta_j = \sum_{i=1}^{n} f_j^i b_i
\end{aligned}
\tag{5.18}
$$

Then we have similar results for the IIR filters:

$$
\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{k=k_S+1}^{k_S+N} \mathrm{z}(k)r(k) = \mathrm{E}[\mathrm{z_I}(k)r_I(k)] = \mathrm{E}\begin{bmatrix} -\xi_{I1e}(k)\sum_{i=0}^{n} a_i\xi_{Fie}(k) \\ \vdots \\ -\xi_{Ine}(k)\sum_{i=0}^{n} a_i\xi_{Fie}(k) \\ -\xi_{I1v}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \\ \vdots \\ -\xi_{Inv}(k)\sum_{i=1}^{n} b_i\xi_{Fiv}(k) \end{bmatrix}
\tag{5.19}
$$

$$\approx -\mathrm{H}_0\sigma_e^2 - \mathrm{H}_F\mathrm{D}\,\theta$$

5.5 BCLS method

From equations (5.12) and (5.15) or (5.19), we have

$$
\theta = \operatorname*{plim}_{N\to\infty}\hat{\theta} + \operatorname*{plim}_{N\to\infty} NP(N)\mathbf{H}_0\sigma_e^2 + \operatorname*{plim}_{N\to\infty} NP(N)\mathbf{H}_F\mathbf{D}\,\theta
\tag{5.20}
$$

which implies that an unbiased estimate of the unknown parameters can be obtained by subtracting an estimate of the bias.

Then the BCLS method is expressed as

$$\widehat{\theta}_{BCLS}(N) \;=\; \widehat{\theta}(N) + \mathbf{M}^{-1}\mathbf{H}_0\sigma_e^2 + \mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}\,\widehat{\theta}_{BCLS}(N-1) \tag{5.21}$$

where

$$\mathbf{M}^{-1} = \operatorname*{plim}_{N\to\infty} NP(N) \;=\; \{\mathrm{E}[\mathbf{z}(k)\mathbf{z}^T(k)]\}^{-1} = \{\mathbf{M_1}+\mathbf{M_2}\}^{-1} \tag{5.22}$$

and

$$\begin{aligned}\mathbf{M_1} &= \mathrm{E}\begin{bmatrix} -\mathbf{z}_x(k) \\ \mathbf{z}_u(k)\end{bmatrix}[-\mathbf{z}_x^T(k), \mathbf{z}_u^T(k)] \\ \mathbf{M_2} &= \mathrm{E}\begin{bmatrix} -\mathbf{z}_e(k) \\ \mathbf{z}_v(k)\end{bmatrix}[-\mathbf{z}_e^T(k), \mathbf{z}_v^T(k)] = \mathbf{H}_F\mathbf{D}\end{aligned} \tag{5.23}$$

It is necessary to investigate whether this algorithm is stable. When the spectral radius of the matrix $\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}$ is less than unity, the algorithm is stable and will converge (Sagara and Wada 1977). Here the spectral radius of $\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}$, $\varrho[\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}]$ is defined as

$$\varrho[\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}] \;=\; \max_{1\le i\le 2n}\left|\lambda_i[\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}]\right| \tag{5.24}$$

where $\lambda_i[\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}]$ denotes an eigen value of $\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}$.

According to **Lemma 2.1** in Stoica and Söderström (1982), if $\mathbf{M}, \mathbf{M_1}$ are positive definite, and $\mathbf{M_2}$ is non-negative definite, then $\varrho[\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}]$ is less than unity, and hence the algorithm is stable and converge (Sagara and Wada 1977).

It is not difficult to understand the fact that if the the input signal is persistently exciting of sufficient orders, $\mathbf{M_1}$ may be positive definite. And it can be shown that $\mathbf{M_2}$ is always non-negative definite.

To take more detailed discussions, we first state **Lemma 5.1**:

Lemma 5.1 Define

$$\left(\frac{T}{2}\right)^m (1+z^{-1})^m(1-z^{-1})^{m-i} = \sum_{j=0}^{m}\widetilde{g}_j^i z^{-j}, \qquad i = 1,\cdots,m \tag{5.25}$$

and the matrix $\widetilde{\mathbf{G}}_{m\times m+1}$

$$\widetilde{\mathbf{G}}_{m\times m+1} = \begin{bmatrix} \widetilde{g}_0^1 & \widetilde{g}_1^1 & \cdots & \widetilde{g}_m^1 \\ \widetilde{g}_0^2 & \widetilde{g}_1^2 & \cdots & \widetilde{g}_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{g}_0^m & \widetilde{g}_1^m & \cdots & \widetilde{g}_m^m \end{bmatrix} \tag{5.26}$$

Then the matrix $\widetilde{\mathbf{G}}_{m\times m+1}$ has full rank, i.e. rank$\widetilde{\mathbf{G}}_{m\times m+1} = m$.

Proof

For an arbitrary z^{-1}, let

$$\sum_{i=1}^{m} c_i \left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)^{m-i} = 0 \tag{5.27}$$

Since $(1-z^{-1})$ and $(1+z^{-1})$ are coprime and hence $\{2/T(1-z^{-1})/(1+z^{-1})\}^{m-i}$ $(i = 1, \cdots, m)$ are linearly independent for an arbitrary z^{-1}, then the above equation holds only when $c_i = 0 (i = 1, \cdots, m)$.

Rewrite equation (5.27) as

$$\sum_{i=1}^{m} c_i \left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)^{m-i} = \sum_{i=1}^{m} c_i(\tilde{g}_0^i + \tilde{g}_1^i z^{-1} + \cdots + \tilde{g}_m^i z^{-m}) = 0 \tag{5.28}$$

Since equation (5.28) holds only when $c_i = 0$, then the elements of

$$\begin{bmatrix} \tilde{g}_0^1 + \tilde{g}_1^1 z^{-1} + \ldots + \tilde{g}_m^1 z^{-m} \\ \tilde{g}_0^2 + \tilde{g}_1^2 z^{-1} + \ldots + \tilde{g}_m^2 z^{-m} \\ \vdots \\ \tilde{g}_0^m + \tilde{g}_1^m z^{-1} + \ldots + \tilde{g}_m^m z^{-m} \end{bmatrix} \tag{5.29}$$

are linearly independent.

Moreover, since

$$\begin{bmatrix} \tilde{g}_0^1 + \tilde{g}_1^1 z^{-1} + \ldots + \tilde{g}_m^1 z^{-m} \\ \tilde{g}_0^2 + \tilde{g}_1^2 z^{-1} + \ldots + \tilde{g}_m^2 z^{-m} \\ \vdots \\ \tilde{g}_0^m + \tilde{g}_1^m z^{-1} + \ldots + \tilde{g}_m^m z^{-m} \end{bmatrix} = \begin{bmatrix} \tilde{g}_0^1 & \tilde{g}_1^1 & \ldots & \tilde{g}_m^1 \\ \tilde{g}_0^2 & \tilde{g}_1^2 & \ldots & \tilde{g}_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{g}_0^m & \tilde{g}_1^m & \ldots & \tilde{g}_m^m \end{bmatrix} \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{-m} \end{bmatrix} \tag{5.30}$$

and the elements of $[1, z, \cdots, z^{-m}]^T$ are linearly independent, then

$$\widetilde{\mathbf{G}}_{m\times m+1} = \begin{bmatrix} \tilde{g}_0^1 & \tilde{g}_1^1 & \ldots & \tilde{g}_m^1 \\ \tilde{g}_0^2 & \tilde{g}_1^2 & \ldots & \tilde{g}_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{g}_0^m & \tilde{g}_1^m & \ldots & \tilde{g}_m^m \end{bmatrix}$$

has full rank.

Notice that from equation (3.4), the noise free output can be generated approximatedly by

$$x(k) = \frac{B'(z^{-1})}{A'(z^{-1})} u(k) \tag{5.31}$$

and hence

$$\begin{aligned}
\xi_{ix}(k) &= Q(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}x(k) \\
&\approx Q(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}\frac{B'(z^{-1})}{A'(z^{-1})}u(k) \\
\xi_{iu}(k) &= Q(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}u(k) \\
&= Q(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}\frac{A'(z^{-1})}{A'(z^{-1})}u(k)
\end{aligned} \tag{5.32}$$

where

$$\begin{aligned}
&Q(z^{-1}) = Q_F(z^{-1}) \quad \text{(FIR filter)} \\
&Q(z^{-1}) = Q_I(z^{-1}) \quad \text{(IIR filter)}
\end{aligned} \tag{5.33}$$

Then through straightforward calculation, we have

$$\begin{aligned}
\begin{bmatrix} -\mathrm{z}_x(k) \\ \mathrm{z}_u(k) \end{bmatrix} &= \frac{Q(z^{-1})}{A'(z^{-1})}
\begin{bmatrix}
-\left(\frac{T}{2}\right)^1 (1+z^{-1})^1(1-z^{-1})^{n-1}B'(z^{-1})u(k) \\
\vdots \\
-\left(\frac{T}{2}\right)^n (1+z^{-1})^n(1-z^{-1})^{n-n}B'(z^{-1})u(k) \\
\left(\frac{T}{2}\right)^1 (1+z^{-1})^1(1-z^{-1})^{n-1}A'(z^{-1})u(k) \\
\vdots \\
\left(\frac{T}{2}\right)^n (1+z^{-1})^n(1-z^{-1})^{n-n}A'(z^{-1})u(k)
\end{bmatrix} \\
&= \mathcal{R}^T(-B,A)\frac{Q(z^{-1})}{A'(z^{-1})}
\begin{bmatrix}
\left(\frac{T}{2}\right)^1 (1+z^{-1})^1(1-z^{-1})^{2n-1}u(k) \\
\vdots \\
\left(\frac{T}{2}\right)^n (1+z^{-1})^n(1-z^{-1})^{2n-n}u(k) \\
\left(\frac{T}{2}\right)^{n+1} (1+z^{-1})^{n+1}(1-z^{-1})^{2n-n-1}u(k) \\
\vdots \\
\left(\frac{T}{2}\right)^{2n} (1+z^{-1})^{2n}(1-z^{-1})^{2n-2n}u(k)
\end{bmatrix} \\
&= \mathcal{R}^T(-B,A)\widetilde{\mathrm{G}}_{2n\times 2n+1}\frac{Q(z^{-1})}{A'(z^{-1})}
\begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-2n) \end{bmatrix}
\end{aligned} \tag{5.34}$$

where

$$\mathcal{R}(-B,A)=\left[\begin{array}{cccc|cccc} 0 & & & & 1 & & & \\ -b_1 & \mathbf{0} & & & a_1 & \mathbf{0} & & \\ \vdots & \ddots & & & \vdots & \ddots & & \\ -b_n & & -b_1 & & a_n & & a_1 & \\ \mathbf{0} & \ddots & \vdots & & \mathbf{0} & \ddots & \vdots & \\ & & -b_n & & & & a_n & \end{array}\right] \tag{5.35}$$

Consequently, we have

$$\begin{aligned} \mathbf{M_1} &= \mathrm{E}\begin{bmatrix} -\mathbf{z}_x(k) \\ \mathbf{z}_u(k) \end{bmatrix} [-\mathbf{z}_x^T(k), \mathbf{z}_u^T(k)] \\ &= \mathcal{R}^T(-B,A)\widetilde{\mathbf{G}}_{2n\times 2n+1}\mathcal{Q}(A',Q,u)\widetilde{\mathbf{G}}^T_{2n\times 2n+1}\mathcal{R}(-B,A) \end{aligned} \tag{5.36}$$

where

$$\mathcal{Q}(A',Q,u)=\mathrm{E}\,\frac{Q(z^{-1})}{A'(z^{-1})}\begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-2n) \end{bmatrix}\frac{Q(z^{-1})}{A'(z^{-1})}[u(k),u(k-1),\cdots,u(k-2n)] \tag{5.37}$$

The determinant of $\mathcal{R}(-B,A)$ is called a resultant. It is known that $\mathcal{R}(-B,A)$ is nonsingular if and only if $A(p)$, $B(p)$ are relatively prime (Söderström and Stoica 1981). **Lemma 5.1** implies that the matrix $\widetilde{\mathbf{G}}_{m\times m+1}$ has full rank; since $\mathbf{M_1}$ is symmetric, and $\mathcal{Q}(A',Q,u)$ have full rank if the input signal is persistently exciting of sufficient orders (Söderström and Stoica 1981, 1989), then it is clear that $\mathbf{M_1}$ is nonsingular and positive definite. Hence, we have

Lemma 5.2 Assume that $u(k)$ is persistently exciting of sufficient orders, then $\mathbf{M_1}$ is positive definite.

Similarly, we have

$$\begin{aligned} \mathbf{M_2} &= \mathrm{E}\begin{bmatrix} -\mathbf{z}_e(k) \\ \mathbf{z}_v(k) \end{bmatrix} [-\mathbf{z}_e^T(k), \mathbf{z}_v^T(k)] \\ &= \begin{bmatrix} \widetilde{\mathbf{G}}_{n\times n+1} & \mathbf{0}_{n\times n+1} \\ \mathbf{0}_{n\times n+1} & \widetilde{\mathbf{G}}_{n\times n+1} \end{bmatrix}\Omega(Q,e,v)\begin{bmatrix} \widetilde{\mathbf{G}}^T_{n\times n+1} & \mathbf{0}_{n+1\times n} \\ \mathbf{0}_{n+1\times n} & \widetilde{\mathbf{G}}^T_{n\times n+1} \end{bmatrix} \end{aligned} \tag{5.38}$$

where

$$\Omega(Q,e,v) = \mathrm{E}\, Q(z^{-1}) \begin{bmatrix} e(k) \\ e(k-1) \\ \vdots \\ e(k-n) \\ v(k) \\ v(k-1) \\ \vdots \\ v(k-n) \end{bmatrix} Q(z^{-1})[e(k), e(k-1), \cdots, e(k-n), v(k), v(k-1), \cdots, v(k-n)] \tag{5.39}$$

Since the white noises are persistently exciting of all orders, we have the following result:

Lemma 5.3

$$\mathbf{M_2} = \mathrm{E}\begin{bmatrix} -\mathbf{z}_e(k) \\ \mathbf{z}_v(k) \end{bmatrix} [-\mathbf{z}_e^T(k), \mathbf{z}_v^T(k)] = \mathbf{H}_F\mathbf{D}$$

is always non-negative definite, for any $\sigma_e^2 \geq 0$ and $\sigma_v^2 \geq 0$.

It is clear that the result of the following Lemma holds:

Lemma 5.4

$$\mathbf{M} = \{\mathrm{E}[\mathbf{z}(k)\mathbf{z}^T(k)]\}^{-1} = \mathbf{M_1} + \mathbf{M_2}$$

is positive definite if the input signal is sufficiently persistently exciting so that $\mathbf{M_1}$ is positive definite.

The above results and **Lemma 2.1** of Stoica and Söderström (1982) imply that the spectral radius of the matrix $\mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}$ is less than unity, if the input signal is sufficiently persistently exciting so that $\mathbf{M_1}$ and $\mathbf{M}$ are positive definite. Hence we have

Theorem 5.1 The BCLS method

$$\hat{\theta}_{BCLS}(N) \;=\; \hat{\theta}(N) + NP(N)\mathbf{H}_0\sigma_e^2 + \mathbf{M}^{-1}\mathbf{H}_F\mathbf{D}\,\hat{\theta}_{BCLS}(N-1)$$

is stable and will converge, if the input signal is sufficiently persistently exciting.

However, in most practical situations, it is difficult to have *a priori* knowledge of σ_e^2 and σ_v^2, the practical applicability of the BCLS method is restricted within narrow limits. If $\hat{\sigma}_e^2$ and $\hat{\sigma}_v^2$ are the consistent estimates of σ_e^2 and σ_v^2, the BCLS method becomes

$$\hat{\theta}_{BCLS}(N) \;=\; \hat{\theta}(N) + NP(N)\mathbf{H}_0\hat{\sigma}_e^2 + \mathbf{M}^{-1}\mathbf{H}_F\hat{\mathbf{D}}\,\hat{\theta}_{BCLS}(N-1) \tag{5.40}$$

Remark 5.1: In the identification process, if $\varrho[\mathbf{M}^{-1}\mathbf{H}_F\widehat{\mathbf{D}}]$ is less than unity, the algorithm would also converge. If we monitor the value of $\varrho[\mathbf{M}^{-1}\mathbf{H}_F\widehat{\mathbf{D}}]$ and keep it less than unity during the identification process, the algorithm will never diverge. Experience shows, however, it is not necessary in most cases to monitor the stability.

5.6 Estimation of σ_e^2 and σ_v^2

The residual $\widehat{r}(k)$ for the LS estimate $\widehat{\boldsymbol{\theta}}(N)$ is given as

$$\widehat{r}(k) = \xi_{0y}(k) - \mathbf{z}^T(k)\widehat{\boldsymbol{\theta}}(N) \tag{5.41}$$

Using equation (5.8), we have

$$\widehat{r}(k) = \mathbf{z}^T(k)\left[\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}(N)\right] + r(k) \tag{5.42}$$

From equation (5.41) and equation (5.42), we have

$$\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\widehat{r}(k) = \mathbf{0} \tag{5.43}$$

Using equation (5.42) and equation (5.43), we have the sum of squared residuals:

$$\begin{aligned} g(N) &= \sum_{k=k_S+1}^{k_S+N} \widehat{r}(k)\widehat{r}(k) \\ &= \sum_{k=k_S+1}^{k_S+N} r^2(k) + \sum_{k=k_S+1}^{k_S+N} \mathbf{z}^T(k)r(k)\left[\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}(N)\right] \end{aligned} \tag{5.44}$$

Since

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} r^2(k) = \mathrm{E}[r^2(k)] = \sum_{j=0}^{M_F+n} \alpha_j^2(\mathbf{a})\sigma_e^2 + \sum_{j=0}^{M_F+n} \beta_j^2(\mathbf{b})\sigma_v^2 \tag{5.45}$$

and

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \mathbf{z}^T(k)r(k) = \mathrm{E}[\mathbf{z}^T(k)r(k)] = -\mathbf{H}_0^T\sigma_e^2 - \boldsymbol{\theta}^T\mathbf{H}_F\mathbf{D} \tag{5.46}$$

then the following result can be obtained:

$$\begin{aligned} \operatorname*{plim}_{N\to\infty} \frac{g(N)}{N} &= \sum_{j=0}^{M_F+n} \alpha_j^2(\mathbf{a})\sigma_e^2 + \sum_{j=0}^{M_F+n} \beta_j^2(\mathbf{b})\sigma_v^2 - \mathbf{h}^e(\mathbf{a})[\mathbf{a} - \widehat{\mathbf{a}}(N)]\sigma_e^2 - \mathbf{h}^v(\mathbf{b})[\mathbf{b} - \widehat{\mathbf{b}}(N)]\sigma_v^2 \\ &= \left[\sum_{j=0}^{M_F+n} f_j^0\alpha_j(\mathbf{a}) + \mathbf{h}^e(\mathbf{a})\widehat{\mathbf{a}}(N)\right]\sigma_e^2 + \mathbf{h}^v(\mathbf{b})\widehat{\mathbf{b}}(N)\sigma_v^2 \end{aligned} \tag{5.47}$$

where

$$\mathbf{h}^e(\mathbf{a}) = [h_1^e(\mathbf{a}), \cdots, h_n^e(\mathbf{a})], \quad \mathbf{h}^v(\mathbf{b}) = [h_1^v(\mathbf{b}), \cdots, h_n^v(\mathbf{b})]$$

$$h_i^e(\mathbf{a}) = \sum_{j=0}^{M_F+n} f_j^i \alpha_j(\mathbf{a}), \quad h_i^v(\mathbf{b}) = \sum_{j=0}^{M_F+n} f_j^i \beta_j(\mathbf{b}) \tag{5.48}$$

where $\alpha_j(\mathbf{a})$ and $\beta_j(\mathbf{b})$ are defined in equation (5.5) or equation (5.18).

Similar to the above discussions, define $\widetilde{\boldsymbol{\theta}}(N)$ by

$$\widetilde{\boldsymbol{\theta}}(N) = \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\mathbf{z}^T(k-l)\right]^{-1} \cdot \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\xi_{0y}(k-l)\right] \tag{5.49}$$

where l is a natural number.

The residual for $\widetilde{\boldsymbol{\theta}}(N)$ is defined by

$$\widetilde{r}(k) = \xi_{0y}(k) - \mathbf{z}^T(k)\widetilde{\boldsymbol{\theta}}(N) \tag{5.50}$$

and it can also be shown that

$$\widetilde{r}(k) = \mathbf{z}^T(k)[\boldsymbol{\theta} - \widetilde{\boldsymbol{\theta}}(N)] + r(k) \tag{5.51}$$

and

$$\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\widetilde{r}(k-l) = \mathbf{0} \tag{5.52}$$

Hence we have

$$\begin{aligned} f(N) &= \sum_{k=k_S+1}^{k_S+N} \widetilde{r}(k)\widetilde{r}(k-l) \\ &= \sum_{k=k_S+1}^{k_S+N} r(k)r(k-l) + \sum_{k=k_S+1}^{k_S+N} \mathbf{z}^T(k-l)r(k)\left[\boldsymbol{\theta} - \widetilde{\boldsymbol{\theta}}(N)\right] \end{aligned} \tag{5.53}$$

and thus

$$\begin{aligned} \operatorname*{plim}_{N\to\infty} \frac{f(N)}{N} &= \sum_{j=0}^{M_F+n-l} \alpha_j(\mathbf{a})\alpha_{j+l}(\mathbf{a})\sigma_e^2 + \sum_{j=0}^{M_F+n-l} \beta_j(\mathbf{b})\beta_{j+l}(\mathbf{b})\sigma_v^2 \\ &\quad -\widetilde{\mathbf{h}}^e(\mathbf{a})[\mathbf{a} - \widetilde{\mathbf{a}}(N)]\sigma_e^2 - \widetilde{\mathbf{h}}^v(\mathbf{b})[\mathbf{b} - \widetilde{\mathbf{b}}(N)]\sigma_v^2 \\ &= \left[\sum_{j=0}^{M_F+n-l} f_j^0 \alpha_{j+l}(\mathbf{a}) + \widetilde{\mathbf{h}}^e(\mathbf{a})\widetilde{\mathbf{a}}(N)\right]\sigma_e^2 + \widetilde{\mathbf{h}}^v(\mathbf{b})\widetilde{\mathbf{b}}(N)\sigma_v^2 \end{aligned} \tag{5.54}$$

where

$$\widetilde{\mathbf{h}}^e(\mathbf{a}) = [\widetilde{h}_1^e(\mathbf{a}), \cdots, \widetilde{h}_n^e(\mathbf{a})], \quad \widetilde{\mathbf{h}}^v(\mathbf{b}) = [\widetilde{h}_1^v(\mathbf{b}), \cdots, \widetilde{h}_n^v(\mathbf{b})]$$

$$\widetilde{h}_i^e(\mathbf{a}) = \sum_{j=0}^{M_F+n-l} f_j^i \alpha_{j+l}(\mathbf{a}), \quad \widetilde{h}_i^v(\mathbf{b}) = \sum_{j=0}^{M_F+n-l} f_j^i \beta_{j+l}(\mathbf{b}) \tag{5.55}$$

Then the estimates of the unknown variances σ_e^2 and σ_v^2 are given by the solution of the following simultaneous equation:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \sigma_e^2 \\ \sigma_v^2 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \sigma_e^2 \\ \sigma_v^2 \end{bmatrix} = \frac{1}{N} \begin{bmatrix} g(N) \\ f(N) \end{bmatrix} \tag{5.56}$$

where

$$\begin{aligned} a_{11} &= \sum_{j=0}^{M_F+n} f_j^0 \alpha_j (\hat{\mathbf{a}}(N-1)) + \mathbf{h}^e (\hat{\mathbf{a}}_{BCLS}(N-1)) \hat{\mathbf{a}}(N) \\ a_{12} &= \mathbf{h}^v \left(\hat{\mathbf{b}}_{BCLS}(N-1)\right) \hat{\mathbf{b}}(N) \\ a_{21} &= \sum_{j=0}^{M_F+n-l} f_j^0 \alpha_{j+l} (\hat{\mathbf{a}}(N-1)) + \tilde{\mathbf{h}}^e (\hat{\mathbf{a}}_{BCLS}(N-1)) \tilde{\mathbf{a}}(N) \\ a_{22} &= \tilde{\mathbf{h}}^v \left(\hat{\mathbf{b}}_{BCLS}(N-1)\right) \tilde{\mathbf{b}}(N) \end{aligned} \tag{5.57}$$

Remark 5.2: It should be noted that the delay l should be chosen such that $\mathbf{A}$ and $[\sum_{k=k_S+1}^{k_S+N} \mathbf{z}(k)\mathbf{z}^T(k-l)]$ are nonsingular. To our experiences, the results are not so sensitive to l.

It is still necessary to calculate $g(N)$ and $f(N)$ for equation (5.56). It can be shown using the previous results that

$$\begin{aligned} &\sum_{k=k_S+1}^{k_S+N} \begin{bmatrix} -\xi_{0y}(k) \\ \mathbf{z}(k) \end{bmatrix} [-\xi_{0y}(k),\ \mathbf{z}^T(k)] \begin{bmatrix} 1 \\ \hat{\theta}(N) \end{bmatrix} \\ &= \sum_{k=k_S+1}^{k_S+N} \begin{bmatrix} -\xi_{0y}(k) \\ \mathbf{z}(k) \end{bmatrix} [-\hat{r}(k)] = \begin{bmatrix} g(N) \\ \mathbf{0} \end{bmatrix} \end{aligned} \tag{5.58}$$

and

$$\begin{aligned} &\sum_{k=k_S+1}^{k_S+N} \begin{bmatrix} -\xi_{0y}(k) \\ \mathbf{z}(k) \end{bmatrix} [-\xi_{0y}(k-l),\ \mathbf{z}^T(k-l)] \begin{bmatrix} 1 \\ \tilde{\theta}(N) \end{bmatrix} \\ &= \sum_{k=k_S+1}^{k_S+N} \begin{bmatrix} -\xi_{0y}(k) \\ \mathbf{z}(k) \end{bmatrix} [-\tilde{r}(k-l)] = \begin{bmatrix} f(N) \\ \mathbf{0} \end{bmatrix} \end{aligned} \tag{5.59}$$

Hence we can express $g(N)$, $f(N)$ as

$$\begin{aligned} g(N) &= \sum_{k=k_S+1}^{k_S+N} \xi_{0y}^2(k) - \left[\sum_{k=k_S+1}^{k_S+N} \xi_{0y}(k)\mathbf{z}^T(k)\right] \hat{\theta}(N) \\ f(N) &= \sum_{k=k_S+1}^{k_S+N} \xi_{0y}(k)\xi_{0y}(k-l) - \left[\sum_{k=k_S+1}^{k_S+N} \xi_{0y}(k)\mathbf{z}^T(k-l)\tilde{\theta}(N)\right] \end{aligned} \tag{5.60}$$

5.7 Implementation of the algorithm

This section describes the implementation techniques of the BCLS method. Both the off-line version and on-line version can be considered. However, usually, the recursive on-line identification algorithms have contrastively a small requirement of primary memory since only a modest amount of information is stored. Considering the practical applications, it is preferable to implement the BCLS method in an on-line manner.

The on-line recursive BCLS algorithm is described as follows:

1: Calculate the LS estimate $\widehat{\theta}(N)$ and the estimate $\widetilde{\theta}(N)$:

$$\begin{aligned}
\widehat{\theta}(N) &= \widehat{\theta}(N-1) + L(N)[\xi_{0y}(k_S+N) - \mathbf{z}^T(k_S+N)\widehat{\theta}(N-1)] \\
L(N) &= \frac{P(N-1)\mathbf{z}(k_S+N)}{\rho(N) + \mathbf{z}^T(k_S+N)P(N-1)\mathbf{z}(k_S+N)} \\
P(N) &= \frac{1}{\rho(N)}\left[P(N-1) - \frac{P(N-1)\mathbf{z}(k_S+N)\mathbf{z}^T(k_S+N)P(N-1)}{\rho(N) + \mathbf{z}^T(k_S+N)P(N-1)\mathbf{z}(k_S+N)}\right]
\end{aligned} \tag{5.61}$$

and

$$\begin{aligned}
\widetilde{\theta}(N) &= \widetilde{\theta}(N-1) + \widetilde{L}(N)[\xi_{0y}(k_S+N-L) - \mathbf{z}^T(k_S+N-L)\widetilde{\theta}(N-1)] \\
\widetilde{L}(N) &= \frac{\widetilde{P}(N-1)\mathbf{z}(k_S+N)}{\rho(N) + \mathbf{z}^T(k_S+N-L)\widetilde{P}(N-1)\mathbf{z}(k_S+N)} \\
\widetilde{P}(N) &= \frac{1}{\rho(N)}\left[\widetilde{P}(N-1) - \frac{\widetilde{P}(N-1)\mathbf{z}(k_S+N)\mathbf{z}^T(k_S+N-L)\widetilde{P}(N-1)}{\rho(N) + \mathbf{z}^T(k_S+N-L)\widetilde{P}(N-1)\mathbf{z}(k_S+N)}\right]
\end{aligned} \tag{5.62}$$

where $\rho(N)$ is the forgetting factor and is chosen to be

$$\rho(N) = (1-0.01)\rho(N-1) + 0.01, \qquad \rho(k_S) = 0.95 \tag{5.63}$$

2: Calculate $g(N)$ and $f(N)$:

$$\begin{aligned}
g(N) &= \sum_{k=k_S+1}^{k_S+N} \xi_{0y}^2(k) - \left[\sum_{k=k_S+1}^{k_S+N} \xi_{0y}(k)\mathbf{z}^T(k)\right]\widehat{\theta}(N) \\
f(N) &= \sum_{k=k_S+1}^{k_S+N} \xi_{0y}(k)\xi_{0y}(k-l) - \left[\sum_{k=k_S+1}^{k_S+N} \xi_{0y}(k)\mathbf{z}^T(k-l)\right]\widetilde{\theta}(N)
\end{aligned} \tag{5.64}$$

3: Calculate the variance estimates $\widehat{\sigma}_e^2(N)$ and $\widehat{\sigma}_v^2(N)$ for $N > N_0$:

$$\begin{bmatrix} \sigma_e^2(N) \\ \sigma_v^2(N) \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{N}\begin{bmatrix} g(N) \\ f(N) \end{bmatrix} \tag{5.65}$$

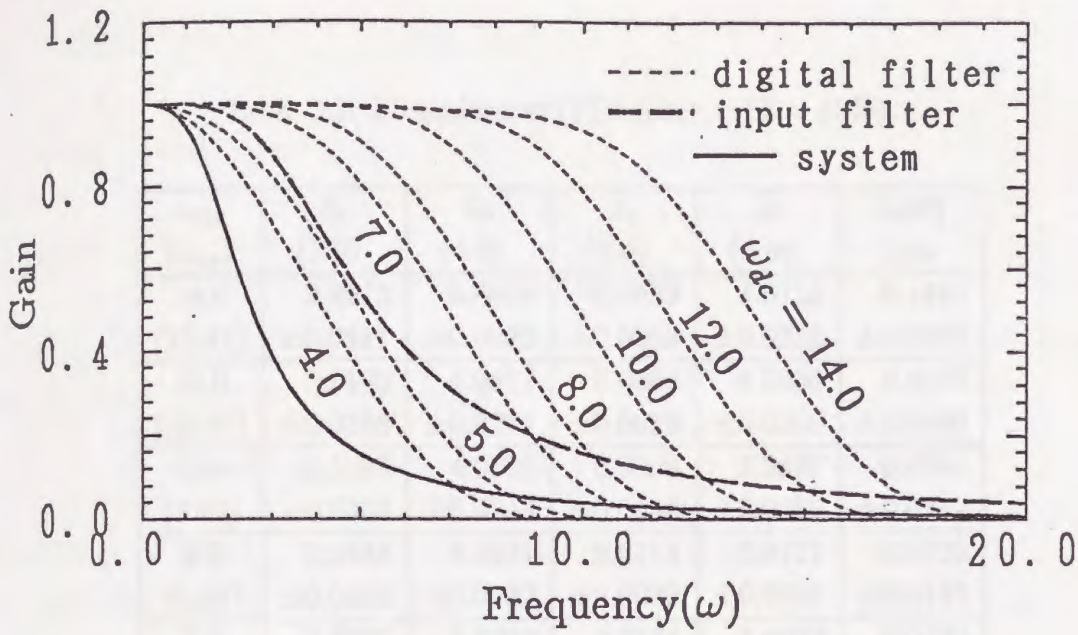


Figure 5.1: Frequency responses of the FIR filters for Table 5.1.

where

$$
\begin{aligned}
a_{11} &= \sum_{j=0}^{M_F+n} f_j^0 \alpha_j \left(\hat{\mathbf{a}}(N-1)\right) + \mathbf{h}^e \left(\hat{\mathbf{a}}_{BCLS}(N-1)\right) \hat{\mathbf{a}}(N) \\
a_{12} &= \mathbf{h}^v \left(\hat{\mathbf{b}}_{BCLS}(N-1)\right) \hat{\mathbf{b}}(N) \\
a_{21} &= \sum_{j=0}^{M_F+n-l} f_j^0 \alpha_{j+l} \left(\hat{\mathbf{a}}(N-1)\right) + \tilde{\mathbf{h}}^e \left(\hat{\mathbf{a}}_{BCLS}(N-1)\right) \tilde{\mathbf{a}}(N) \\
a_{22} &= \tilde{\mathbf{h}}^v \left(\hat{\mathbf{b}}_{BCLS}(N-1)\right) \tilde{\mathbf{b}}(N)
\end{aligned}
\qquad (5.66)
$$

4: Compensate the bias of the LS estimate $\widehat{\theta}(N)$:

$$
\begin{cases}
\widehat{\theta}_{BCLS}(N) = \widehat{\theta}(N) & N \le N_0 \\
\widehat{\theta}_{BCLS}(N) = \widehat{\theta}(N) + \mathbf{M}^{-1} \left[\mathbf{H}_0 \widehat{\sigma}_e^2(N-1) + \mathbf{H}_F \widehat{\mathbf{D}}(N-1)\, \widehat{\theta}_{BCLS}(N-1)\right] & N > N_0
\end{cases}
\qquad (5.67)
$$

5: Increase the recursion number N and return to 1 until convergence.

Remark 5.3: It should be remembered that all the key equations of the BCLS algorithm are derived by the assumption that sufficiently long data samples have been processed, i.e. N is sufficiently large. Hence it is often a practical policy to delay the start of the bias compensation procedure by $N_0 + 1$ recursions.

Table 5.1: LS estimates (FIR filter, NSR≈ 10%).

ω_{dc} (ω_{ac})	$\widehat{a}_1$ (3.0)	$\widehat{a}_2$ (4.0)	$\widehat{b}_1$ (0.0)	$\widehat{b}_2$ (4.0)	$\Delta\|\theta\|$ $\pm\sigma_\theta$
14.0 (12.47)	2.8915 ±0.0411	4.0890 ±0.0298	0.0089 ±0.0058	4.0118 ±0.0338	0.1410 ±0.0276
12.0 (10.57)	2.9180 ±0.0368	4.0517 ±0.0271	0.0090 ±0.0056	4.0038 ±0.0300	0.0973 ±0.0249
10.0 (8.62)	2.9298 ±0.0303	4.0196 ±0.0248	0.0099 ±0.0046	3.9887 ±0.0250	0.0743 ±0.0212
8.0 (6.69)	2.9343 ±0.0251	3.9931 ±0.0232	0.0113 ±0.0039	3.9721 ±0.0208	0.0725 ±0.0182
7.0 (5.69)	2.9339 ±0.0236	3.9803 ±0.0230	0.0125 ±0.0041	3.9629 ±0.0199	0.0792 ±0.0176
5.0 (3.78)	2.9199 ±0.0240	3.9449 ±0.0234	0.0184 ±0.0052	3.9326 ±0.0222	0.1196 ±0.0187
4.0 (3.01)	2.90176 ±0.0263	3.9143 ±0.0245	0.0248 ±0.0059	3.9038 ±0.0255	0.1633 ±0.0206

Table 5.2: LS estimates (IIR filter, NSR≈ 10%).

ω_c	$\widehat{a}_1$ (3.0)	$\widehat{a}_2$ (4.0)	$\widehat{b}_1$ (0.0)	$\widehat{b}_2$ (4.0)	$\Delta\|\theta\|$ $\pm\sigma_\theta$
10.0	2.8570 ±0.0391	4.0999 ±0.0290	0.0100 ±0.0058	3.9884 ±0.0297	0.1751 ±0.0259
8.0	2.9060 ±0.0348	4.0496 ±0.0273	0.0091 ±0.0051	3.9896 ±0.0260	0.1071 ±0.0233
6.0	2.9361 ±0.0296	4.0127 ±0.0254	0.0088 ±0.0043	3.9839 ±0.0223	0.0676 ±0.0204
4.0	2.9435 ±0.0242	3.9788 ±0.0231	0.0107 ±0.0036	3.9652 ±0.0189	0.0703 ±0.0174
3.0	2.9300 ±0.0242	3.9490 ±0.0223	0.0150 ±0.0036	3.9376 ±0.0198	0.1076 ±0.0175
2.0	2.8749 ±0.0295	3.8656 ±0.0241	0.0302 ±0.0049	3.8516 ±0.0265	0.2379 ±0.0212

5.8 Illustrative examples

Consider a second-order system described by

$$\begin{aligned} &\ddot{x}(t) + a_1\dot{x}(t) + a_2x(t) = b_1\dot{u}(t) + b_2u(t) \\ &a_1 = 3.0,\ a_2 = 4.0,\ b_1 = 0.0,\ b_2 = 4.0 \end{aligned} \tag{5.68}$$

The input $u(t)$ is the output of a second-order continuous-time Butterworth input filter driven by a stationary random signal $\zeta(t)$:

$$u(t) = L(p)\zeta(t) = \frac{1}{(p/\omega_c)^2 + \sqrt{2}(p/\omega_c) + 1}\,\zeta(t), \quad \omega_c = 4.0 \tag{5.69}$$

Simulation experiments are carried out when the sampling interval is taken to be $T = 0.04$, and in this case, $\sigma_u = 2.38, \sigma_x = 0.69$.

Example 5.1: Effects of the filter characteristics.

The LS estimates for the case of low noises where $\sigma_v = 0.24, \sigma_e = 0.07$ (NSR$\approx$ 10%) are shown in Table 5.1 for the FIR filters($M_F = 50$), and Table 5.2 for the IIR filters($m = 2$). The frequency responses of the system, the digital pre-filters used in Tables 5.1~5.2 and the input filter $L(p)$ in (5.69) are shown in Figures 5.1~5.2. It is clear that accurate estimates can be obtained if the pass-band of the pre-filters includes that of the low-pass input filter $L(p)$ in equation (5.69). Therefore, for the case of low input-output measurement noises, if the pass-band of the digital low-pass filters is chosen such that it *includes the main frequencies of the input signals which excite the system modes*, the noise effects are sufficiently reduced, and thus the LS estimates are still acceptable. For the case where only the output is corrupted by a measurement noise, it is known that the pass-band of the filters should be chosen such that it matches that of the system under study as closely. This suggestion is not appropriate in the presence of input measurement noise.

Example 5.2: Parameter estimates in the presence of high noises using the FIR filters.

The LS estimates and the BCLS estimates($l = 5$) for the FIR filters are shown in Tables 5.3~5.4, when $\sigma_v = 0.60, \sigma_e = 0.17$ (NSR $\approx$ 25%).

In the presence of high input-output measurement noises, it is difficult to obtain accurate estimates with the LS method. However, the BCLS method is very efficient in this case. Figure 5.3 plot the examples of the LS estimates and the BCLS estimates (when $\omega_{dc} = 7.0$). For short samples, it is clear that the BCLS estimates are slightly more sensitive than the LS

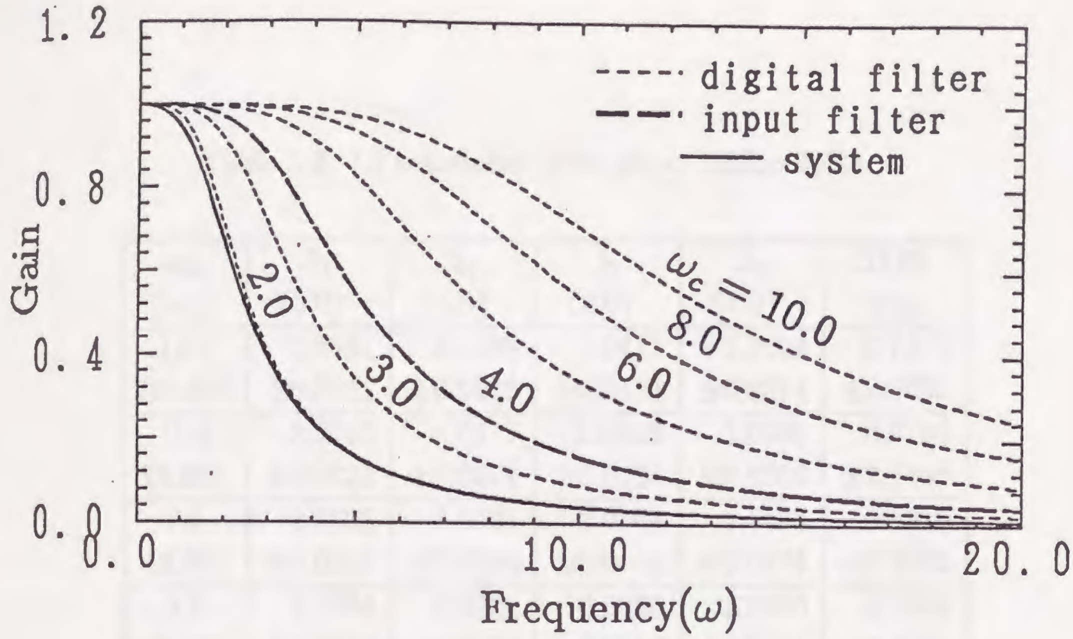


Figure 5.2: Frequency responses of the IIR filters for Table 5.2.

estimates. However, when sufficiently long samples are taken, the BCLS estimates converge to their true values in a very accurate manner.

Example 5.3: Parameter estimates in the presence of high noises using the IIR filters.

The LS estimates and the BCLS estimates ($L = 5$) for the IIR filters are shown in Tables 5.5~5.6 respectively, when $\sigma_v = 0.60, \sigma_e = 0.17$ (NSR $\approx 25\%$). And the examples of the LS estimates and the BCLS estimates (when $\omega_c = 6.0$) are shown in Figure 5.4. For the IIR filters, a third-order Butterworth filter

$$F_{I3}(p) = \frac{1}{(p/\omega_c)^3 + 2(p/\omega_c)^2 + 2\sqrt{2}(p/\omega_c) + 1}$$

is used, and $Q_I(z^{-1})$ is approximated by an FIR filter$Q_{FI}(z^{-1})$ with a sufficiently large length M_F for calculating the bias of the LS estimate. Although in most identification problems, a second-order filter is sufficient, it is found that when it is necessary to estimate the variances of both the input and output noises, for the second-order system under study, the third-order filter yields obviously better results than the second-order filter. The reason for this fact was mentioned in **Remark 3.2**. It should be pointed out that for a filter having slower transient response characteristics, a larger M_F is required to have accurate approximations. However, too large a M_F may require a great amount of computational burden.

Table 5.3: LS estimates (FIR filter, NSR≈ 25%).

ω_{dc}	$\widehat{a}_1$	$\widehat{a}_2$	$\widehat{b}_1$	$\widehat{b}_2$	$\Delta\|\theta\|$
(ω_{ac})	(3.0)	(4.0)	(0.0)	(4.0)	$\pm\sigma_\theta$
12.0	2.3631	4.2188	0.0411	3.7028	0.7372
(10.57)	±0.0721	±0.0625	±0.0131	±0.0674	±0.0538
10.0	2.4765	4.0417	0.0445	3.6969	0.6709
(8.62)	±0.0623	±0.0571	±0.0108	±0.0558	±0.0465
7.0	2.5698	3.8391	0.0564	3.6485	0.5810
(5.69)	±0.0510	±0.0534	±0.0099	±0.0426	±0.0392
5.0	2.5364	3.6724	0.0825	3.5350	0.7383
(3.78)	±0.0487	±0.0526	±0.0116	±0.0444	0.0393
4.0	2.4711	3.5447	0.1097	3.4195	0.9142
(3.01)	±0.0510	±0.0525	±0.0124	±0.0496	0.0414

Table 5.4: BCLS estimates (FIR filter, NSR≈ 25%).

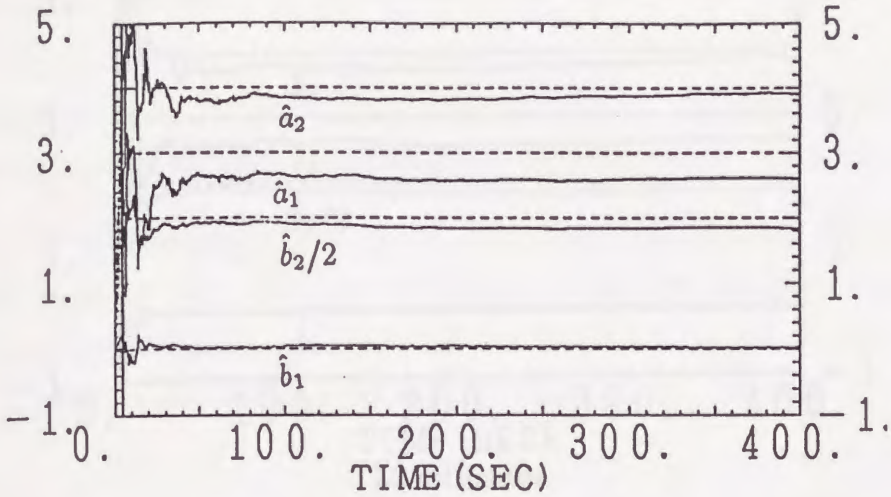
ω_{dc}	$\widehat{a}_1$	$\widehat{a}_2$	$\widehat{b}_1$	$\widehat{b}_2$	$\Delta\|\theta\|$	σ_e	σ_v
(ω_{ac})	(3.0)	(4.0)	(0.0)	(4.0)	$\pm\sigma_\theta$	(0.17)	(0.60)
12.0	3.0760	4.0295	0.0007	4.0936	0.1242	0.1699	0.5416
(10.57)	±0.0964	±0.0723	±0.0151	±0.0767	±0.0651	±0.0051	±0.1568
10.0	3.0521	4.0244	0.0020	4.0715	0.0918	0.1702	0.5773
(8.62)	±0.0774	±0.0665	±0.0124	±0.0609	±0.0543	±0.0047	±0.0828
7.0	3.0342	4.0233	0.0016	4.0527	0.0671	0.1706	0.5954
(5.69)	±0.0612	±0.0703	±0.0104	±0.0510	±0.0482	±0.0044	±0.0366
5.0	3.0372	4.0306	−0.0002	4.0571	0.0747	0.1707	0.6005
(3.78)	±0.0620	±0.0615	±0.0134	±0.0606	±0.0494	±0.0054	±0.0255
4.0	3.0393	4.0345	−0.0009	4.0611	0.0805	0.1703	0.6011
(3.01)	±0.0709	±0.0660	±0.0161	±0.0721	±0.0562	±0.0058	±0.0249

Table 5.5: LS estimates (IIR filter, NSR≈ 25%).

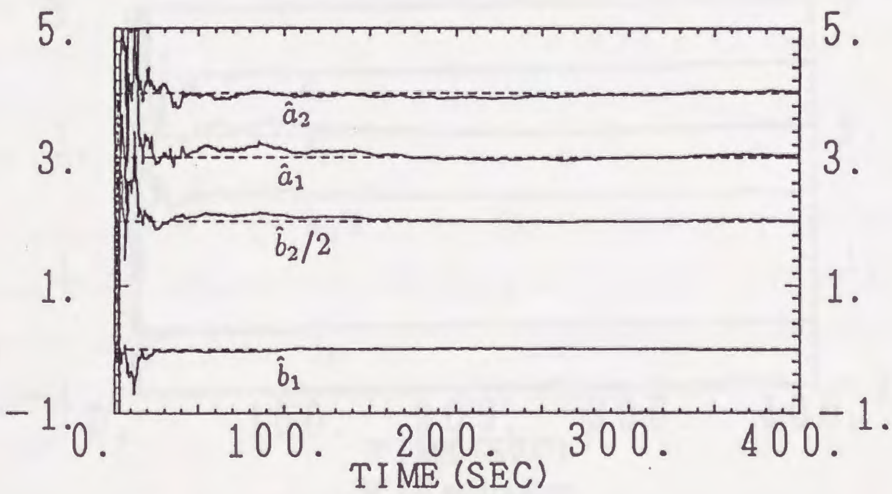
ω_c	$\widehat{a}_1$ (3.0)	$\widehat{a}_2$ (4.0)	$\widehat{b}_1$ (0.0)	$\widehat{b}_2$ (4.0)	$\Delta\|\theta\|$ $\pm\sigma_\theta$
10.0	2.3686 ±0.0652	4.2330 ±0.0639	0.0377 ±0.0109	3.6949 ±0.0602	0.7399 ±0.0505
8.0	2.5107 ±0.0604	4.0429 ±0.0598	0.0391 ±0.0096	3.7055 ±0.0526	0.5740 ±0.0456
6.0	2.5922 ±0.0522	3.8931 ±0.0559	0.0464 ±0.0088	3.6821 ±0.0419	0.5301 ±0.0397
4.0	2.5507 ±0.0450	3.6863 ±0.0509	0.0825 ±0.0085	3.5350 ±0.0387	0.7201 0.0358
3.0	2.3890 ±0.0520	3.4105 ±0.0506	0.1423 ±0.0103	3.2711 ±0.0501	1.1280 0.0408

Table 5.6: BCLS estimates (IIR filter, NSR≈ 25%).

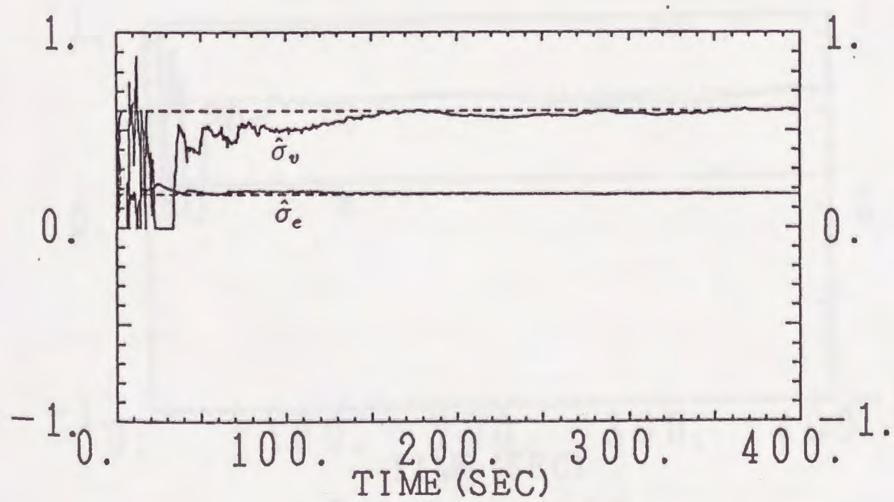
ω_c	$\widehat{a}_1$ (3.0)	$\widehat{a}_2$ (4.0)	$\widehat{b}_1$ (0.0)	$\widehat{b}_2$ (4.0)	$\Delta\|\theta\|$ $\pm\sigma_\theta$	σ_e (0.17)	σ_v (0.60)
10.0 ($M_F = 50$)	3.0872 ±0.0914	4.0320 ±0.0713	0.0004 ±0.0130	4.1076 ±0.0712	0.1421 ±0.0617	0.1697 ±0.0030	0.5815 ±0.1858
8.0 ($M_F = 50$)	3.0602 ±0.0741	4.0287 ±0.0654	0.0012 ±0.0112	4.0807 ±0.0555	0.1047 ±0.0512	0.1698 ±0.0029	0.5956 ±0.0919
6.0 ($M_F = 50$)	3.0419 ±0.0599	4.0264 ±0.0592	0.0011 ±0.0095	4.0610 ±0.0464	0.0786 ±0.0438	0.1699 ±0.0028	0.5999 ±0.0489
4.0 ($M_F = 50$)	3.0389 ±0.0552	4.0319 ±0.0571	−0.0002 ±0.0104	4.0602 ±0.0479	0.0784 ±0.0427	0.1678 ±0.0029	0.6049 ±0.0276
3.0 ($M_F = 50$)	3.0549 ±0.0756	4.0687 ±0.0691	−0.0077 ±0.0165	4.1059 ±0.0741	0.1379 ±0.0588	0.1345 ±0.0012	0.6716 ±0.0214
3.0 ($M_F = 80$)	3.0469 ±0.0754	4.0478 ±0.0692	−0.0036 ±0.0165	4.0788 ±0.0739	0.1035 ±0.0587	0.1429 ±0.0037	0.6238 ±0.0238
3.0 ($M_F = 100$)	3.0435 ±0.0753	4.0383 ±0.0692	−0.0018 ±0.0164	4.0660 ±0.0739	0.0883 ±0.0587	0.1708 ±0.0043	0.6010 ±0.0251
3.0 ($M_F = 120$)	3.0436 ±0.0753	4.0384 ±0.0692	−0.0018 ±0.0164	4.0661 ±0.0739	0.0884 ±0.0587	0.1705 ±0.0043	0.6012 ±0.0251



LS estimates



BCLS estimates



Estimates of σ_e and σ_v

Figure 5.3: Parameter estimates using the FIR filtering approach ($\omega_{dc} = 7.0$, NSR$\approx$ 25%).

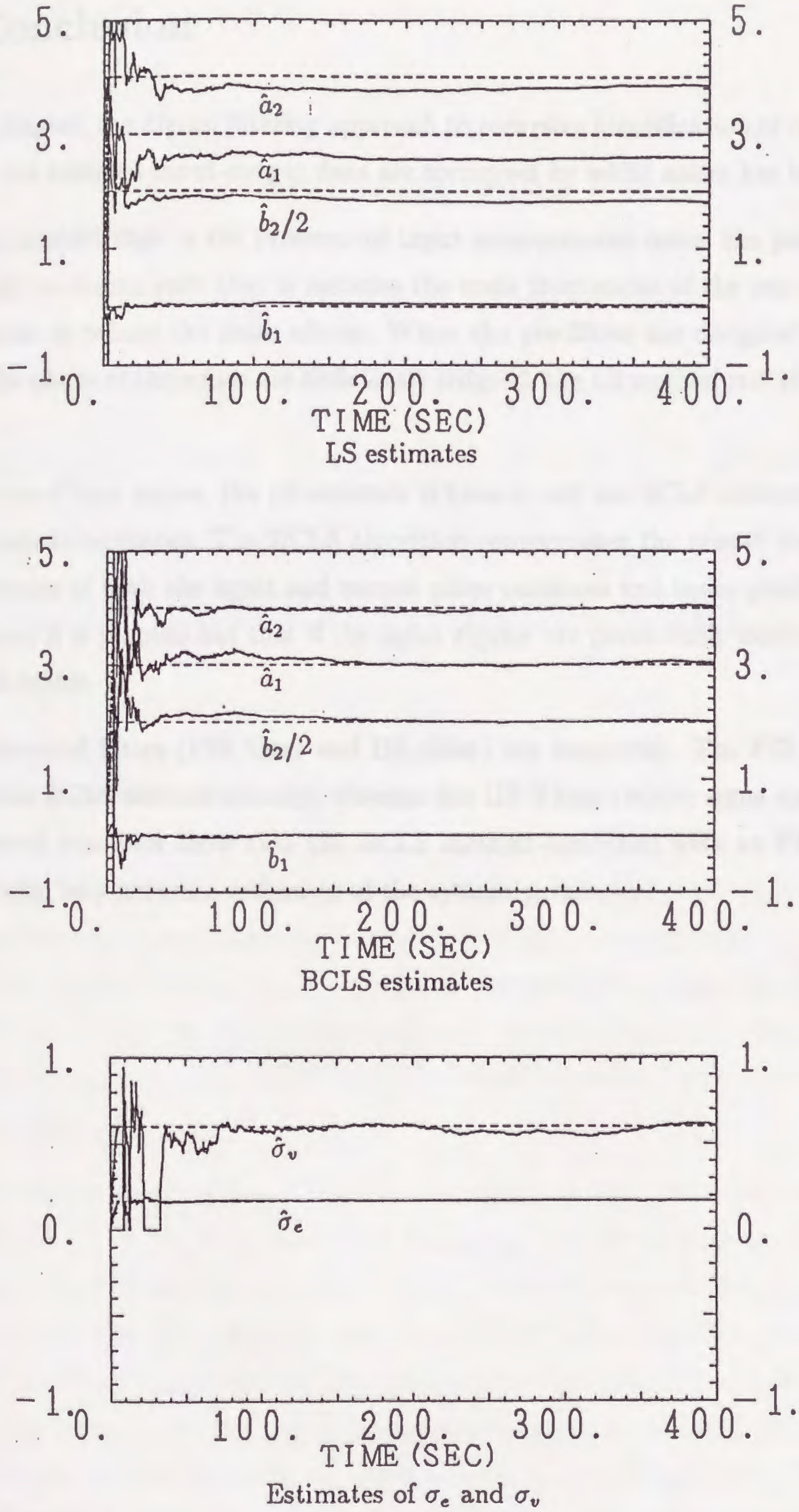


Figure 5.4: Parameter estimates using the IIR filtering approach ($\omega_c = 6.0$, NSR$\approx$ 25%).

5.9 Conclusion

In this chapter, the digital filtering approach to recursive identification of continuous systems when the sampled input-output data are corrupted by white noises has been discussed.

It is emphasized that in the presence of input measurement noise, the pass-band of the filters should be chosen such that it includes the main frequencies of the real system input-output signals to reduce the noise effects. When the pre-filters are designed appropriately such that the effects of the noises are sufficiently reduced, the LS method still gives acceptable results.

In the case of high noises, the LS estimate is biased, and the BCLS method is utilized to obtain consistent estimates. The BCLS algorithm compensates the bias of the LS estimate by the estimates of both the input and output noise variances and hence yields a consistent estimate. And it is pointed out that if the input signals are persistently exciting, the BCLS algorithm is stable.

Both classes of filters (FIR filter and IIR filter) are employed. The FIR filters can be applied to the BCLS method directly, whereas the IIR filters require some approximations. And numerical examples show that the BCLS method combined with an FIR filter or an IIR filter yields very accurate estimates of the system parameter.

Chapter 6

Identification in the Presence of Input-Output Measurement Noises Using Bias-Compensated Instrumental Variable Method

6.1 Introduction

In chapter 5, the BCLS method proposed by (Wada and Eguchi 1986) for common discrete-time systems was extended to the problem of continuous systems by making use of a digital low-pass pre-filter. The BCLS algorithm compensates the bias of the LS estimate by the estimates of both the input and output noise variances and hence yields a consistent estimate. This approach seems to be more convenient than the early works discussed by some other researchers.

In a lot of applications, the digitizing noise is one of the most important noise contribution, and in this case it might be reasonable to assume that the input measurement noise is white (Schoukens, Pintelon and Renneboog 1988). However, since the output measurement noise includes not only the digitizing noise, but also the noises created by the environment such as external disturbances and the modeling error in some practical situations, it is more appropriate to assume that the output measurement noise is coloured. Because of these considerations, in the present work, we discuss the problem of parameter identification of dynamic processes in the case where the discrete input and output measurements are corrupted by a white noise and by a noise which may be coloured respectively (Yang, Sagara and Wada 1991).

As mentioned in the previuous chapters, considering using digital computers, it is preferable to estimate the parameters of a continuous system using an approximated discrete-time estimation model with the continuous system parameters. In this chapter, the discrete-time estimation model combined with an adaptive digital IIR filter described in chapter 4 is utilized. It is obvious that for the estimation model derived by the adaptive pre-filter, the LS estimates are biased in general due to the effects of both the input and output noises. To avoid the effects of the output noise we introduce an IV method with filtered inputs and delayed filtered outputs as instrumental variables as suggested in chapter 4. The bias of the IV estimate due to the input noise is then analysed. By compensating the bias of the IV estimate with the estimated variance of the input noise, a bias compensated IV (BCIV) method is proposed to obtain a consistent estimate. The BCIV method treats only the variance of the input noise and is therefore more convenient than the BCLS method which requires estimating the variances of both the input and output noises. Approximated on-line implementation techniques of the proposed method are also described. Numerical examples show the results of the proposed BCIV mehtod are quite satisfactory.

6.2 Statement of the problem

Consider the following SISO continuous system

$$\begin{aligned} A(p)x(t) &= B(p)u(t) \\ A(p) &= \sum_{i=0}^{n} a_i p^{n-i} \quad (a_0 = 1) \\ B(p) &= \sum_{i=1}^{n} b_i p^{n-i} \end{aligned} \tag{6.1}$$

The sampled measurements of the input-output signals are described as

$$\begin{aligned} y(k) &= x(k) + C(z^{-1})e(k) \\ w(k) &= u(k) + v(k) \end{aligned} \tag{6.2}$$

where $C(z^{-1})$ denotes an MA process with known finite length r:

$$C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_r z^{-r} \tag{6.3}$$

$v(k)$ and $e(k)$ are white noises such that

$$\begin{gathered} \mathrm{E}[e(k)] = 0,\ \mathrm{E}[e(k)^2] = \sigma_e^2 \\ \mathrm{E}[v(k)] = 0,\ \mathrm{E}[v(k)^2] = \sigma_v^2 \\ \mathrm{E}[e(k)v(k)] = 0,\ \mathrm{E}[u(k)v(k)] = 0,\ \mathrm{E}[u(k)e(k)] = 0 \end{gathered} \tag{6.4}$$

Our goal is to identify the system parameters from the noisy sampled input-output data.

Using the results in chapter 4 and taking the noisy measurements into account, we have the following discrete-time estimation model:

$$\begin{gathered} \xi_{A0y}(k) + \sum_{i=1}^{n} a_i \xi_{Aiy}(k) = \sum_{i=1}^{n} b_i \xi_{Aiw}(k) + C(z^{-1})e(k) - \frac{B'(z^{-1})}{A'(z^{-1})}v(k) \\ \xi_{Aiw}(k) = Q_{IA}(z^{-1}) \left(\frac{T}{2}\right)^i (1+z^{-1})^i (1-z^{-1})^{n-i} w(k) \\ \xi_{Aiy}(k) = Q_{IA}(z^{-1}) \left(\frac{T}{2}\right)^i (1+z^{-1})^i (1-z^{-1})^{n-i} y(k) \end{gathered} \tag{6.5}$$

where

$$Q_{IA}(z^{-1}) = \frac{1}{A'(z^{-1})} \tag{6.6}$$

The low-pass pre-filter $1/A'(z^{-1})$ is introdued to attempt to filter off unwanted high frequency components of the noises and also to let the output noise remain in its original form in the equation error (Young and Jakeman 1980).

6.3 IV method and its bias

The discrete-time estimation model of equation (6.5) can be written in vector form as

$$\begin{aligned} \xi_{A0y}(k) &= \mathbf{z}^T(k)\boldsymbol{\theta} + r(k) \\ \mathbf{z}^T(k) &= [-\xi_{A1y}(k), \cdots, -\xi_{Any}(k), \xi_{A1w}(k), \cdots, \xi_{Anw}(k)] \\ &= [-\mathbf{z_y}^T(k), \mathbf{z_w}^T(k)] = [-\{\mathbf{z_x}^T(k) + \mathbf{z_e}^T(k)\}, \{\mathbf{z_u}^T(k) + \mathbf{z_v}^T(k)\}] \\ \boldsymbol{\theta}^T &= [a_1, \cdots, a_n, b_1, \cdots, b_n] = [\mathbf{a}^T, \mathbf{b}^T] \end{aligned} \tag{6.7}$$

where

$$\begin{aligned} r(k) &= C(z^{-1})e(k) - \frac{B'(z^{-1})}{A'(z^{-1})}v(k) \\ \mathbf{z_e}^T(k) &= [\xi_{1e}(k), \cdots, \xi_{ne}(k)] \\ \mathbf{z_x}^T(k) &= [\xi_{1x}(k), \cdots, \xi_{nx}(k)] \\ \mathbf{z_v}^T(k) &= [\xi_{1v}(k), \cdots, \xi_{nv}(k)] \\ \mathbf{z_u}^T(k) &= [\xi_{1u}(k), \cdots, \xi_{nu}(k)] \end{aligned} \tag{6.8}$$

and

$$\begin{aligned} \xi_{Aie}(k) &= Q_{IA}(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}C(z^{-1})e(k) \\ \xi_{Aiv}(k) &= Q_{IA}(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}v(k) \end{aligned} \tag{6.9}$$

It should be noted that owing to using the pre-filter $1/A'(z^{-1})$, the output measurement noise $C(z^{-1})e(k)$ remains in its original form in the equation error $r(k)$. This fact implies that it is possible to avoid the effects of the output noise by introducing an IV vector $\mathbf{m}^T(k)$ whose elements are highly correlated with the real system signals, but totally uncorrelated with the noise $C(z^{-1})e(k)$. A straightforward approach is to choose an IV vector with filtered inputs and delayed filtered outputs to avoid the correlations between $\xi_{Aiy}(k)(i = 1, \cdots, n)$ and $C(z^{-1})e(k)$ (Sagara, Yang and Wada 1991c):

$$\begin{aligned} \mathbf{m}^T(k) &= [-\xi_{A1y}(k-l), \cdots, -\xi_{Any}(k-l), \xi_{A1w}(k), \cdots, \xi_{Anw}(k)] \\ &= [-\mathbf{z_y}^T(k-l), \mathbf{z_w}^T(k)], \qquad l > r \end{aligned} \tag{6.10}$$

The motivation of the choice of such an IV vector stems from the IV-3 variant discussed by Söderström and Stoica (1981).

When the input signal is sufficiently rich so that

$$\left[\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)\mathbf{z}^T(k)\right]^{-1}$$

exists, we can estimate the continuous system parameters by the following IV method:

$$\widehat{\boldsymbol{\theta}}_{IV} = \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)\mathbf{z}^T(k)\right]^{-1} \cdot \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)\xi_{A0y}(k)\right] \tag{6.11}$$

It is necessary to analyse the limiting behavior of the IV estimate when the number of data tends to infinity. Using equation (6.7), it can be shown that

$$\operatorname*{plim}_{N\to\infty} \widehat{\boldsymbol{\theta}}_{IV} = \boldsymbol{\theta} + \operatorname*{plim}_{N\to\infty} NP(N)\left[\operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)r(k)\right] \tag{6.12}$$

where

$$P(N) = \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)\mathbf{z}^T(k)\right]^{-1} \tag{6.13}$$

Furthermore, we obtain the following result due to the fact that $\xi_{Aiw}(k)(i = 1, \cdots, n)$ are correlated with $r(k)$:

$$\begin{aligned}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{k=k_S+1}^{k_S+N}\mathbf{m}(k)r(k) &= \mathrm{E}\begin{bmatrix}-\xi_{A1y}(k-l)\\ \vdots\\ -\xi_{Any}(k-l)\\ \xi_{A1w}(k)\\ \vdots\\ \xi_{Anw}(k)\end{bmatrix}\left[C(z^{-1})e(k)-\frac{B'(z^{-1})}{A'(z^{-1})}v(k)\right]\\ &= \mathrm{E}\begin{bmatrix}0\\ \vdots\\ 0\\ \xi_{A1w}(k)r_v(k)\\ \vdots\\ \xi_{Anw}(k)r_v(k)\end{bmatrix}\neq 0\end{aligned} \tag{6.14}$$

where

$$r_v(k)=-\frac{B'(z^{-1})}{A'(z^{-1})}v(k) \tag{6.15}$$

Therefore the IV estimator is asymptotically biased in general due to the presence of the input noise.

Remark 6.1: Someone may suggest to choose an IV vector as

$$\mathbf{m}^T(k) = [-\bar{\xi}_{A1y}(k-l),\cdots,-\bar{\xi}_{Any}(k-l),\bar{\xi}_{A(n+1)y}(k-l),\cdots,\bar{\xi}_{A2ny}(k-l)], \qquad l>r \tag{6.16}$$

where

$$\bar{\xi}_{Aiy}(k)=\frac{1}{A'^2(z^{-1})}\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{2n-i}y(k) \tag{6.17}$$

Then he may conclude that there will be no asymptotic bias (and no need for a bias compensation described in the next section). It should be commented that this suggestion is not appropriate in general, since such a choice of $\mathbf{m}^T(k)$ may destroy the existence of $P(N)=[\sum_{k=k_S+1}^{k_S+N}\mathbf{m}(k)\mathbf{z}^T(k)]^{-1}$ (Söderström 1981, Söderström and Stoica 1981)

Now we will consider the method to express the result of equation (6.14) with the input noise variance σ_v^2. It is well-known that the pre-filter $1/A'(z^{-1})$ can be approximated by an FIR filter $Q_{FA}(z^{-1})$ with a sufficiently large length M_A:

$$Q_{FA}(z^{-1})=\sum_{m=0}^{M_A}s_m z^{-m}\approx\frac{1}{A'(z^{-1})}=\frac{1}{\sum_{i=0}^{n}\alpha_i z^{-i}} \tag{6.18}$$

where

$$\begin{cases} 1 = s_0\alpha_0 \\ 0 = s_m\alpha_0 + s_{m-1}\alpha_1 + s_{m-2}\alpha_2 + \ldots + s_{m-n}\alpha_n, \quad 1 \le m \le M_A \end{cases} \tag{6.19}$$

If the length M_A is sufficiently large and hence the frequency response of $Q_{FA}(z^{-1})$ is compatiable to that of $1/A'(z^{-1})$, then the approximation error is neglectable compared to the stochastic noise sources.

Hence we have the following approximations:

$$\begin{aligned} \xi_{Aiv}(k) &= \frac{\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i}}{A'(z^{-1})} v(k) \\ &\approx Q_{FA}(z^{-1})\left(\frac{T}{2}\right)^i (1+z^{-1})^i(1-z^{-1})^{n-i} v(k) = \sum_{j=0}^{M_A+n} f_j^i z^{-j} v(k) \end{aligned} \tag{6.20}$$

and

$$r_v(k) = -\frac{B'(z^{-1})}{A'(z^{-1})} v(k) = -\sum_{i=1}^{n} b_i \xi_{Aiv}(k) \approx -\sum_{j=0}^{M_A+n} \beta_j z^{-j} v(k) \tag{6.21}$$

where

$$\beta_j = \sum_{i=1}^{n} f_j^i b_i \tag{6.22}$$

Based on the above discussions, we have the following results through straightforward calculations:

$$\begin{aligned} \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k) r(k) = E[\mathbf{m}(k)r(k)] &= \mathrm{E} \begin{bmatrix} \mathbf{0}_{n\times 1} \\ -\xi_{A1v}(k) \sum_{i=1}^{n} b_i \xi_{Aiv}(k) \\ \vdots \\ -\xi_{Anv}(k) \sum_{i=1}^{n} b_i \xi_{Aiv}(k) \end{bmatrix} \\ &= -\mathrm{E} \begin{bmatrix} \mathbf{0}_{n\times 1} \\ \\ \mathbf{z}_v(k) \end{bmatrix} [\mathbf{0}_{1\times n},\ \mathbf{z}_v(k)] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \\ &= -\sigma_v^2 \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{F}\mathbf{F}^T \end{bmatrix} \theta \\ &= -\sigma_v^2 \mathbf{D}(\mathbf{F})\, \theta \end{aligned} \tag{6.23}$$

where

$$\mathbf{F} = \begin{bmatrix} f_0^1 & f_1^1 & \cdots & f_{M_A+n}^1 \\ f_0^2 & f_1^2 & \cdots & f_{M_A+n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ f_0^n & f_1^n & \cdots & f_{M_A+n}^n \end{bmatrix}, \qquad \mathbf{D}(\mathbf{F}) = \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{F}\mathbf{F}^T \end{bmatrix} \tag{6.24}$$

6.4 BCIV method

Inspection of equations (6.12) and (6.23) leads to

$$\boldsymbol{\theta} \;=\; \operatorname*{plim}_{N\to\infty} \widehat{\boldsymbol{\theta}}_{IV}(N) \;+ \operatorname*{plim}_{N\to\infty} N\sigma_v^2 P(N)\mathbf{D}(\mathbf{F})\boldsymbol{\theta} \tag{6.25}$$

which implies that a consistent estimate of the unknown system parameters can be obtained by substracting an estimate of the bias term from the IV estimate.

Motivated by the BCLS method of Sagara and Wada (1977), we can obtain the consistent estimate by the BCIV method:

$$\widehat{\boldsymbol{\theta}}_{BCIV}(N) \;=\; \widehat{\boldsymbol{\theta}}_{IV}(N) \;+ \widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})\widehat{\boldsymbol{\theta}}_{BCIV}(N-1) \tag{6.26}$$

where

$$\widetilde{\mathbf{M}}^{-1} = \operatorname*{plim}_{N\to\infty} NP(N) \;= \{E[\mathbf{m}(k)\mathbf{z}^T(k)]\}^{-1} = \{\widetilde{\mathbf{M}}_1 + \widetilde{\mathbf{M}}_2\}^{-1} \tag{6.27}$$

and

$$\begin{aligned} \widetilde{\mathbf{M}}_1 &= \mathrm{E}\begin{bmatrix} -\mathbf{z}_x(k-l) \\ \mathbf{z}_u(k) \end{bmatrix} [-\mathbf{z}_x^T(k), \mathbf{z}_u^T(k)] \\ \widetilde{\mathbf{M}}_2 &= \mathrm{E}\begin{bmatrix} \mathbf{0}_{n\times 1} \\ \mathbf{z}_v(k) \end{bmatrix} [\mathbf{0}_{1\times n}, \mathbf{z}_v^T(k)] \approx \sigma_v^2\mathbf{D}(\mathbf{F}) \end{aligned} \tag{6.28}$$

It is necessary to investigate whether this algorithm is stable. Similar to the BCLS method, when the spectral radius of the matrix $\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})$ is less than unity, the algorithm may converge (Sagara and Wada 1977). Here the spectral radius of $\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})$, $\varrho[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})]$ is defined as

$$\varrho[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})] \;=\; \max_{1\le i\le 2n} \left|\lambda_i[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})]\right| \tag{6.29}$$

where $\lambda_i[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})]$ denotes an eigen value of $\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})$.

As mentioned in chapter 5, according to **Lemma 2.1** in Stoica and Söderström (1982), if $\widetilde{\mathbf{M}}$, $\widetilde{\mathbf{M}}_1$ are positive definite and $\widetilde{\mathbf{M}}_2 \approx \sigma^2{}_v\mathbf{D}(\mathbf{F})$ is non-negative definite, then $\varrho[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})]$ is less than unity, and hence the algorithm is stable and converges (Sagara and Wada 1977). That is a sufficient condition for stability. However, in contrast to the case of the BCLS method (Sagara and Wada 1977), unfortunately, for the delay parameter $l \geq 1$, $\widetilde{\mathbf{M}}_1$ is not symmetric, and hence we cannot expect $\widetilde{\mathbf{M}}$, $\widetilde{\mathbf{M}}_1$ to be positive definite. So far, we have not found general conclusions about that under what conditions $\varrho[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})]$ is less than unity and hence the algorithm is stable. Empirical numerical studies tell that when the noise levels are not extremely high, $\varrho[\widetilde{\mathbf{M}}^{-1}\sigma^2{}_v\mathbf{D}(\mathbf{F})]$ may be less than unity and the algorithm (6.26) is stable in most cases.

However, in most practical situations, it is difficult to have *a priori* knowledge of σ_v^2, the practical applicability of the proposed method is restricted within narrow limits. If $\widehat{\sigma}_v^2$ is an consistent estimate of σ_v^2, the BCIV method becomes

$$\widehat{\boldsymbol{\theta}}_{BCIV}(N) \;=\; \widehat{\boldsymbol{\theta}}_{IV}(N) \;+\widetilde{\mathbf{M}}^{-1}\widehat{\sigma}_v^2(N)\mathbf{D}(\mathbf{F})\widehat{\boldsymbol{\theta}}_{BCIV}(N-1) \qquad (6.30)$$

In the identification process, if $\varrho[\widetilde{\mathbf{M}}^{-1}\widehat{\sigma}_v^2(N)\mathbf{D}(\mathbf{F})]$ is less than unity, the algorithm would also converge. If we monitor the value of $\varrho[\widetilde{\mathbf{M}}^{-1}\widehat{\sigma}_v^2(N)\mathbf{D}(\mathbf{F})]$ and keep it less than unity during the identification process, the algorithm will never diverge.

Remark 6.2: So far, we have not found any counter-example where $\varrho[\widetilde{\mathbf{M}}^{-1}\sigma_v^2\mathbf{D}(\mathbf{F})] \geq 1$ if σ_v^2 is known. However, it should be noted that as shown in the numerical example in this chapter, in some 'rare' cases where the noise levels are extremely high, with the estimate of σ_v^2, we may have $\varrho[\widetilde{\mathbf{M}}^{-1}\widehat{\sigma}_v^2(N)\mathbf{D}(\mathbf{F})] \geq 1$ during the whole identification process, and in these cases, we have to give up the bias compensation procedure. A simple way is to set $\widehat{\sigma}_v^2(N)$ to zero if $\varrho[\widetilde{\mathbf{M}}^{-1}\widehat{\sigma}_v^2(N)\mathbf{D}(\mathbf{F})] \geq 1$. A strict convergency analysis of the BCIV method is not available. However, simulations show the BCIV estimates converge to their true values in most cases if the noises are not extremely high.

Remark 6.3: Notice that equation (6.25) may be written as

$$\widehat{\boldsymbol{\theta}}_{BCIV}(N) = [\mathbf{I} - N\widehat{\sigma}_v^2(N)P(N)\mathbf{D}(\mathbf{F})]^{-1}\widehat{\boldsymbol{\theta}}_{IV}(N) \qquad (6.31)$$

if $[\mathbf{I} - N\widehat{\sigma}_v^2 P(N)\mathbf{D}(\mathbf{F})]^{-1}$ exists. This means that the consistent estimate can be obtained by solving a set of simultaneous linear equations. This method is similar to the bias compensation procedure investigated by James, Souter and Dixson (1972), Stoica (1980), Stoica and Söderström (1982) by simulations and theoritical analysis. And it was pointed by Stoica (1980) that this method is not always globally convergent and may explode. Although our simulations show that the results of algorithms (6.30) and (6.31) are not significantly different in typical cases, we will restrict our discussions on procedure (6.30) instead of (6.31), since in the presence of extremely high noises, we have observed that (6.31) can easily explode.

6.5 Estimation of σ_v^2

The BCIV method requires the estimate of σ_v^2. We will show the method to find $\widehat{\sigma}_v^2$. It is a good idea to utilize the information of the equation error $r(k)$, and the residual of the IV estimate $\widehat{r}(k)$. The residual $\widehat{r}(k)$ is given by

$$\widehat{r}(k) = \xi_{A0y}(k) - \mathbf{z}^T(k)\widehat{\theta}_{IV}(N) \tag{6.32}$$

Using equation (6.7), we have

$$\widehat{r}(k) = \mathbf{z}^T(k)[\theta - \widehat{\theta}_{IV}(N)] + r(k) \tag{6.33}$$

Using equation (6.33), we have

$$\begin{aligned} f(N) &= \sum_{k=k_S+1}^{k_S+N} r(k)\widehat{r}(k-l) \\ &= \sum_{k=k_S+1}^{k_S+N} r(k)\ \{\ \mathbf{z}^T(k-l)[\theta - \widehat{\theta}_{IV}(N)] + r(k-l)\ \} \\ &= \sum_{k=k_S+1}^{k_S+N} \mathbf{z}^T(k-l)r(k)\ [\theta - \widehat{\theta}_{IV}(N)] + \sum_{k=k_S+1}^{k_S+N} r(k)r(k-l) \end{aligned} \tag{6.34}$$

Hence, we have

$$\begin{aligned} \operatorname*{plim}_{N\to\infty} \frac{1}{N} f(N) &= \mathrm{E}[r(k)\widehat{r}(k-l)] \\ &= \mathrm{E}[\mathbf{z}^T(k-l)r(k)][\theta - \widehat{\theta}_{IV}(N)] + \mathrm{E}[r(k)r(k-l)] \\ &= \mathrm{E}[-\mathbf{z_y}^T(k-l)r(k), \mathbf{z_w}^T(k-l)r(k)]\ [\theta - \widehat{\theta}_{IV}(N)] + \mathrm{E}[r(k)r(k-l)] \\ &= \mathrm{E}[\mathbf{z_w}^T(k-l)r(k)][\mathbf{b} - \widehat{\mathbf{b}}_{IV}(N)] + \mathrm{E}[r(k)r(k-l)] \\ &= -\bar{\mathbf{h}}^v(\mathbf{b})[\mathbf{b} - \widehat{\mathbf{b}}_{IV}(N)]\sigma_v^2 + \sum_{j=0}^{M_A+n-l} \beta_j(\mathbf{b})\beta_{j+l}(\mathbf{b})\sigma_v^2 \\ &= \bar{\mathbf{h}}^v(\mathbf{b})\widehat{\mathbf{b}}_{IV}(N)\sigma_v^2 \end{aligned} \tag{6.35}$$

where

$$\bar{\mathbf{h}}^v(\mathbf{b}) = [\bar{h}_1^v(\mathbf{b}), \cdots, \bar{h}_n^v(\mathbf{b})], \quad \bar{h}_i^v(\mathbf{b}) = \sum_{j=0}^{M_A+n-l} f_j^i \beta_{j+l}(\mathbf{b}) \tag{6.36}$$

and β_j is defined in equation (6.22).

Consequently, the estimate of the unknown variance σ_v^2 is given as

$$\widehat{\sigma}_v^2(N) = \frac{\frac{1}{N}\sum_{k=k_S+1}^{k_S+N} \widehat{r}(k)\widehat{r}(k-l)}{\bar{\mathbf{h}}^v\left(\widehat{\mathbf{b}}_{BCIV}(N-1)\right)\widehat{\mathbf{b}}_{IV}(N)} \tag{6.37}$$

where

$$\begin{aligned} \bar{r}(k) &= \xi_{A0y}(k) - \mathbf{z}^T(k)\widehat{\theta}_{BCIV}(N-1) \\ \widehat{r}(k-l) &= \xi_{A0y}(k-l) - \mathbf{z}^T(k-l)\widehat{\theta}_{IV}(N) \end{aligned} \tag{6.38}$$

6.6 Implementation of the algorithm

This section describes the implementation techniques of the proposed method. Both the off-line and on-line algorithms are described. Based on the discussions in the previous sections, we can establish the following off-line iterative procedure:

1: Determine the pre-filter $1/A'\left(\widehat{\mathbf{a}}^{(i-1)}_{BCIV}(N), z^{-1}\right)$ where $i = 1, 2, \cdots$ and $\widehat{a}^{(0)}_{BCIV}(N)$ are given by an appropriate guess.

2: Calculate the IV estimate $\widehat{\theta}^{(i)}_{IV}(N)$:

$$\widehat{\theta}^{(i)}_{IV}(N) = \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)\mathbf{z}^T(k)\right]^{-1} \cdot \left[\sum_{k=k_S+1}^{k_S+N} \mathbf{m}(k)\xi_{A0y}(k)\right] \tag{6.39}$$

3: Estimate the input noise variance:

$$\begin{aligned} \widehat{\sigma}^{2(i)}_v(N) &= \frac{\frac{1}{N}\sum_{k=k_S+1}^{k_S+N} \bar{r}(k)\widehat{r}(k-l)}{\bar{\mathbf{h}}^v\left(\widehat{\mathbf{b}}^{(i-1)}_{BCIV}(N)\right)\widehat{\mathbf{b}}^{(i)}_{IV}(N)} \\ \bar{r}(k) &= \xi_{A0y}(k) - \mathbf{z}^T(k)\widehat{\theta}^{(i-1)}_{BCIV}(N) \\ \widehat{r}(k-l) &= \xi_{A0y}(k-l) - \mathbf{z}^T(k-l)\widehat{\theta}^{(i)}_{IV}(N) \end{aligned} \tag{6.40}$$

where $\widehat{\theta}^{(0)}_{BCIV}(N)$ is replaced by $\widehat{\theta}^{(1)}_{IV}(N)$.

4: Compensate the bias of the IV estimate $\widehat{\theta}^{(i)}_{IV}(N)$:

$$\widehat{\theta}^{(i)}_{BCIV}(N) = \widehat{\theta}^{(i)}_{IV}(N) + N\widehat{\sigma}^{2(i)}_v(N)P^{(i)}(N)\mathbf{DF}\left(\widehat{\mathbf{a}}^{(i-1)}_{BCIV}(N)\right)\widehat{\theta}^{(i-1)}_{BCIV}(N) \tag{6.41}$$

where $\widehat{\theta}^{(0)}_{BCIV}(N)$ is replaced by $\widehat{\theta}^{(1)}_{IV}(N)$.

5: Increase the iteration number i and return to **1** untill convergence.

Obviously, the above off-line procedure requires a great amount of computational burden and is hence not practical. Usually, the recursive on-line identification algorithms have contrastively a small requirement of primary memory since only a modest amount of information is stored. Considering the practical applications, it is still necessary to derive an on-line version of the above off-line estimation procedure. Many recursive identification methods are derived as approximations of off-line methods. It may therefore happen that the price paid for the approximation is a reduction in accuracy. For long data records, however, the statistical efficiency will approach that of the off-line algorithm (Young and Jakeman 1979, Ljung and Söderström 1983, Söderström and Stoica 1989) and hence the difference is not significant.

The approximated recursive BCIV algorithm is described as follows.

1: Determine the pre-filter $1/A'(\bar{\mathbf{a}}_{BCIV}(N), z^{-1})$ as mentioned above, for $N = 1, 2, \cdots$:

$$\begin{cases} \bar{\mathbf{a}}(N) = \bar{\mathbf{a}}(0) & N \leq N_0 \\ \bar{\mathbf{a}}(N) = (1-\mu)\bar{\mathbf{a}}(N-1) + \mu\hat{\mathbf{a}}_{BC}(N-d) & N > N_0 \end{cases} \tag{6.42}$$

where μ and d can be chosen to be

$$\begin{aligned} \mu &= 0.05 \\ d &= 10 \end{aligned} \tag{6.43}$$

2: Calculate the IV estimate $\hat{\boldsymbol{\theta}}_{IV}(N)$:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{IV}(N) &= \hat{\boldsymbol{\theta}}_{IV}(N-1) + L(N)[\xi_{AF0y}(k_S+N) - \mathbf{z}^T(k_S+N)\hat{\boldsymbol{\theta}}_{IV}(N-1)] \\ L(N) &= \frac{P(N-1)\mathbf{m}(k_S+N)}{\rho(N) + \mathbf{z}^T(k_S+N)P(N-1)\mathbf{m}(k_S+N)} \\ P(N) &= \frac{1}{\rho(N)}\left[P(N-1) - \frac{P(N-1)\mathbf{m}(k_S+N)\mathbf{z}^T(k_S+N)P(N-1)}{\rho(N) + \mathbf{z}^T(k_S+N)P(N-1)\mathbf{m}(k_S+N)}\right] \end{aligned} \tag{6.44}$$

where $\rho(N)$ is the forgetting factor. A typical choice of the forgetting factor is given as

$$\rho(N) = (1-0.01)\rho(N-1) + 0.01, \quad \rho(k_S) = 0.95 \tag{6.45}$$

3: Estimate the input noise variance σ_v^2 and monitor the stability for $N > N_0$:

$$\hat{\sigma}_v^{2\prime}(N) = \frac{\frac{1}{N}f(N)}{\bar{\mathbf{h}}^v\left(\hat{\mathbf{b}}_{BCIV}(N-1)\right)\hat{\mathbf{b}}_{IV}(N)} \tag{6.46}$$

and

$$\begin{cases} \hat{\sigma}_v^2(N) = \hat{\sigma}_v^{2\prime}(N) & \text{if} \quad \varrho\left[NP_B(N)\hat{\sigma}_v^{2\prime}(N)\mathbf{DF}\left(\bar{\mathbf{a}}_{BCIV}(N)\right)\right] < 1 \\ \hat{\sigma}_v^2(N) = 0 & \text{if} \quad \varrho\left[NP_B(N)\hat{\sigma}_v^{2\prime}(N)\mathbf{DF}\left(\bar{\mathbf{a}}_{BCIV}(N)\right)\right] \geq 1 \end{cases} \tag{6.47}$$

where

$$\begin{cases} f(N) = \\ \qquad \sum_{k=k_S+1}^{k_S+N} [\xi_{A0_y}(k) - \mathbf{z}^T(k)\widehat{\theta}_{BCIV}(N-1)][\xi_{A0_y}(k-l) - \mathbf{z}^T(k-l)\widehat{\theta}_{IV}(N)] & N = N_0 \\ f(N) = f(N-1) + \bar{r}(N)\hat{r}(N-l) & N > N_0 \end{cases} \tag{6.48}$$

and

$$\begin{aligned} \bar{r}(N) &= \xi_{A0_y}(k_S+N) - \mathbf{z}^T(k_S+N)\widehat{\theta}_{BCIV}(N-1) \\ \hat{r}(N-l) &= \xi_{A0_y}(k_S+N-l) - \mathbf{z}^T(k_S+N-l)\widehat{\theta}_{IV}(N) \end{aligned} \tag{6.49}$$

4: Compensate the bias of the IV estimate $\widehat{\theta}_{IV}(N)$:

$$\begin{cases} \widehat{\theta}_{BCIV}(N) = \widehat{\theta}_{IV}(N) & N \leq N_0 \\ \widehat{\theta}_{BCIV}(N) = \widehat{\theta}_{IV}(N) + N\widehat{\sigma}_v^2(N)P(N)\mathbf{DF}(\bar{\mathbf{a}}_{BCIV}(N))\,\widehat{\theta}_{BCIV}(N-1) & N > N_0 \end{cases} \tag{6.50}$$

5: Increase the recursion number N and return to **1** untill convergence.

Remark 6.4: Due to the adaptions of the pre-filter and the small initial value of the forgetting factor, the IV estimate $\widehat{\theta}_{IV}(N)$ may change roughly during the beginning recursions (Sagara, Yang and Wada 1991c). Hence it is nonsense to start the BCIV algorihm together with the IV algorithm. It is often a practical policy to delay the start of the BCIV algorithm by $N_0 + 1$ recursions. To the experience, $N_0 = 100 \sim 500$ are reasonable when the system signals are corrupted by high measurement noises.

Remark 6.5: As mentioned previously, it is necessary to monitor the stability of the algorithm and contract the estimates within the stable region. The simplest way of this modification is to keep the estimates in the stable region during the recursions, as described in equation (6.47). This means that when instabilty of the algorithm is detected, we have to avoid the bias compensation procedure by setting $\widehat{\sigma}_v^2(N)$ to zero. Experience shows that the zero setting as shown in equation (6.47) typically takes place at only few samples in the beginng recursions. The information loss by ignoring certain samples, as in equation (6.47) is therefore moderate.

6.7 Illustrative examples

To illustrate the effectiveness of the proposed estimation algorithm, we consider a second-order system described by

$$\begin{gathered} \ddot{x}(t) + a_1\dot{x}(t) + a_2x(t) = b_1\dot{u}(t) + b_2u(t) \\ a_1 = 3.0, \quad a_2 = 4.0, \quad b_1 = 0.0, \quad b_2 = 4.0 \end{gathered} \tag{6.51}$$

The input $u(t)$ is the output of a second-order continuous-time Butterworth input filter driven by a stationary random signal $\zeta(t)$:

$$u(t) = L(p)\zeta(t) = \frac{1}{(p/\omega_c)^2 + \sqrt{2}(p/\omega_c) + 1}\,\zeta(t), \quad \omega_c = 4.0 \tag{6.52}$$

Simulation experiments are carried out when the sampling interval is taken to be $T = 0.04$, and in this case, $\sigma_u = 2.38, \sigma_x = 0.69$.

The on-line recursive algorithm is applied in the simulation experiments. The bias compensation procedure is performed when $N > k_S + N_0 = k_S + 100$, and the pre-filter $Q_{IA}(z^{-1})$ is initialized as

$$\begin{aligned} Q_{IA}(z^{-1}) &= \frac{1}{\sum_{i=0}^{2} a_i(\frac{T}{2})^i(1+z^{-1})^i(1-z^{-1})^{2-i}} \\ & a_0 = 1.0, \; a_1 = 11.0, \; a_2 = 10.0 \end{aligned} \tag{6.53}$$

Example 6.1: Effects of the values of l and M_A.

The effects of the delay parameter l of the delayed filtered outputs in the instrumental variables, and the order M_A of the FIR filter $Q_{FA}(z^{-1})$ which approximates the pre-filter $1/A'(z^{-1})$ are first investigated. To take comparison of the BCIV method with the BCLS method described in chapter 5, we first perform the simulation experiments under the same conditions as given in chapter 5. The input and output noises are assumed to be white:

$$\begin{aligned} y(k) &= x(k) + e(k) \\ w(k) &= u(k) + v(k) \end{aligned} \tag{6.54}$$

and the noise variances are given as

$$\sigma_e = 0.17 \qquad \sigma_v = 0.60 \tag{6.55}$$

which let N/S ratio(NSR)= $\sigma_e/\sigma_x \approx \sigma_v/\sigma_u \approx 25\%$.

Table 6.1: Estimates for different values of l ($M_A = 50$).

	$l = 1$	$l = 2$	$l = 4$	$l = 8$	$l = 12$
$\widehat{a}_{1IV}$ (3.0)	2.5626 ±0.0618	2.6146 ±0.0621	2.7045 ±0.0670	2.8235 ±0.0751	2.8899 ±0.0857
$\widehat{a}_{2IV}$ (4.0)	3.5073 ±0.0615	3.5539 ±0.0595	3.6376 ±0.0601	3.7570 ±0.0621	3.8291 ±0.0704
$\widehat{b}_{1IV}$ (0.0)	0.0921 ±0.0111	0.0839 ±0.0109	0.0695 ±0.0113	0.0488 ±0.0121	0.0365 ±0.0136
$\widehat{b}_{2IV}$ (4.0)	3.4117 ±0.0589	3.4689 ±0.0580	3.5689 ±0.0608	3.7047 ±0.0638	3.7832 ±0.0797
$\widehat{a}_{1BC}$ (3.0)	3.0396 ±0.0747	3.0324 ±0.0673	3.0245 ±0.0756	3.0215 ±0.0818	3.0178 ±0.0938
$\widehat{a}_{2BC}$ (4.0)	4.0414 ±0.0695	4.0328 ±0.0604	4.0233 ±0.0668	4.0204 ±0.0676	4.0149 ±0.0798
$\widehat{b}_{1BC}$ (0.0)	0.0003 ±0.0121	0.0016 ±0.0110	0.0033 ±0.0117	0.0038 ±0.0124	0.0047 ±0.0148
$\widehat{b}_{2BC}$ (4.0)	4.0655 ±0.0707	4.0567 ±0.0610	4.0464 ±0.0677	4.0438 ±0.0767	4.0368 ±0.0912
$\widehat{\sigma}_v$ (0.60)	0.5924 ±0.0298	0.5904 ±0.0285	0.5897 ±0.0259	0.5939 ±0.0300	0.5892 ±0.0301
$\Delta\|\theta\|_{IV}$	0.8880	0.7979	0.6398	0.4240	0.2994
$\Delta\|\theta\|_{BC}$	0.0870	0.0731	0.0575	0.0519	0.0438
$\sigma_{\theta IV}$	0.0483	0.0476	0.0498	0.0542	0.0623
$\sigma_{\theta BC}$	0.0567	0.0499	0.0555	0.0589	0.0699

Table 6.2: Estimates for different values of M_A ($l = 8$).

	$M_A = 20$	$M_A = 30$	$M_A = 40$	$M_A = 50$	$M_A = 60$	$M_A = 100$
$\hat{a}_{1BC}$ (3.0)	3.1007 $\pm$0.0877	3.0393 $\pm$0.0824	3.0240 $\pm$0.0814	3.0211 $\pm$0.0812	3.0208 $\pm$0.0811	3.0208 $\pm$0.0811
$\hat{a}_{2BC}$ (4.0)	4.1274 $\pm$0.0754	4.0442 $\pm$0.0679	4.0239 $\pm$0.0667	4.0201 $\pm$0.0665	4.0198 $\pm$0.0664	4.0198 $\pm$0.0665
$\hat{b}_{1BC}$ (0.0)	-0.0170 $\pm$0.0162	-0.0005 $\pm$0.0133	0.0030 $\pm$0.0127	0.0037 $\pm$0.0126	0.0037 $\pm$0.0126	0.0037 $\pm$0.0126
$\hat{b}_{2BC}$ (4.0)	4.1771 $\pm$0.0860	4.0736 $\pm$0.0786	4.0478 $\pm$0.0780	4.0429 $\pm$0.0755	4.0424 $\pm$0.0753	4.0425 $\pm$0.0754
$\hat{\sigma}_v$ (0.60)	0.8236 $\pm$0.0394	0.6430 $\pm$0.0321	0.6012 $\pm$0.0303	0.5939 $\pm$0.0300	0.5935 $\pm$0.0301	0.5957 $\pm$0.0300
$\Delta\|\theta\|_{BC}$	0.2409	0.0941	0.0587	0.0519	0.0514	0.0514
$\sigma_{\theta BC}$	0.0663	0.0603	0.0592	0.0589	0.0589	0.0589

We have shown that due to the considerable high noises described above, the conventional LS estimates are greatly biased, and in this case the BCLS method gives very accurate estimates.

The results for different values of l and M_A are shown in Table 6.1 and Table 6.2 respectively. It is clear from Table 6.1, that the IV estimates are greatly biased as pointed out through theoretical analysis in section 6.3. In contrast, the BCIV estimates are quite accurate. Furthermore, it seems that the efficiency of the BCIV estimates is comparable to that of the BCLS estimates in the presence of white noises. It should be noted that in this case, when the delay parameter $l \geq 1$, the BCIV estimates are consistent according to the theoretical analysis in section 6.3. However, during the estimation process, the pre-filter $1/A'(z^{-1})$ is constructed adaptively by the BCIV estimates, and any error of the BCIV estimates may make the IV vector $\mathbf{m}(k)$ be slightly correlated with the output noise $e(k)$. These possible correlations can be reduced by taking a considerable large $l (l \geq 4)$, and hence the BCIV estimates are more accurate than in the case of small $l (l \leq 2)$. However, an unnecessarily large l should be avoided, since a large l may destroy the existence of $P(N) = [\sum_{k=ks+1}^{ks+N} \mathbf{m}(k)\mathbf{z}^T(k)]^{-1}$ (Söderström and Stoica 1981) and hence may make the algorithm very sensitive or even unstable. It seems that selection of an appropriate l is not a difficult task since the results are not sensitive to l.

Table 6.2 shows that the estimation error of the BCIV estimates decreases when the order

Table 6.3: Estimates for different noise levels($M_A = 50, l = 12$).

NSR	40%	50%	60%	80%	100%
$\widehat{a}_{1IV}$ (3.0)	2.7290 ±0.1280	2.6168 ±0.1419	2.5058 ±0.1412	2.2313 ±0.2527	1.9929 ±0.3225
$\widehat{a}_{2IV}$ (4.0)	3.5864 ±0.1006	3.4250 ±0.1347	3.2617 ±0.1414	2.8344 ±0.2522	2.4778 ±0.3401
$\widehat{b}_{1IV}$ (0.0)	0.0772 ±0.0207	0.1035 ±0.0272	0.1250 ±0.0792	0.1856 ±0.0550	0.2232 ±0.0712
$\widehat{b}_{2IV}$ (4.0)	3.4579 ±0.1081	3.2200 ±0.1286	2.9797 ±0.1245	2.4146 ±0.2579	1.9303 ±0.3011
$\widehat{a}_{1BC}$ (3.0)	3.0167 ±0.1460	3.0330 ±0.1952	3.0503 ±0.2007	2.9469 ±0.4703	2.8118 ±0.7634
$\widehat{a}_{2BC}$ (4.0)	4.0031 ±0.1234	4.0258 ±0.1960	4.0459 ±0.1900	3.8522 ±0.5245	3.6294 ±0.9437
$\widehat{b}_{1BC}$ (0.0)	0.0082 ±0.0245	0.0072 ±0.0369	0.0035 ±0.0372	0.0408 ±0.10045	0.0766 ±0.1561
$\widehat{b}_{2BC}$ (4.0)	4.0299 ±0.1445	4.0522 ±0.2168	4.0778 ±0.2178	3.8483 ±0.7001	3.5844 ±1.1982
σ_v	0.96	1.20	1.44	1.92	2.40
$\widehat{\sigma}_v$	0.9312 ±0.0553	1.1608 ±0.0726	1.3916 ±0.0851	1.7719 ±0.4340	1.9464 ±0.8735
$\Delta\|\theta\|_{IV}$	0.7378	1.0472	1.3586	2.1207	2.7686
$\Delta\|\theta\|_{BC}$	0.0353	0.0673	0.1035	0.2222	0.5928
$\sigma_{\theta IV}$	0.0875	0.1081	0.1091	0.2045	0.2587
$\sigma_{\theta BC}$	0.1096	0.1612	0.1615	0.4488	0.7653

M_A of the FIR filter $Q_{FA}(z^{-1})$ defined in equation (6.18) increases. This is due to the fact that, in general, the FIR filter $Q_{FA}(z^{-1})$ becomes a better approximation of the pre-filter $1/A'(z^{-1})$ as the order increases. A large M_A, however, also increases the computation and storage requirement. The results in Table 6.2 show that for the system under study, when M_A exceeds about 40, the estimates do not vary significantly, therefore it is not necessary to choose too large an M_A.

Example 6.2: Parameter estimates in the presence of extremely high noises.

To investigate the performance of the BCIV method in the presence of high noises, some simulation experiments are performed for different values of NSR. The results are

Table 6.4: Estimates under correlated output noise(NSR≈ 25%, $M_A = 50$).

	$l=1$	$l=8$	$l=10$	$l=12$
$\widehat{a}_{1IV}$ (3.0)	2.6036 ±0.0748	2.8327 ±0.0766	2.8692 ±0.0814	2.8951 ±0.0860
$\widehat{a}_{2IV}$ (4.0)	3.5943 ±0.0692	3.7700 ±0.0655	3.8091 ±0.0685	3.8373 ±0.0735
$\widehat{b}_{1IV}$ (0.0)	0.0703 ±0.0132	0.0469 ±0.0145	0.0403 ±0.0151	0.0355 ±0.0159
$\widehat{b}_{2IV}$ (4.0)	3.4944 ±0.0733	3.7171 ±0.0671	3.7600 ±0.0703	3.7910 ±0.0781
$\widehat{a}_{1BC}$ (3.0)	4.1379 ±0.1976	3.0222 ±0.0839	3.0197 ±0.0922	3.0182 ±0.0965
$\widehat{a}_{2BC}$ (4.0)	5.2572 ±0.1833	4.0234 ±0.0733	4.0915 ±0.0818	4.0167 ±0.0850
$\widehat{b}_{1BC}$ (0.0)	−0.1991 ±0.0257	0.0034 ±0.0155	0.0043 ±0.0166	0.0048 ±0.0172
$\widehat{b}_{2BC}$ (4.0)	5.5878 ±0.2345	4.0442 ±0.0762	4.0390 ±0.0888	4.0365 ±0.0932
$\widehat{\sigma}_v$ (0.60)	0.9989 ±0.0213	0.5900 ±0.0348	0.5836 ±0.0351	0.5820 ±0.0326
$\Delta\|\theta\|_{IV}$	0.7631	0.4038	0.3358	0.2871
$\Delta\|\theta\|_{BC}$	2.3413	0.0548	0.0481	0.0443
$\sigma_{\theta IV}$	0.0578	0.0557	0.0583	0.0634
$\sigma_{\theta BC}$	0.1603	0.0622	0.0699	0.0730

shown in Table 6.3. It is clear from Table 6.3 that in the presence of high noises, the IV estimates are not acceptable at all due to the great effects of the noises. However, when the noises are not extremely high (NSR≤ 60%), the BCIV estimates still converge to their true values. Unfortunately, when NSR exceed 80%, even the BCIV estimates are not satisfactory. Especially when NSR= 100%, we have observed that some realizations of the BCIV estimates fail to converge, if we do not monitor the stability of the algorithm and contract the estimates within the stable region as described previously. It should, however, be stressed that this fact by no means implies that the BCIV method is useless, since practically the situations where NSR of both the input and output noises exceed 80% are 'rare'. Rather, it means it is important to monitor the stability of the algorithm during the identification process in the presence of very high noises.

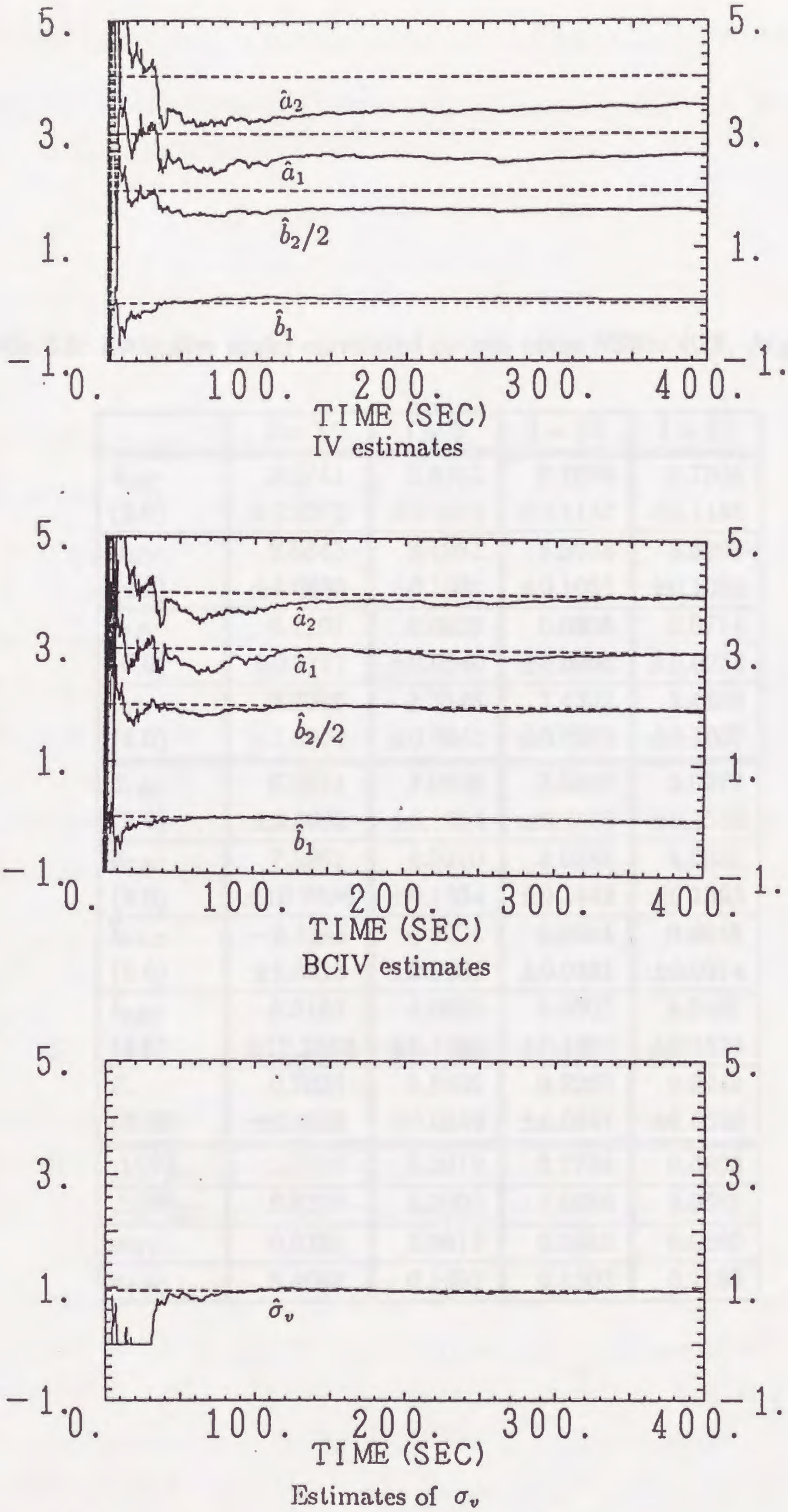


Figure 6.1: Parameter estimates under correlated noise (NSR≈ 40%, $l = 8$).

Table 6.5: Estimates under correlated output noise(NSR≈ 40%, $M_A = 50$).

	$l=1$	$l=8$	$l=10$	$l=12$
$\widehat{a}_{1IV}$ (3.0)	2.2741 ±1.2253	2.6353 ±0.1064	2.7025 ±0.1135	2.7504 ±0.1156
$\widehat{a}_{2IV}$ (4.0)	2.9843 ±1.0835	3.4961 ±0.1031	3.5678 ±0.1051	3.6203 ±0.1068
$\widehat{b}_{1IV}$ (0.0)	0.2107 ±0.1777	0.0929 ±0.0240	0.0805 ±0.0236	0.0714 ±0.0232
$\widehat{b}_{2IV}$ (4.0)	2.7936 ±1.2576	3.3546 ±0.0941	3.4302 ±0.0998	3.4859 ±0.1007
$\widehat{a}_{1BC}$ (3.0)	6.3674 ±9.5672	3.0408 ±0.1354	3.0309 ±0.1505	3.0272 ±0.1515
$\widehat{a}_{2BC}$ (4.0)	7.5261 ±10.2896	4.0410 ±0.1334	4.0285 ±0.1442	4.0237 ±0.1393
$\widehat{b}_{1BC}$ (0.0)	−0.4535 ±1.5415	0.0024 ±0.0281	0.0044 ±0.0281	0.0048 ±0.0274
$\widehat{b}_{2BC}$ (4.0)	8.3185 ±12.2363	4.0688 ±0.1395	4.0502 ±0.1582	4.0461 ±0.1574
$\widehat{\sigma}_v$ (0.96)	0.7025 ±0.9664	0.9405 ±0.0549	0.9260 ±0.0547	0.9249 ±0.0546
$\Delta\|\theta\|_{IV}$	1.7488	0.9012	0.7788	0.6898
$\Delta\|\theta\|_{BC}$	6.5276	0.0900	0.0656	0.0588
$\sigma_{\theta IV}$	0.9360	0.0819	0.0855	0.0866
$\sigma_{\theta BC}$	8.4087	0.1091	0.1202	0.1189

Example 6.3: Parameter estimates when the output noise is correlated.

The effects of the correlations of the output noise are investigated. The measurements of the input and output signals are given as

$$\begin{aligned} y(k) &= x(k) + C(z^{-1})e(k) \\ w(k) &= u(k) + v(k) \end{aligned} \tag{6.56}$$

where $C(z^{-1})$ is

$$C(z^{-1}) = \frac{1}{1 - 0.5z^{-1}} \tag{6.57}$$

and can be approximated by an MA process model with finite order r:

$$\begin{aligned} C(z^{-1}) &\approx 1 + 0.5z^{-1} + 0.5^2z^{-2} + 0.5^3z^{-3} + \cdots + 0.5^rz^{-r} \\ &\approx 1 + 0.5z^{-1} + 0.25z^{-2} + 0.125z^{-3} + 0.0625z^{-4} + 0.03125z^{-5} + 0.015625z^{-6} \end{aligned} \tag{6.58}$$

Although in section 6.2 it is assumed that $C(z^{-1})$ is an MA process with finite length r, to show that the proposed method is still efficient even when $C(z^{-1})$ is an stable AR process which can be approximated by an MA process with finite length, we choose $C(z^{-1})$ as equation (6.57) for simulation study. Therefore, for delay parameter $l > r = 6$, we may expect the BCIV estimates to be consistent. However, it might be noted that the BCLS method does not give consistent estimates when the output noise is correlated.

Simulation experiments are performed for two different noise levels:

$$\begin{aligned} &(i) \quad \sigma_e = 0.15, \quad \sigma_v = 0.60, \quad \text{NSR} \approx 25\% \\ &(ii) \quad \sigma_e = 0.24, \quad \sigma_v = 0.96, \quad \text{NSR} \approx 40\% \end{aligned} \tag{6.59}$$

The results for the two different noise levels are shown in Table 6.4 and Table 6.5 respectively. From Tables 6.4~6.5, we know that if we neglect the correlations of the output noise and hence let $l = 1$, the BCIV estimates are greatly biased. Therefore it is dangerous to use the bias compensating techniques when the correlations of the output noise are neglected. However, for large values of $l = 8, 10, 12$, i.e. $l > r$, the BCIV estimates are very accurate. To show the convergency properties of the BCIV algorithm, the associated estimates of one example where NSR$\approx$ 40% are plotted in Figure 6.1. As will be seen from the figures, the BCIV estimates are much more accurate than the IV estimates, however in the presence of very high noises the BCIV estimates require long samples to ensure convergence. Additionally, it is found that the estimate of the standard deviation σ_v of the input noise approaches the true value very quickly.

6.8 Conclusion

In this chapter, a new method has been proposed for identification of continuous systems in the case where the discrete input measurement is corrupted by a white noise and the discrete output measurement is corrupted by a noise which may be coloured.

The continuous system is identified through the approximated discrete-time estimation model with continuous system parameters involving adaptive IIR filters. The adaptive IIR filter is introduced to attempt to filter off unwanted high frequency components of the noises and also to let the output noise remain in its original form in the equation error.

Then a BCIV method is proposed to obtain consistent estimates. The effects of the output noise are avoided by introducing an IV method with filtered inputs and delayed filtered outputs as instrumental variables, while the bias of the IV estimate due to the input noise is compensated with the estimated variance of the input noise. The proposed method does not require the assumption that the output noise is white and is hence more practical.

The approximated recursive version of the proposed method is also described in detail. And it is pointed out that the proposed algorithm is stable under weak conditions.

The proposed recursive identification algorithm is easy to implement on digital computers and numerical examples show that the proposed recursive identification algorithm is quite efficient even in the presence of high input-output noises.

Chapter 7

Parameter Identification of Distributed Parameter Systems in the Presence of Measurement Noise

7.1 Introduction

In practical situations, we often come across dynamical systems described by partial differential equations. Therefore identification of distributed parameter systems (DPSs) is very important in practice. Applications of orthogonal functions in the field of DPS identification are studied by several researchers such as double general orthogonal polynomials (Lee and Chang 1986), Laguerre polynomials (Ranganathan, Jha and Rajamani 1984,1986,1987), Chebyshev series (Horng, Chou and Tsai 1986), Walsh functions (Sinha, Rajamani and Sinha 1980), block-pulse functions (Hsu and Cheng 1982, Jiang Ning and Jiang Jiong 1988), Laguerre operational matrices (Jha and Zaman 1985) and shifted Legendre polynomials (Mohan and Datta 1987) etc. A unified approach of these methods was given by Mohan and Datta (1991). These approaches first derive some operational matrices for double integrations from the orthogonal functions, then the partial differential equation which characterizes the dynamics of the DPS under study is converted into a set of over-determined linear algebraic equations by the operational matrices. Therefore if the input-output data can be observed, the unknown parameters can be estimated directly by the LS algorithm without using direct partial differentiations which may accentuate the measurement noise. However, the dimensions of these operational matrices may grow drastically according to the number of observed data, thus these algorithms are not suitable for on-line processing. With these methods, it is also difficult to consider the consistency problem in the presence of stochastic measurement noise. Furthermore, with the orthogonal function expansion approaches, the initial and boundary conditions should be estimated as unknown parameters. Since much more parameters including those concerning the initial and boundary conditions are estimated by the LS type algorithms, it should be careful to choose the input signals to obtain unique solution to the parameter estimation problem. Hara, Yamamoto and Ougita (1988) also proposed a direct method for identification of DPSs. In their method, the noise contaminated observed data are smoothed by splines to get an approximation function of the real output before the parameters are identified. However, although their method can avoid treatment of the initial and boundary conditions, the smoothing procedure requires complex calculations and should be implemented in a batch manner. Although the integration operational matrices using orthogonal functions or the smoothing procedure using splines are introduced to reduce the noise effects, the consistency problem, that is whether these estimation methods can give consistent estimates in the presence of high noise were not discussed.

In this chapter, the problem of direct parameter identification of a second-order DPS is discussed under unknown initial condition and boundary condition, when input and noise corrupted output are available at a finite number of discrete points equidistantly located in the space domain. A two-dimensional low-pass filter which is designed in continuous time-space domain is introduced to reduce the effects of measurement noise. A class of two-dimensional low-pass IIR filters is then obtained by discretizing the designed low-pass filter via the bilinear transformation. Thus a discrete estimation model of the DPS under study is constructed with filtered input-output data for recursive on-line identification algorithms. The main advantage of this discrete estimation model is that the parameter identification procedure does not treat any initial and boundary conditions and thus the identification algorithm can be implemented in an on-line manner. If the pass-band of the filter is designed appropriately, the noise effects can be reduced and therefore accurate estimates can be obtained by the LS method. It is shown by simulation that the estimates are not sensitive to the selection of the filter parameters in some cases. When the filter is not designed appropriately, or when the output is corrupted by high noise, the LS method cannot give unbiased estimates. In these cases, we consider a recursive IV method with filtered input data as instrumental variables. No restrictions about specific types of the stochastic properties of measurement noise are imposed. Theoretical analysis shows that the IV method gives consistent estimates. Simulation results show that the IV method is quite efficient when the output data is corrupted by high noise or the filter parameters are not selected successfully.

7.2 Estimation model

Consider the following second-order partial differential equation (Mohan and Datta 1987, Hara, Yamamoto and Ougita 1988):

$$c_1\frac{\partial^2 y(x,t)}{\partial t^2}+c_2\frac{\partial^2 y(x,t)}{\partial x\partial t}+c_3\frac{\partial^2 y(x,t)}{\partial x^2}+c_4\frac{\partial y(x,t)}{\partial t}+c_5\frac{\partial y(x,t)}{\partial x}+c_6 y(x,t)=u(x,t)$$

$$(t\geq 0,\quad 0\leq x\leq L) \tag{7.1}$$

where $u(x,t)$, $y(x,t)$ and $c_i(i=1,\cdots,6)$ denote the system input, real system output and the unknown parameters respectively.

The time sampling instants for the system signals are assumed to be

$$t=N_ST,\ (N_S+1)T,\ \cdots,nT,\ \cdots,\ N_FT$$

where T is the time sampling interval and N_S, N_F are natural numbers. And for each time sampling instant, the system signals are available at a finite number of equidistant observation points in the space domain located at

$$x=M_LX,\ (M_L+1)X,\ \cdots,mX,\ \cdots,\ M_RX$$

where X is the spatial observation interval and M_L, M_R are natural numbers.

Practically, the output is corrupted by a measurement noise. Denoting $y(mX,nT)$ as $y(m,n)$ for convenience of notation, the noisy output observation vector is defined to be

$$\begin{aligned}\mathbf{z}(n)&=\mathbf{y}(n)+\mathbf{v}(n)\\ \mathbf{z}(n)&=[z(M_L,n),z(M_L+1,n),\cdots,z(m,n),\cdots,z(M_R,n)]^T\\ \mathbf{y}(n)&=[y(M_L,n),y(M_L+1,n),\cdots,y(m,n),\cdots,y(M_R,n)]^T\\ \mathbf{v}(n)&=[v(M_L,n),v(M_L+1,n),\cdots,v(m,n),\cdots,v(M_R,n)]^T\end{aligned} \tag{7.2}$$

Considering the practical situations, the restrictions about the stochastic properties of noise process vector $\mathbf{v}(n)$ should be as mild as possible. Here we only assume that the noise vector is uncorrelated with the input and is a widely stationary stochastic process with zero-mean, i.e.

$$E[\mathbf{v}(n)]=0 \tag{7.3}$$

Since partial differential operations may accentuate the measurement noise, it is critical to avoid using direct approximations of the partial differentiations from the discrete observed input-output data. A lot of efforts have been made to this problem. The methods using operational properties of orthogonal functions have been studied by several authors in recent years. However, the unknown initial and boundary conditions should also be estimated as unknown parameters and thus the number of the parameters to be estimated may increase drastically according to the number of the observed data. Therefore these methods are not convenient and not suitable for recursive on-line identification. Instead of taking double integrations by orthogonal functions, we introduce the following two-dimensional filter to remove the direct partial differentiations:

$$F(p_1,p_2) = \frac{1}{(\tau_1 p_1 + 1)^2(\tau_2 p_2 + 1)^2} \tag{7.4}$$

where τ_1,τ_2 are space constant and time constant, and p_1, p_2 denote the partial differential operators:

$$\begin{aligned} p_1 &= \frac{\partial}{\partial x} \\ p_2 &= \frac{\partial}{\partial t} \end{aligned} \tag{7.5}$$

Multiplying both sides of the partial differential equation (7.1) by the two-dimensional filter leads to

$$\begin{aligned} c_1\ F_{20}(p_1,p_2)\ y(x,t) + c_2\ F_{11}(p_1,p_2)\ y(x,t) + c_3\ F_{02}(p_1,p_2)\ y(x,t) + c_4\ F_{21}(p_1,p_2)\ y(x,t) + \\ c_5\ F_{12}(p_1,p_2)\ y(x,t) + c_6\ F_{22}(p_1,p_2)\ y(x,t) = F_{22}(p_1,p_2)\ u(x,t) \end{aligned} \tag{7.6}$$

where

$$F_{ij}(p_1,p_2) = F(p_1,p_2)p_1^{2-i}p_2^{2-j}, \qquad i,j = 0,1,2 \tag{7.7}$$

It is known that the partial differential operator p_1, p_2 can be replaced by the bilinear transformation as

$$\begin{aligned} p_1 &= \frac{2}{X}\frac{1-z_1^{-1}}{1+z_1^{-1}} \\ p_2 &= \frac{2}{T}\frac{1-z_2^{-1}}{1+z_2^{-1}} \end{aligned} \tag{7.8}$$

where z_1^{-1}, z_2^{-1} are the shift operators which let

$$z_1^{-i}z_2^{-j}y(x,t) = y(x-iX, t-jT) \tag{7.9}$$

The filters in equation (7.6) can be discretized into two-dimensional IIR digital filters by the bilinear transformation:

$$F'_{ij}(z_1^{-1},z_2^{-1}) = \frac{\left(\frac{X}{2}\right)^i (1-z_1^{-1})^{2-i}(1+z_1^{-1})^i \left(\frac{T}{2}\right)^j (1-z_2^{-1})^{2-j}(1+z_2^{-1})^j}{\left\{\left(\tau_1+\frac{X}{2}\right)+\left(\frac{X}{2}-\tau_1\right)z_1^{-1}\right\}^2 \left\{\left(\tau_2+\frac{T}{2}\right)+\left(\frac{T}{2}-\tau_2\right)z_2^{-1}\right\}^2} \tag{7.10}$$

Hence we have the discrete estimation model from equation (7.6):

$$
\begin{aligned}
&c_1\xi_{20y}(m,n) + c_2\xi_{11y}(m,n) + c_3\xi_{02y}(m,n) + c_4\xi_{21y}(m,n) + c_5\xi_{12y}(m,n) + c_6\xi_{22y}(m,n) \\
&= \xi_{22u}(m,n)
\end{aligned}
\tag{7.11}
$$

where $\xi_{ijy}(m,n), \xi_{iju}(m,n)$ are the outputs of the two-dimensional filters:

$$
\begin{aligned}
\xi_{ijy}(m,n) &= F'_{ij}(z_1^{-1}, z_2^{-1})y(m,n) \\
\xi_{iju}(m,n) &= F'_{ij}(z_1^{-1}, z_2^{-1})u(m,n)
\end{aligned}
\tag{7.12}
$$

Substituting the observations of equation (7.2) into the discrete estimation model, we have

$$
\begin{aligned}
&c_1\xi_{20z}(m,n) + c_2\xi_{11z}(m,n) + c_3\xi_{02z}(m,n) + c_4\xi_{21z}(m,n) + c_5\xi_{12z}(m,n) + c_6\xi_{22z}(m,n) \\
&\quad = \xi_{22u}(m,n) + w(m,n)
\end{aligned}
\tag{7.13}
$$

where

$$
\xi_{ijz}(m,n) = F'_{ij}(z_1^{-1}, z_2^{-1})z(m,n) \tag{7.14}
$$

and the noise term $w(m,n)$ is

$$
\begin{aligned}
&w(m,n) = \\
&c_1\xi_{20v}(m,n) + c_2\xi_{11v}(m,n) + c_3\xi_{02v}(m,n) + c_4\xi_{21v}(m,n) + c_5\xi_{12v}(m,n) + c_6\xi_{22v}(m,n) \\
&= \frac{\sum_{i=0}^{2}\sum_{j=0}^{2} r_{ij} z_1^{-i} z_2^{-j}}{\sum_{i=0}^{2}\sum_{j=0}^{2} q_{ij} z_1^{-i} z_2^{-j}} v(m,n)
\end{aligned}
\tag{7.15}
$$

where r_{ij}, q_{ij} are the resulting coefficients and

$$
\xi_{ijv}(m,n) = F'_{ij}(z_1^{-1}, z_2^{-1})v(m,n) \tag{7.16}
$$

7.3 Estimation methods

The discrete estimation model can be written in vector form:

$$
\begin{aligned}
\xi_{22u}(m,n) &= \varphi^T(m,n)\theta - w(m,n) \\
\varphi^T(m,n) &= [\xi_{20z}(m,n), \xi_{11z}(m,n), \xi_{02z}(m,n), \xi_{21z}(m,n), \xi_{12z}(m,n), \xi_{22z}(m,n)] \quad (7.17) \\
\theta^T &= [c_1, c_2, , c_3, c_4, c_5, c_6]
\end{aligned}
$$

For computing the filters' outputs on digital filters, we may set the boundary values and the initial values of the filters to be zero, i.e.

$$\begin{aligned} \xi_{ijz}(M_L, n) &= 0, & \xi_{iju}(M_L, n) &= 0 \\ \xi_{ijz}(M_L+1, n) &= 0, & \xi_{iju}(M_L+1, n) &= 0 \end{aligned} \tag{7.18}$$

and

$$\begin{aligned} \xi_{ijz}(m, N_S) &= 0, & \xi_{iju}(m, N_S) &= 0 \\ \xi_{ijz}(m, N_S+1) &= 0, & \xi_{iju}(m, N_S+1) &= 0 \end{aligned} \tag{7.19}$$

Therefore the outputs of the filters can be computed out using the observed input-output data respectively for $t \geq N_S T + 2$, $\quad x \geq M_L X + 2$. Then using the outputs of the two-dimensional filters for

$$t = (N_S+3)T,\ (N_S+4)T,\ \cdots, nT,\ \cdots,\ N_F T$$

and

$$x = (M_L+3)X,\ (M_L+4)X,\ \cdots,\ mX,\ \cdots,\ M_R X$$

the system parameters can be estimated directly by the LS method:

$$\hat{\theta} = \left[\sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \varphi(m,n)\varphi^T(m,n) \right]^{-1} \cdot \left[\sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \varphi(m,n)\xi_{22u}(m,n) \right] \tag{7.20}$$

If the filter parameters are selected appropriately, the effects of the measurement noise are reduced and thus the LS method can give accurate estimates in the presence of low measurement noise. It will be shown by simulation results that the parameter estimates are not sensitive to the selection of the filter parameters.

When the observed data is corrupted by high measurement noise, or the filter parameters are not correctly selected since little *a priori* knowledge of the unknown systems can be obtained, the LS method will give biased estimates due to the measurement noise effects. To obtain consistent estimates in these cases, we can introduce an IV vector $\mathbf{m}(m,n)$ which includes elements chosen to be highly correlated with the real system output, but totally uncorrelated with the measurement noise added to the real system output (Söderström and Stoica 1981). For the estimation model of equation (7.17), since the vector $\varphi(m,n)$ includes only filtered noisy observations of the system output, we can use filtered input data to construct the IV vector as

$$\mathbf{m}^T(m,n) = [\xi_{20u}(m,n), \xi_{11u}(m,n), \xi_{02u}(m,n), \xi_{21u}(m,n), \xi_{12u}(m,n), \xi_{22u}(m,n)] \tag{7.21}$$

where $\xi_{iju}(m,n)$ is defined in equation (7.12).

If

$$\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)\varphi^T(m,n)$$

exists and is nonsingular, we have the following IV estimate (Söderström and Stoica 1981):

$$\hat{\boldsymbol{\theta}} = \left[\sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)\varphi^T(m,n)\right]^{-1} \cdot \left[\sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)\xi_{22u}(m,n)\right] \quad (7.22)$$

Using equation (7.17) we have

$$\begin{aligned}
\operatorname*{plim}_{N_F\to\infty} \hat{\boldsymbol{\theta}} &= \left[\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)\varphi^T(m,n)\right]^{-1} \cdot \\
&\quad \left[\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)\ \{\ \varphi^T(m,n)\boldsymbol{\theta} - w(m,n)\ \}\right] \\
&= \boldsymbol{\theta} - \left[\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)\varphi^T(m,n)\right]^{-1} \cdot \\
&\quad \left[\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)w(m,n)\right]
\end{aligned} \quad (7.23)$$

It is known that if

$$\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)w(m,n) = 0 \quad (7.24)$$

the IV method gives a consistent estimate (Söderström and Stoica 1981).

From equation (7.15) we have

$$\begin{aligned}
&\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \sum_{m=M_L+3}^{M_R} \mathbf{m}(m,n)w(m,n) \\
&= \sum_{m=M_L+3}^{M_R} \left[\operatorname*{plim}_{N_F\to\infty} \frac{1}{N_F-(N_S+2)} \sum_{n=N_S+3}^{N_F} \mathbf{m}(m,n) \frac{\sum_{i=0}^{2}\sum_{j=0}^{2} r_{ij} z_1^{-i} z_2^{-j}}{\sum_{i=0}^{2}\sum_{j=0}^{2} q_{ij} z_1^{-i} z_2^{-j}} v(m,n)\right]
\end{aligned} \quad (7.25)$$

Since the noise vector $\mathbf{v}(n) = [v(M_L,n), v(M_L+1,n), \cdots, v(m,n), \cdots, v(M_R,n)]^T$ is assumed to be a widely stationary stochastic process with zero-mean, and the IV vector $\mathbf{m}(m,n)$

Table 7.1: Estimates of c_1 in **Example 7.1** (true value of $c_1 = 1.0$).

τ_1 / τ_2	0.1	0.2	0.3	0.4	0.6	0.8	1.0	1.2	1.6	2.0
0.1	0.4004	0.2465	0.1888	0.1578	0.1220	0.1019	0.0892	0.0805	0.0696	0.0631
0.2	1.2840	0.8303	0.7153	0.6561	0.5845	0.5376	0.5043	0.4797	0.4460	0.4243
0.3	1.5814	1.0484	0.9997	0.8965	0.8503	0.8212	0.7998	0.7835	0.7600	0.7439
0.4	1.6255	1.0940	0.9982	0.9646	0.9343	0.9174	0.9056	0.8966	0.8838	0.8750
0.5	1.6261	1.1049	1.0162	0.9870	0.9634	0.9518	0.9442	0.9387	0.9309	0.9257
0.6	1.6298	1.1113	1.0248	0.9972	0.9763	0.9670	0.9612	0.9518	0.9442	0.9387
0.7	1.6443	1.1195	1.0316	1.0039	0.9838	0.9753	0.9704	0.9672	0.9518	0.9483
0.8	1.6692	1.1309	1.0389	1.0100	0.9898	0.9811	0.9765	0.9735	0.9698	0.9675
0.9	1.7025	1.1457	1.0473	1.0163	0.9947	0.9859	0.9813	0.9784	0.9749	0.9728
1.0	1.7418	1.1635	1.0570	1.0232	0.9994	0.9904	0.9856	0.9826	0.9791	0.9770
1.2	1.8316	1.2081	1.0812	1.0397	1.0106	0.9971	0.9945	0.9906	0.9867	0.9845
1.4	1.9268	1.2634	1.1112	1.0602	1.0238	1.0102	1.0033	0.9992	0.9944	0.9918
1.6	2.0198	1.3282	1.1477	1.0851	1.0397	1.0227	1.0140	1.0088	1.0030	0.9998
1.8	2.1063	1.4013	1.1908	1.1147	1.0585	1.0372	1.0264	1.0200	1.0127	1.0088
2.0	2.1838	1.4814	1.2407	1.1493	1.0806	1.0543	1.0409	1.0320	1.0238	1.0190

includes only filtered input signals which are uncorrelated with the measurement noise, the following result holds:

$$\sum_{m=M_L+3}^{M_R}\left[\operatorname*{plim}_{N_F\to\infty}\frac{1}{N_F-(N_S+2)}\sum_{n=N_S+3}^{N_F}\cdot\,\mathbf{m}(m,n)\frac{\sum_{i=0}^{2}\sum_{j=0}^{2} r_{ij}z_1^{-i}z_2^{-j}}{\sum_{i=0}^{2}\sum_{j=0}^{2} q_{ij}z_1^{-i}z_2^{-j}}v(m,n)\right]=0 \qquad (7.26)$$

Therefore the IV estimate is consistent. Mohan and Datta (1987), Hara, Yamamoto and Ougita (1988) also considered the situations of the presence of measurement noise. In their methods, although the noise effects are reduced by multiple integrations or spline functions, it is still difficult to obtain consistent estimates in the presence of high measurement noise. Sagara and Zhao (1990) also used the IV technique combined with the two-dimensional finite integral filters, however the measurement noise was strictly restricted to be uncorrelated in time domain or in space domain. To the best of the knowledge of the authors, it is a new approach in the literature to introduce the filtered input data as instrumental variables to obtain consistent estimatess when the noise effects cannot be neglected.

Table 7.2: Estimates of $-c_3$ in **Example 7.1** (true value of $-c_3 = 2.0$).

τ_1 / τ_2	0.1	0.2	0.3	0.4	0.6	0.8	1.0	1.2	1.6	2.0
0.1	1.5070	2.2685	2.3845	2.4018	2.3923	2.3797	2.3688	2.3605	2.3490	2.3416
0.2	1.3009	1.9989	2.1202	2.1563	2.1821	2.1927	2.1985	2.2021	2.2063	2.2086
0.3	1.2185	1.8893	2.0014	2.0348	2.0591	2.0698	2.0763	2.0808	2.0867	2.0904
0.4	1.2370	1.8714	1.9720	2.0009	2.0204	2.0282	2.0327	2.0356	2.0395	2.0419
0.5	1.2370	1.8694	1.9634	1.9900	2.0071	2.0134	2.0167	2.0189	2.0216	2.0232
0.6	1.2409	1.8669	1.9592	1.9850	2.0012	2.0069	2.0098	2.0115	2.0136	2.0149
0.7	1.2289	1.8613	1.9553	1.9815	1.9978	2.0033	2.0060	2.0076	2.0094	2.0105
0.8	1.2040	1.8358	1.9505	1.9780	1.9915	2.0008	2.0035	2.0051	2.0068	2.0078
0.9	1.1692	1.8396	1.9446	1.9743	1.9926	1.9986	2.0015	2.0031	2.0049	2.0059
1.0	1.1273	1.8240	1.9376	1.9699	1.9900	1.9965	1.9996	2.0014	2.0033	2.0043
1.2	1.0303	1.7847	1.9199	1.9594	1.9841	1.9921	1.9964	1.9980	2.0003	2.0015
1.4	0.9263	1.7353	1.8973	1.9460	1.9769	1.9870	1.9917	1.9944	1.9972	1.9986
1.6	0.8236	1.6770	1.8696	1.9295	1.9681	1.9809	1.9868	1.9902	1.9937	1.9955
1.8	0.7272	1.6108	1.8368	1.9098	1.9576	1.9736	1.9811	1.9852	1.9897	1.9920
2.0	0.6396	1.5380	1.7987	1.8866	1.9452	1.9650	1.9743	1.9795	1.9851	1.9879

7.4 Illustrative examples

The recursive identification algorithms can all be described by an algorithm of the following form:

$$\begin{aligned}
\widehat{\theta}(k) &= \widehat{\theta}(k-1) + L(k)\varepsilon(k) \\
\varepsilon(k) &= \xi_{22u}(m,n) - \phi^T(m,n)\widehat{\theta}(k-1) \\
L(k) &= \frac{P(k-1)\psi(m,n)}{\rho(k) + \phi^T(m,n)P(k-1)\psi(m,n)} \qquad (7.27) \\
P(k) &= \frac{1}{\rho(k)}\left[P(k-1) - \frac{P(k-1)\psi(k)\phi^T(m,n)P(k-1)}{\rho(k) + \phi^T(m,n)P(k-1)\psi(m,n)}\right]
\end{aligned}$$

where k is the recursion number and $\rho(k)$ is the forgetting factor. Here the forgetting factor is chosen to be

$$\rho(k) = (1 - 0.01)\rho(k-1) + 0.01, \qquad \rho(0) = 0.95 \qquad (7.28)$$

The initial setting of the algorithm is

$$\begin{aligned} \hat{\theta}(0) &= \mathbf{0} \\ P(0) &= 10^6\mathbf{I} \end{aligned} \tag{7.29}$$

Now the methods under discussion will be obtained as special cases of the recursive algorithm.

The LS method is obtained with

$$\begin{aligned} \phi(m,n) = \varphi(m,n) \\ \psi(m,n) = \varphi(m,n) \end{aligned} \tag{7.30}$$

and the proposed IV method is obtained with

$$\begin{aligned} \phi(m,n) = \varphi(m,n) \\ \psi(m,n) = \mathbf{m}(m,n) \end{aligned} \tag{7.31}$$

Some illustrative examples are given here to illustrate the effectiveness of the proposed identification algorithms.

Example 7.1: Effects of the pre-filters (example 1).

The following system is considered in order to investigate the effects of the selection of the filter parameters (τ_1, τ_2) using the LS method:

$$\begin{aligned} & c_1 \frac{\partial^2 y(x,t)}{\partial t^2} + c_3 \frac{\partial^2 y(x,t)}{\partial x^2} = u(x,t) \qquad (t \geq 0, \;\; 0 \leq x \leq 2) \\ & c_1 = 1, \;\; c_3 = -2 \end{aligned} \tag{7.32}$$

The input and the output are

$$\begin{aligned} u(x,t) &= -4\ (x^2+4)\cos(t) \\ y(x,t) &= 4\ x^2\cos(t) \end{aligned} \tag{7.33}$$

The sampling intervals are taken to be

$$\begin{aligned} X &= 0.25 \\ T &= 0.1 \end{aligned} \tag{7.34}$$

The input and the output are observed at nine points:

$$x = 0.00,\ 0.25,\ 0.50,\ 0.75,\ 1.00,\ 1.25,\ 1.50,\ 1.75,\ 2.00$$

And at each observation point the output is corrupted by a white noise with zero-mean and variance $\sigma_x^2(\sigma_x = 0.1x^2)$. The results are shown in Table 7.1 and Table 7.2. It is shown that

Table 7.3: Estimates of c_1 in **Example 7.2** (true value of $c_1 = 1.0$).

τ_1 / τ_2	0.1	0.2	0.3	0.4	0.6	0.8	1.0	1.2	1.6	2.0
0.1	0.7428	0.6926	0.6791	0.6798	0.6839	0.6868	0.6887	0.6900	0.6917	0.6927
0.2	1.0745	0.9754	0.9597	0.9587	0.9609	0.9629	0.9644	0.9655	0.9669	0.9678
0.3	1.1262	1.0124	0.9952	0.9936	0.9956	0.9976	0.9991	1.0002	1.0017	1.0026
0.4	1.1511	1.0259	1.0060	1.0041	1.0062	1.0085	1.0102	1.0115	1.0131	1.0142
0.5	1.1726	1.0371	1.0134	1.0110	1.0136	1.0163	1.0184	1.0199	1.0219	1.0232
0.6	1.1926	1.0495	1.0209	1.0180	1.0212	1.0245	1.0270	1.0289	1.0313	1.0329
0.7	1.2108	1.0636	1.0295	1.0260	1.0299	1.0340	1.0371	1.0393	1.0423	1.0442
0.8	1.2266	1.0793	1.0394	1.0351	1.0400	1.0451	1.0488	1.0514	1.0550	1.0573
0.9	1.2400	1.0962	1.0506	1.0457	1.0516	1.0576	1.0620	1.0651	1.0693	1.0719
1.0	1.2511	1.1378	1.0630	1.0575	1.0645	1.0715	1.0765	1.0801	1.0849	1.0879

Table 7.4: Estimates of $-c_3$ in **Example 7.2** (true value of $-c_3 = 1.0$).

τ_1 / τ_2	0.1	0.2	0.3	0.4	0.6	0.8	1.0	1.2	1.6	2.0
0.1	0.6466	1.6165	1.7472	1.7519	1.7289	1.7101	1.6971	1.6877	1.6755	1.6678
0.2	0.3862	0.9823	1.0617	1.0675	1.0574	1.0478	1.0409	1.0359	1.0293	1.0249
0.3	0.3374	0.8998	0.9755	0.9822	0.9739	0.9654	0.9592	0.9546	0.9485	0.9446
0.4	0.3042	0.8682	0.9489	0.9566	0.9481	0.9392	0.9327	0.9279	0.9214	0.9172
0.5	0.2662	0.8381	0.9292	0.9381	0.9287	0.9187	0.9113	0.9058	0.8985	0.8938
0.6	0.2663	0.8025	0.9077	0.9183	0.9073	0.8955	0.8868	0.8804	0.8719	0.8664
0.7	0.1887	0.7610	0.8824	0.8950	0.8818	0.8678	0.8575	0.8500	0.8400	0.8336
0.8	0.1556	0.7143	0.8530	0.8676	0.8518	0.8353	0.8232	0.8144	0.8027	0.7954
0.9	0.1277	0.6640	0.8194	0.8362	0.8174	0.7981	0.7842	0.7741	0.7607	0.7523
1.0	0.1047	0.6117	0.7822	0.8010	0.7791	0.7570	0.7412	0.7298	0.7148	0.7055

Table 7.5: Estimates of **Example 7.3** by LS method ($\tau_1 = 0.1, \tau_2 = 0.1$).

σ_0	0.00	0.03	0.05	0.10	0.20	0.30	0.40	0.50
$\hat{c}_1$	0.9916	0.9701	0.9284	0.7428	0.4102	0.2345	0.1465	0.0988
$\hat{c}_3$	−1.0209	−0.9394	−0.8527	−0.6466	−0.3468	−0.1962	−0.1217	−0.0815

Table 7.6: Estimates of **Example 7.3** by IV method ($\tau_1 = 0.1, \tau_2 = 0.1$).

σ_0	0.00	0.03	0.05	0.10	0.20	0.30	0.40	0.50
$\hat{c}_1$	0.9916	0.9916	0.9916	0.9917	0.9918	0.9916	0.9920	0.9921
$\hat{c}_3$	−1.0209	−1.0210	−1.0211	−1.0212	−1.0214	−1.0217	−1.0219	−1.0222

the LS estimates are still accurate in a relative wide range of (τ_1, τ_2), where the noise effects are reduced sufficiently. It is obvious that the estimates are not sensitive to the selection of the filter parameters.

Example 7.2: Effects of the pre-filters (example 2).

Another system which was studied by Hara, Yamamoto and Ougita (1988) is also considered using the LS method like **Example 7.1**:

$$c_1 \frac{\partial^2 y(x,t)}{\partial t^2} + c_3 \frac{\partial^2 y(x,t)}{\partial x^2} = u(x,t) \qquad (t \geq 0, \quad 0 \leq x \leq 1)$$
$$c_1 = 1, \quad c_3 = -1 \tag{7.35}$$

The input and the output are

$$\begin{aligned} u(x,t) &= -13 \exp(-x)\cos(1.5t) - 9.32 \exp(-0.5x)\cos(2.1t) \\ y(x,t) &= 4 \exp(-x)\cos(1.5t) + 2 \exp(-0.5x)\cos(2.1t) \end{aligned} \tag{7.36}$$

The sampling intervals are the same as those in **Example 7.1**. The input and the output are observed at only five points:

$$x = 0.00, \ 0.25, \ 0.50, \ 0.75, \ 1.00$$

And at each observation point the output is corrupted by a white noise with zero-mean and variance $\sigma_x^2(\sigma_x = 0.1 \exp(-0.5x))$. The results are shown in Table 7.3 and Table 7.4. It is shown that the LS estimates are still acceptable if the filter parameters are selected appropriately. However, it seems that the LS estimates of **Example 7.2** are more sensitive to the selection of the filter parameters than the results of **Example 7.1**. From Tables 7.1~7.4, we know that when the filter parameters are selected unsuccessfully, the parameter estimates are greatly biased. In some practical situations, we may fail to design the digital filters successfully, since little *a priori* knowledge of the unknown systems can be obtained. And some times, the output measurement may be corrupted by a high measurement noise. In these cases, the LS method is not efficient.

Example 7.3: Comparison of the IV method with the LS method.

The system in **Example 7.2** is considered in order to compare the LS method and the IV method under the same situations. The filter parameters are selected to be

$$\begin{aligned} \tau_1 &= 0.1 \\ \tau_2 &= 0.1 \end{aligned} \tag{7.37}$$

Considering the results of **Example 7.2**, the selection is a very bad one. At each observation point the output is corrupted by a white noise with zero-mean and variance σ_x^2($\sigma_x = \sigma_0 \exp(-0.5x)$). Simulation experiments are carried out for different values of σ_0. The results are shown in Table 7.5 and Table 7.6. It is shown that even when a bad selection of the filter parameters is taken, the IV method is still effective in the presence of high measurement noise. Therefore the IV method is quite efficient and much superior to the LS method. Furthermore, we can conclude that if the IV method is used, the selection of the filter parameters becomes to be unimportant.

7.5 Conclusion

In this paper, the two-dimensional digital IIR filtering approach to the problem of identification of second-order DPSs in the presence of measurement noise has been considered by the authors.

This approach need not estimate the initial and boundary conditions in contrast to those using orthogonal functions. Compared with the methods using orthogonal functions, the methods proposed here are thought to be simpler, more convenient for on-line implementation by digital computers.

It has been shown that the two-dimensional IIR digital filters can reduce the effects of the measurement noise, and therefore the LS method is still efficient if the filters parameters are selected appropriately. In practice, the estimates are not so sensitive to the selection of the filter parameters in some cases.

An IV method with filtered input signals as instrumental variables has also been proposed when the filter is not designed appropriately or the output is corrupted by high noise. It has been shown that the IV method can give consistent estimates without restrictions about the stochastic properties of measurement noise. Therefore the IV method is thought to be suitable for practical situations.

The proposed algorithms are expected to be applicable to on-line identification of real systems.

Chapter 8

Implementation of Multi-Rate Model Reference Adaptive Control for Continuous-Time Systems

8.1 Introduction

The two major forms of the model reference adaptive control system (MRACS) are continuous-time algorithms and discrete-time algorithms. The continuous-time algorithms are developed by generating adaptive feedback laws to let the error between the system output and the reference model output be zero under stability considerations, whereas the discrete-time algorithms are usually viewed as a combination of a parameter estimator and a feedback controller (Egardt 1979). Reasons for the popularity of discrete-time algorithms stem from rapid development and wide uses of digital computers, and also from the considerable flexibility in choice of both the controller and estimator. However, it is known that the parameters of the discrete-time models of continuous-time processes do not have physical interpretations and may depend on sampling interval. Furthermore, it was pointed by Åström, Hagander and Sternby (1984) that for a discrete-time system sampled at a fast sampling rate with zero-order hold (ZOH) input, the zeros may be outside or on the unit circle. Thus the discrete-time adaptive control algorithms are inappropriate in some practical situations. Gawthrop (1980) pointed out that the 'hybrid' approach to adaptive control is superior to the wholly discrete-time approach. And it was shown by Goodwin, Lozano-Leal, Mayne and Middelton (1986) that using the delta operator which approximates the differentiations directly for a fast sampling rate, there is no essential difference between continuous and discrete MRACS and therefore the nonminimum phase zero problem is easily avoided.

This work is perhaps the best known in the literature. Most recently, the limiting-zero model and the modified delta operator model has been proposed by Mizuno and Fujii (1991). Application of the bilinear transformation based on the trapezoidal rule to implementation of indirect multi-rate MRACS was investigated by Sagara and Yang (1988).

In recent years, considerable attention has been given to the development of algebraic methods for system analysis, identification, order reduction and control by using orthogonal functions. Walsh functions and BPFs have been widely used to control theory, because digital computer control is usually implemented on the basis of the discrete-time models where the step functions produced by sampling and holding can be precisely expressed by finite Walsh functions or BPFs. The BPFs have received more and more attention due to the convenience and simplicity of related computations. Numerous papers have been published on system identification (Palanisamy and Bhattacharya 1981, Cheng and Hsu 1982, Hsu and Cheng 1982, Sagara, Yuan and Wada 1988a etc.), system analysis (Shieh, Yeung and McInnis 1978, Rao and Srinivasan 1978,1980, Chen and Lee 1982, Kawaji 1983 etc.), and optimal control (Hsu and Cheng 1981, Zhu and Lu 1988 etc.). Application of the BPFs to implicit MRACS of Egardt's algorithm (1979) was first described by Sagara, Yuan and Wada (1988b). Most recently, a detailed description of an explicit self-tuning algorithm for continuous systems has been proposed by Patra and Rao (1989) successfully, by making use of the so-called one-shot operational matrix of the BPFs for repeated integration. And it was shown by an example that the continuous-time approach is superior in performance to the corresponding discrete-time version (Patra and Rao 1989).

Motivated by the previous works(Goodwin, Lozano-Leal, Mayne and Middelton 1986, Sagara, Yuan and Wada 1988b, Sagara and Yang 1988, Patra and Rao 1989), the purpose of this chapter is to show that the bilinear transformation based on the BPFs can be applied to adaptive systems with excellent precision and simple related calculations, and to show that the method using the BPFs is much more satisfactory than the others. In this chapter, the BPFs are further investigated in the sense of the bilinear transformation which is closely related with the trapezoidal integrating rule. The discretization or approximation techniques for continuous systems using the well-known delta operator, and the bilinear transformation based on the BPFs or the trapezoidal rule are discussed. It is shown that the bilinear transformation gives more accurate approximations than the delta operator. And it is shown that the approximated discrete-time model obtained by the bilinear transformation based on the BPFs can be viewed as the Páde approximation of the ZOH sampling of the controlled system model, whereas the approximated discrete-time model obtained by the bilinear transformation based on the trapezoidal rule neglects the fact that the input control signal $u(t)$ is constant between the sampling instants when a ZOH is used. Then an algorithm of multi-rate indirect model reference adaptive control for SISO continuous systems is addressed. The algorithm includes: a recursive LS type parameter estimator, a continuous system model and a controller designed in continuous-time domain. The recursive parameter estimator estimates the parameters of the continuous process using sampled system input-output data. The system model with the continuous-time state variable filters, and the continuous-time controller are discretized by the delta operator, the bilinear transformation based on the BPFs or the trapezoidal rule. Therefore the whole algorithm can be implemented on a digital computer. To reduce the computational burden, the algorithm is implemented in a multi-rate manner with a small sampling interval of the system signals and a relatively large parameter estimation interval. Numerical results for an unstable minimum phase continuous system show that the control performance is very excellent. It is also shown that the algorithm gives accurate estimates of the parameters of the unstable continuous system. Comparison of the discretization methods for the adaptive system using the BPFs, the trapezoidal integrating rule and the delta operator is discussed through simulation study. It seems that the BPF method is the most effective one since the BPFs give excellent approximations of the signals of the digital controlled continuous system.

8.2 Brief review of bilinear transformation, delta operator and BPFs

We define the integration operator p^{-1} to be

$$p^{-1}f(t) = \int_{t_0}^{t} f(t)\,dt \tag{8.1}$$

It is well-known that for a small sampling interval, using the trapezoidal rule , p^{-1} is represented to be (Sinha and Zhou 1983, Krishna 1988)

$$p^{-1} = \frac{T}{2}\frac{1+z^{-1}}{1-z^{-1}} \tag{8.2}$$

where T is the sampling interval which is assumed to be sufficiently small and z^{-1} is the shift operator which lets

$$z^{-1}f(kT) = f(kT-T) \tag{8.3}$$

Equivalently the differential operator p can be replaced by the well-known bilinear transformation:

$$p = \frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} \tag{8.4}$$

which maps the semi-infinite left half of the $p-$plane into the unit circle of the $z-$plane and thus keeps the stability of the discretized system.

Consider the following continuous system

$$\begin{aligned} D(p)y(t) &= N(p)u(t) \\ D(p) &= \sum_{i=0}^{n} \alpha_i p^{n-i} \qquad (\alpha_0 = 1) \\ N(p) &= \sum_{i=0}^{n} \beta_i p^{n-i} \end{aligned} \tag{8.5}$$

where $y(t)$, $u(t)$ are system output and input respectively.

Replacing the differential operator p with the bilinear transformation and writing $u(kT), y(kT)$ as $u(k)$, $y(k)$ for convenience of notation, we have (Krishna 1988)

$$\sum_{i=0}^{n} \alpha_i \left(\frac{T}{2}\right)^i (1+z^{-1})^i (1-z^{-1})^{n-i} y(k) = \sum_{i=0}^{n} \beta_i \left(\frac{T}{2}\right)^i (1+z^{-1})^i (1-z^{-1})^{n-i} u(k) \tag{8.6}$$

An alternative approach to discretization is the well-known delta operator (Euler operator) which replaces the differential operator p by δ as (Kuo 1980, Goodwin, Lozano-Leal, Mayne and Middelton 1986, Hori, Nikiforuk and Kanai 1988)

$$\delta = \frac{z-1}{T} \tag{8.7}$$

In this case the discrete-time model becomes to be

$$\sum_{i=0}^{n} \alpha_i \left(\frac{z-1}{T}\right)^{n-i} y(k) = \sum_{i=0}^{n} \beta_i \left(\frac{z-1}{T}\right)^{n-i} u(k) \tag{8.8}$$

Application of the delta operator to MRACS was proposed by Goodwin, Lozano-Leal, Mayne and Middelton (1986) successfully. However, it was also pointed out by Kuo (1980) that the delta operator based on the advanced rectangular approximated integration does not give so accurate approximations as the bilinear transformation which is based on the trapezoidal approximated integration. For a small sampling interval which approaches zero, we have

$$\begin{aligned} \frac{z-1}{T} &= \frac{e^{pT}-1}{T} = p + \frac{p^2T}{2!} + \frac{p^3T^2}{3!} + \cdots \\ &= p + O(T) \end{aligned} \tag{8.9}$$

and

$$
\begin{aligned}
\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} &= \frac{2}{T}\frac{e^{pT}-1}{e^{pT}+1} = \frac{2}{T}\frac{pT+\frac{p^2T^2}{2!}+\frac{p^3T^3}{3!}+\cdots}{2+pT+\frac{p^2T^2}{2!}+\frac{p^3T^3}{3!}+\cdots} \\
&= \frac{p\left(1+\frac{pT}{2}+\frac{p^2T^2}{4}\right)-\frac{p^3T^2}{12}+\cdots}{\left(1+\frac{pT}{2}+\frac{p^2T^2}{4}\right)+\frac{1}{2}\frac{p^3T^3}{3!}+\cdots} \\
&= p+O(T^2)
\end{aligned}
\tag{8.10}
$$

Therefore, the bilinear transformation gives more accurate approximations than the delta operator. Numerical examples will show that the bilinear transformation is more satisfactory than the delta operator in implementation of the adaptive system.

It can be shown that when the sampling interval decreases, both the discrete-time models by the bilinear transformation and by the delta operator converge in some sense to the original continuous-time system described in equation (8.5) (Hori, Mori and Nikiforuk 1989). For detailed discussions about discrete-time models of a continuous system, the reader is referred to the work of Hori, Mori and Nikiforuk (1989).

A set of BPFs $\phi_k(t),\ k=1,\cdots,N$ is defined in the interval $[0,NT)$ by

$$
\phi_k(t)=\begin{cases}1 & (k-1)T\le t<kT \\ 0 & \text{otherwise}\end{cases}
\tag{8.11}
$$

An arbitrary function $f(t)$ which is square integrable in the interval $[0,NT)$ can be approximated by the BPFs:

$$
f(t)\approx\sum_{k=1}^{N}\bar{f}(k)\phi_k(t)=\bar{\mathbf{f}}^T\Phi_N(t)
\tag{8.12}
$$

where

$$
\begin{aligned}
\bar{\mathbf{f}} &= [\bar{f}(1),\bar{f}(2),\cdots,\bar{f}(N)]^T \\
\Phi_N(t) &= [\phi_1(t),\phi_2(t),\cdots,\phi_N(t)]^T
\end{aligned}
\tag{8.13}
$$

$\bar{f}(k)$ which is called the kth BPF expansion coefficient of the function $f(t)$ is chosen such that

$$
\epsilon=\int_0^{NT}[f(t)-\bar{\mathbf{f}}^T\Phi_N(t)]^2\,dt
\tag{8.14}
$$

is minimized. Thus $\bar{f}(k)$ is given by

$$
\begin{aligned}
\bar{f}(k) &= \frac{1}{T}\int_{(k-1)T}^{kT}f(t)\,dt \\
&\approx \frac{1}{2}[f(k-1)+f(k)]
\end{aligned}
\tag{8.15}
$$

Remark 8.1: If $f(t)$ is the input signal of the continuous-time system preceded by a ZOH, which is held constant between the sampling instants:

$$f(t) = f(kT - T), \qquad (k-1)T \le t < kT \tag{8.16}$$

then the BPF value $\bar{f}(k)$ of piece-wise constant $f(t)$ is given by

$$\bar{f}(k) = f(k-1) \tag{8.17}$$

It is known that the integral of the BPF vector

$$\int_0^{NT} \Phi_N(t)\, dt \approx \mathbf{P}_N \Phi_N(t) \tag{8.18}$$

where $\mathbf{P}_N$ is the $N \times N$ operational matrix of integration:

$$\mathbf{P}_N = \frac{T}{2} \begin{bmatrix} 1 & 2 & 2 & \cdots & 2 \\ & 1 & 2 & \cdots & 2 \\ & & 1 & \ddots & \vdots \\ & \mathbf{0} & & \ddots & 2 \\ & & & & 1 \end{bmatrix} \tag{8.19}$$

With the BPFs, we have

$$p^{-1} f(t) \;=\; \int_0^{NT} f(t)\, dt \approx p^{-1} \bar{\mathbf{f}}^T \Phi_N(t) \tag{8.20}$$

and with the operational matrix of integration $\mathbf{P}_N$, we have

$$\int_0^{NT} f(t)\, dt \;\approx\; \bar{\mathbf{f}}^T \mathbf{P}_N \Phi_N(t) \tag{8.21}$$

Hence we have

$$p^{-1} \bar{\mathbf{f}}^T \Phi_N(t) \;=\; \bar{\mathbf{f}}^T \mathbf{P}_N \Phi_N(t) \tag{8.22}$$

The kth $(k = 1, 2, \cdots, N)$ elements of $p^{-1}\bar{\mathbf{f}}$ can be computed by

$$\begin{aligned} p^{-1}\bar{f}(1) &= \frac{T}{2}\bar{f}(1) \\ p^{-1}\bar{f}(k) &= p^{-1}\bar{f}(k-1) + \frac{T}{2}\left[\bar{f}(k) + \bar{f}(k-1)\right] \end{aligned} \tag{8.23}$$

Then we have

$$p^{-1} \;=\; \frac{T}{2}\frac{1 + z^{-1}}{1 - z^{-1}} \tag{8.24}$$

or equivalently

$$p \;=\; \frac{2}{T}\frac{1 - z^{-1}}{1 + z^{-1}} \tag{8.25}$$

Therefore the BPF method is closely related to the trapezoidal method. Similarly to the trapezoidal method, replacing the differential operator p by the bilinear transformation based on the BPFs, we have the following piecewise constant system model:

$$\sum_{i=0}^{n}\alpha_i\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}\bar{y}(k)=\sum_{i=0}^{n}\beta_i\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n-i}\bar{u}(k) \quad (8.26)$$

where $\bar{u}(k),\bar{y}(k)$ are the kth BPF expansion coefficient of $u(k),y(k)$ respectively:

$$\begin{aligned}\bar{u}(k) &= \frac{1}{T}\int_{(k-1)T}^{kT}u(t)\,dt\\ \bar{y}(k) &= \frac{1}{T}\int_{(k-1)T}^{kT}y(t)\,dt\end{aligned} \quad (8.27)$$

Comment 8.1: It is clear that approximation of a continuous system using the BPFs is very similar to that using the trapezoidal rule. It should be noted that the discrete-time models of equation (8.6) and equation (8.8) represent the relations of the discrete values of the system input-output signals at sampling time instants while the piecewise constant model of equation (8.26) represents the relations of the average values of the input-output signals over one sampling interval. For modeling of a digital controlled system, we would use the BPFs rather than the trapezoidal rule and the delta operator, since the step functions produced by sampling and holding employed in digital control can be precisely expressed by the BPF coefficients, while the trapezoidal rule and the delta operator only give discrete values of continuous signals at sampling time instants.

8.3 Relation between the BPF model and the ZOH sampled model

To show the advantage of the BPF method, the relation of the approximated discretized model with the common discrete-time model is analyzed here.

Consider a strictly proper SISO system

$$\begin{aligned}A(p)y(t) &= B(p)u(t)\\ A(p) &= \sum_{i=0}^{n}a_ip^{n-i} \quad (a_0=1)\\ B(p) &= \sum_{i=1}^{n}b_ip^{n-i}\end{aligned} \quad (8.28)$$

Assuming that all the initial values of the signals of the system are zero, we can write it into the following companion form:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}^T\mathbf{x}(t) \end{aligned} \tag{8.29}$$

where

$$\begin{aligned} \mathbf{x}(t) &= [x_1(t), x_2(t), \cdots, x_n(t)]^T \\ \mathbf{A} &= \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & 0 & 1 \\ a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix} \\ \mathbf{b} &= [0, 0, \cdots, 0, 1]^T \\ \mathbf{c} &= [b_n, b_{n-1}, \cdots, b_2, b_1]^T \end{aligned} \tag{8.30}$$

The transfer function of the companion form is given as

$$G(p) = \mathbf{c}^T(p\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \tag{8.31}$$

A common situation in computer control is that the input signal $u(t)$ is constant between the sampling instants:

$$u(t) = u(kT - T), \qquad (k-1)T \leq t < kT \tag{8.32}$$

hence the BPF value $\bar{u}(k)$ of piece-wise constant input signal $u(t)$ is given by

$$\bar{u}(k) = u(k-1) \tag{8.33}$$

Then the ZOH equivalent of system (8.29) is given as (Åström and Wittenmark 1984)

$$\begin{cases} \mathbf{X}(k) &= \Phi\mathbf{X}(k-1) + \mathbf{r}\bar{u}(k) \\ y(k) &= \mathbf{C}^T\mathbf{X}(k) \end{cases} \tag{8.34}$$

where

$$\Phi = e^{\mathbf{A}T}, \; \mathbf{r} = (\Phi - I)A^{-1}\mathbf{b} \tag{8.35}$$

The transfer function of sampled system model (8.34) is given as

$$y(k) = \mathbf{c}^T(\mathbf{I} - z^{-1}\Phi)^{-1}\mathbf{r}\bar{u}(k) \tag{8.36}$$

Notice that equation (8.34) does not involve any approximations. It gives the exact values of the state values and the output at the sampling instants because the control signal is constant between the sampling instants. The model (8.34) is therefore called a ZOH sampling of system (8.29) or the ZOH equivalent of (8.29).

For a sufficiently small sampling interval T, using the famous Páde approximation we have (Haykin 1972, Haberland and Rao 1973, Sinha and Lastman 1981)

$$e^{\mathbf{A}T} \approx \left(\mathbf{I}+\frac{1}{2}\mathbf{A}T\right)\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1} = \left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1}\left(\mathbf{I}+\frac{1}{2}\mathbf{A}T\right) \tag{8.37}$$

and

$$\begin{aligned}
\mathbf{r} &\approx \left[\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1}\left(\mathbf{I}+\frac{1}{2}\mathbf{A}T\right)-\mathbf{I}\right]\mathbf{A}^{-1}\mathbf{b} \\
&= \left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1}\left[\left(\mathbf{I}+\frac{1}{2}\mathbf{A}T\right)-\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)\right]\mathbf{A}^{-1}\mathbf{b} \qquad (8.38) \\
&= T\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1}\mathbf{b}
\end{aligned}$$

Then the transfer function of sampled system model (8.36) is approximated as

$$\begin{aligned}
y(k) &\approx \mathbf{c}^T\left[\mathbf{I}-z^{-1}\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1}\left(\mathbf{I}+\frac{1}{2}\mathbf{A}T\right)\right]^{-1} T\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)^{-1}\mathbf{b}\bar{u}(k) \\
&= \mathbf{c}^T\left[\left(\mathbf{I}-\frac{1}{2}\mathbf{A}T\right)-z^{-1}\left(\mathbf{I}+\frac{1}{2}\mathbf{A}T\right)\right]^{-1} T\,\mathbf{b}\bar{u}(k) \\
&= \mathbf{c}^T\left[\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\mathbf{I}-\mathbf{A}\right]^{-1}\left[(1+z^{-1})\frac{T}{2}\right]^{-1} T\,\mathbf{b}\bar{u}(k) \qquad (8.39) \\
&= \frac{2}{1+z^{-1}}\mathbf{c}^T\left[\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\mathbf{I}-\mathbf{A}\right]^{-1}\mathbf{b}\bar{u}(k)
\end{aligned}$$

Consequently, we have

$$\begin{aligned}
\bar{y}(k) = \frac{1+z^{-1}}{2}y(k) &= \mathbf{c}^T\left[\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\mathbf{I}-\mathbf{A}\right]\mathbf{b}\bar{u}(k) \\
&= G(p)\Big|_{p=\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}}\bar{u}(k)
\end{aligned} \tag{8.40}$$

It is interesting to find that equation (8.40) is equivalent to the BPF model (8.27), i.e. the BPF model can be viewed as the Páde approximation of a ZOH sampling of the controlled system model.

On the other hand, the trapezoidal model (8.6) is equivalent to

$$y(k) = G(p)\Big|_{p=\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}}\bar{u}(k) \tag{8.41}$$

which represents the relations of the discrete values of the system input-output signals at sampling time instants. Hence we can say that the trapezoidal model neglects the fact that $u(t)$ is constant between the sampling instants. Therefore, for digital control systems with ZOH, the BPF model is more appropriate than the trapezoidal model.

8.4 Basic design of the indirect MRACS

The unknown minimum phase SISO system is described by

$$\begin{aligned} A(p)y(t) &= B(p)u(t) \\ A(p) &= \sum_{i=0}^{n} a_i p^{n-i} \quad (a_0 = 1) \\ B(p) &= \sum_{i=0}^{m} b_i p^{m-i} \\ & \qquad n > m \end{aligned} \tag{8.42}$$

The desired reference model is given by

$$\begin{aligned} A^m(p)y_m(t) &= B^m(p)u_m(t) \\ A^m(p) &= \sum_{i=0}^{n} a_i^m p^{n-i} \quad (a_0^m = 1) \\ B^m(p) &= \sum_{i=0}^{m} b_i^m p^{m-i} \end{aligned} \tag{8.43}$$

where $u_m(t)$ is the command input.

The block diagram of the closed-loop of the MRACS is shown in Figure 8.1 where $R(p), S(p)$ and $T(p)$ are polynomials in the differential operator. The stable polynomial $T(p)$ can be chosen freely without changing the closed-loop transfer function. However, it is of importance for the transient properties and the effect of disturbance (Egardt 1979). Usually, $T(p)$ is chosen to be

$$T(p) = p^{n_T} + t_1 p^{n_T - 1} + \cdots + t_{n_T} \qquad n_T \geq n - m - 1 \tag{8.44}$$

If the equation

$$T(p)A^m(p) = A(p)S(p) + R(p) \tag{8.45}$$

has the unique solutions $R(p)$ and $S(p)$, defined by

$$\begin{aligned} R(p) &= r_1 p^{n-1} + \cdots + r_n \\ S(p) &= p^{n_T} + s_1 p^{n_T - 1} + \cdots + s_{n_T} \end{aligned} \tag{8.46}$$

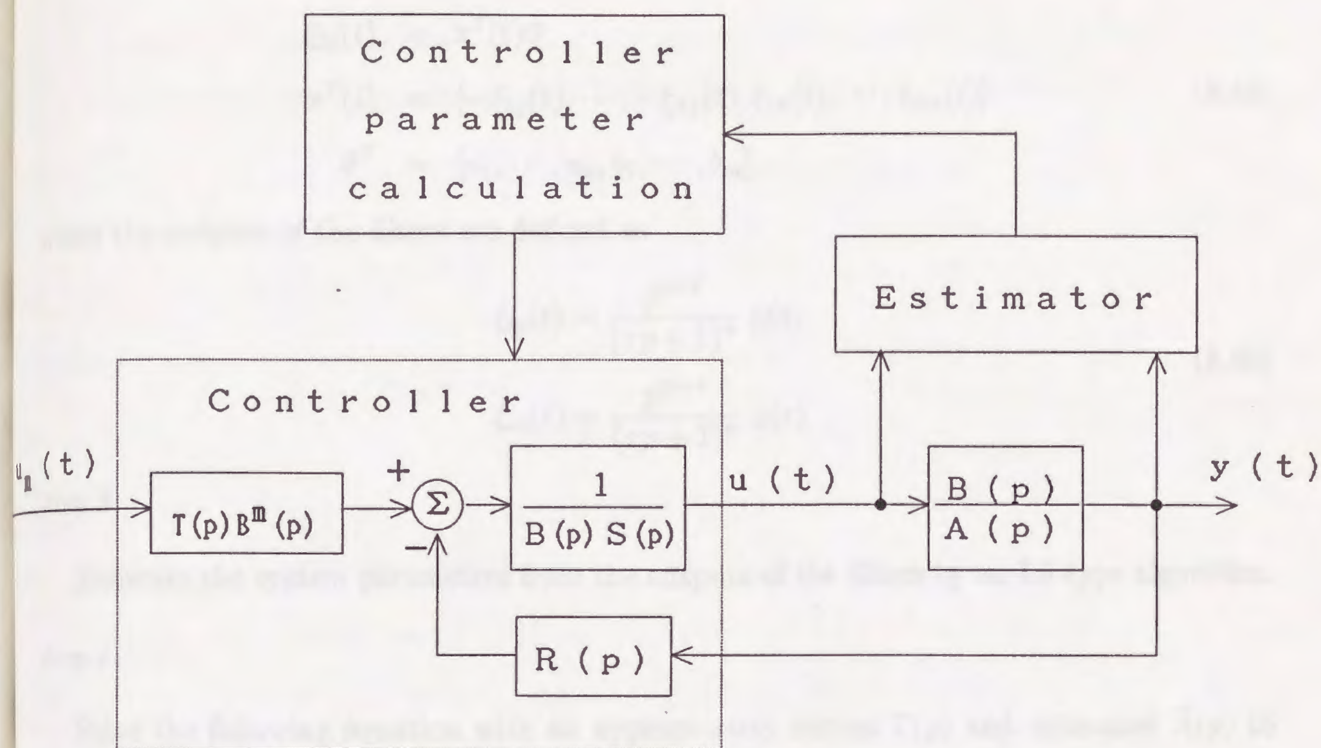


Figure 8.1: Block diagram of the closed-loop of the adaptive system.

then the closed-loop transfer function becomes to be

$$\frac{y(t)}{u_m(t)} = \frac{T(p)B^m(p)}{A(p)S(p) + R(p)} = \frac{B^m(p)}{A^m(p)} \tag{8.47}$$

The adaptive algorithm consists of the following steps:

Step 1

Introduce a low-pass filter to remove the direct signal derivatives:

$$F(p) = \frac{1}{(\tau p + 1)^n} \tag{8.48}$$

where τ is the time constant which determines the pass-band of $F(p)$.

Step 2

Filter both sides of the system model to get

$$\begin{aligned} \xi_{0y}(t) &= \mathbf{z}^T(t)\boldsymbol{\theta} \\ \mathbf{z}^T(t) &= [-\xi_{1y}(t), \cdots, -\xi_{ny}(t), \xi_{1u}(t), \cdots, \xi_{mu}(t)] \\ \boldsymbol{\theta}^T &= [a_1, \cdots, a_n, b_1, \cdots, b_m] \end{aligned} \tag{8.49}$$

where the outputs of the filters are defined as

$$\begin{aligned} \xi_{iy}(t) &= \frac{p^{n-i}}{(\tau p + 1)^n}\, y(t) \\ \xi_{iu}(t) &= \frac{p^{m-i}}{(\tau p + 1)^n}\, u(t) \end{aligned} \tag{8.50}$$

Step 3

Estimate the system parameters from the outputs of the filters by an LS type algorithm.

Step 4

Solve the following equation with an appropriately chosen $T(p)$ and estimated $\widehat{A}(p)$ to have $\widehat{R}(p)$ and $\widehat{S}(p)$:

$$T(p)A^m(p) = \widehat{A}(p)\widehat{S}(p) + \widehat{R}(p) \tag{8.51}$$

Step 5

Generate the control signal $u(t)$:

$$u(t) = \frac{1}{\widehat{B}(p)\widehat{S}(p)}\{T(p)B^m(p)u_m(t) - \widehat{R}(p)y(t)\ \} \tag{8.52}$$

Step 6

Go to *Step 2.*

8.5 Digital implementation of the algorithm

In this section we will describe the digital implementation techniques of the MRACS using the bilinear transformation based on the BPFs. To reduce the computational burden, the algorithm is implemented in a multi-rate manner with a very small sampling interval of the system signals and a relatively large parameter estimation interval.

When the adaptive system has been designed in continuous-time domain, the digital implementation procedure of the adaptive algorithm by the BPFs includes the following.

Step 1

Compute the outputs of the following digital low-pass state variable filters derived from the bilinear transformation based on the BPFs with a small sampling interval T:

$$\begin{aligned}\bar{\xi}_{iy}(k) &= F'(z^{-1})\left(\frac{T}{2}\right)^{i}(1+z^{-1})^{i}(1-z^{-1})^{n-i}\bar{y}(k)\\ \bar{\xi}_{iu}(k) &= F'(z^{-1})\left(\frac{T}{2}\right)^{n-m+i}(1+z^{-1})^{n-m+i}(1-z^{-1})^{m-i}\bar{u}(k)\end{aligned} \tag{8.53}$$

where

$$F'(z^{-1}) = \frac{1}{\left[\tau(1-z^{-1})+\frac{T}{2}(1+z^{-1})\right]^{n}} \tag{8.54}$$

Step 2

Construct the estimation model

$$\begin{aligned}\bar{\xi}_{0y}(k) &= \bar{\mathbf{z}}^{T}(k)\boldsymbol{\theta}\\ \bar{\mathbf{z}}^{T}(k) &= [-\bar{\xi}_{1y}(k),\cdots,-\bar{\xi}_{ny}(k),\bar{\xi}_{1u}(k),\cdots,\bar{\xi}_{mu}(k)]\\ \boldsymbol{\theta}^{T} &= [a_1,\cdots,a_n,b_1,\cdots,b_m]\end{aligned} \tag{8.55}$$

and then estimate the parameters by the recursive LS algorithm.

The estimation procedure need not be carried out at every sampling time instant. Define an estimation interval T_e to be

$$T_e = lT \tag{8.56}$$

where l is a natural number. The system parameters are estimated at every estimation time instant ($t = k_eT_e$) by the following algorithm:

$$\begin{aligned}\widehat{\theta}(k_e) &= \widehat{\theta}(k_e-1)+L(k_e)\varepsilon(k_e)\\ \varepsilon(k_e) &= \bar{\xi}_{0y}(k)-\bar{\mathbf{z}}^{T}(k)\widehat{\theta}(k_e-1)\\ L(k_e) &= \frac{P(k_e-1)\bar{\mathbf{z}}(k)}{\rho(k_e)+\bar{\mathbf{z}}^{T}(k)P(k_e-1)\bar{\mathbf{z}}(k)}\\ P(k_e) &= \frac{1}{\rho(k_e)}\left[P(k_e-1)-\frac{P(k_e-1)\bar{\mathbf{z}}(k)\bar{\mathbf{z}}^{T}(k)P(k_e-1)}{\rho(k_e)+\bar{\mathbf{z}}^{T}(k)P(k_e-1)\bar{\mathbf{z}}(k)}\right]\end{aligned} \tag{8.57}$$

where $\rho(k_e)$ is the forgetting factor and in this chapter it is chosen to be

$$\rho(k_e) = (1-0.01)\rho(k_e-1)+0.01, \quad \rho(k_s) = 0.98 \tag{8.58}$$

Step 3

Solve the following equation with an appropriately chosen $T(p)$ and estimated $\widehat{A}(p)$ to have $\widehat{R}(p)$ and $\widehat{S}(p)$ at every estimation time instant:

$$T(p)A^m(p) = \widehat{A}(p)\widehat{S}(p) + \widehat{R}(p) \tag{8.59}$$

Step 4

Generate the piecewise constant control signal $\bar{u}(k)$:

$$\begin{aligned}\bar{u}(k) &= \frac{T\left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)B^m\left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)\bar{u}_m(k) - \widehat{R}\left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)\bar{y}(k)}{\widehat{B}\left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)\widehat{S}\left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)} \\ &= \frac{T'(z^{-1})B'^m(z^{-1})\bar{u}_m(k) - \widehat{R}'(z^{-1})\bar{y}(k)}{\widehat{B}'(z^{-1})\widehat{S}'(z^{-1})}\end{aligned} \tag{8.60}$$

where

$$\begin{aligned}T'(z^{-1}) &= (1-z^{-1})^{n_T} + \sum_{i=1}^{n_T} t_i\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n_T-i} \\ B'^m(z^{-1}) &= \sum_{i=0}^{m} b_i^m\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{m-i} \\ R'(z^{-1}) &= \sum_{i=1}^{n} r_i\left(\frac{T}{2}\right)^{n_T+m-(n-i)}(1+z^{-1})^{n_T+m-(n-i)}(1-z^{-1})^{n-i} \\ B'(z^{-1}) &= \sum_{i=0}^{m} b_i\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{m-i} \\ S'(z^{-1}) &= (1-z^{-1})^{n_T} + \sum_{i=1}^{n_T} s_i\left(\frac{T}{2}\right)^i(1+z^{-1})^i(1-z^{-1})^{n_T-i}\end{aligned} \tag{8.61}$$

Step 5

Go to *Step 1.*

The digital implementation techniques of the MRACS by the delta operator (Goodwin, Lozano-Leal, Mayne and Middelton 1986) or the trapezoidal rule (Sagara and Yang 1988) can be described similarly as the above procedure.

In practice, it is critical to choose the sampling interval T and the estimation interval T_e considering the trade-off between the truncation error and the computational burden. Here we outline the methods of selection of T and T_e, following the error analysis of the BPFs studied by Rao and Srinivasan (1978). Basically, the most important element of the adaptive loop is the estimator. Therefore it is important to choose a small sampling interval T so that $\bar{\xi}_{iy}(k)$ and $\bar{\xi}_{iu}(k)$ approximate $\xi_{iy}(t)$ and $\xi_{iu}(t)$ accurately. If a function $f(t)$ is represented by a series of the BPFs $\phi_k(t)$ in the subinterval $[(k-1)T, kT)$, the representation error is

$$e_k(t) = \bar{f}(k)\phi_k(t) - f(t), \quad (k-1)T \le t < kT \tag{8.62}$$

where $\bar{f}(k)$ is given by

$$\bar{f}(k) = \frac{1}{T}\int_{(k-1)T}^{kT} f(t)\,dt \tag{8.63}$$

to minimize the squared error $e_k^2(t)$. It can be shown (Rao and Srinivasan 1978) that the minimum integral squared error ϵ_k^2 is

$$\epsilon_k^2 = \frac{T^3}{12}[f'(t_k)]^2, \quad (k-1)T \le t_k < kT \tag{8.64}$$

Hence the mean squared error over the interval $[0, NT)$ is

$$\begin{aligned} E^2 &= \frac{T^3}{12}\sum_{k=1}^{N}[f'(t_k)]^2, \quad (k-1)T \le t_k < kT \\ &\le \frac{T^3}{12} N\, f'^2_{max} \end{aligned} \tag{8.65}$$

where f'_{max} is the largest among all the $f'(t_k)$.

Taking the mean squared error over the normal interval, we have

$$E_n^2 \le \frac{T^2}{12} f'^2_{max} \tag{8.66}$$

If a signal $f(t) = A\sin(\omega t)$ has to be approximated by a set of the BPFs over the normal interval by subinterval T, and if the allowable error should be such that

$$\frac{E_n}{A} \le \sigma_r \tag{8.67}$$

where σ_r is a constant specified error bound, then from equation (8.66), we have

$$T \le \frac{2\sqrt{3}\sigma_r}{\omega_{max}} \tag{8.68}$$

For the low-pass filters in equation (8.50), usually τ is selected such that the pass-band of the filters includes the frequency band of importance to the analysis. Suppose that the highest frequency of the main components of the outputs of the filters is $1/\tau$ radians per second, then we have

$$T \le 2\sqrt{3}\sigma_r\,\tau \tag{8.69}$$

In our case, we choose $\sigma_r = 5\%$. Then if the sampling interval is chosen such that

$$T \leq 0.173\,\tau \tag{8.70}$$

$\bar{\xi}_{iy}(k)$ and $\bar{\xi}_{iu}(k)$ may approximate $\xi_{iy}(t)$ and $\xi_{iu}(t)$ accurately.

Now we will consider how to choose the estimation interval T_e. Actually, the parameters are estimated at every estimation time instant $t = k_e T_e$. That is, we sample the values of the digital filters' outputs $\bar{\xi}_{iy}(k), \bar{\xi}_{iu}(k)$ with the estimation interval T_e to estimate the parameters by the LS algorithm. Therefore to keep the information of $\bar{\xi}_{iy}(k)$ and $\bar{\xi}_{iu}(k)$, T_e should be chosen to agree with the sampling theorem. It has been mentioned that the highest frequency of the main components of the outputs of the filters is $1/\tau$ radians per second, thus T_e should be chosen such that

$$T_e \leq \pi\,\tau \tag{8.71}$$

8.6 Numerical examples

The effectiveness of the discussed methods is demonstrated by applying them to an unknown unstable minimum phase system governed by the following differential equation:

$$\begin{aligned} &(p^2 + a_1 p + a_2)y(t) = b_0 u(t) \\ &a_1 = 1.0,\ a_2 = -3.0,\ b_0 = -1.0 \end{aligned} \tag{8.72}$$

with unknown nonzero initial conditions:

$$y(0) = 1.0, \quad y'(0) = 1.0, \quad u(0) = 0.0 \tag{8.73}$$

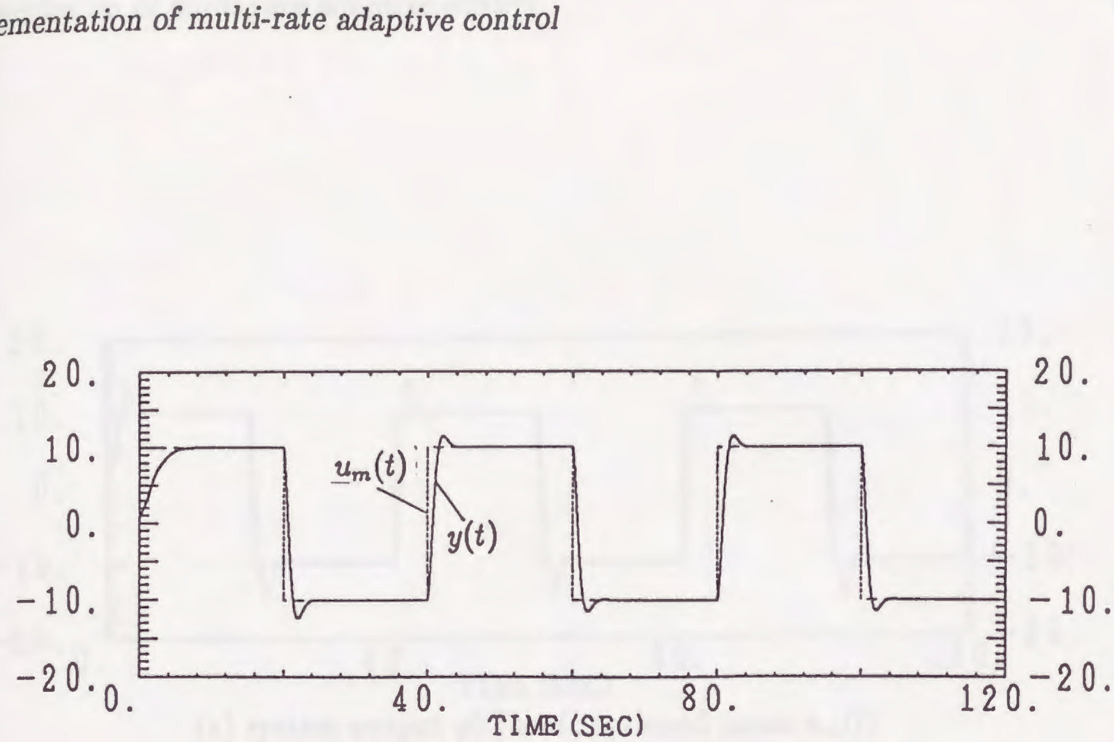
The reference model is

$$\begin{aligned} &(p^2 + a_1^m p + a_2^m)y_m(t) = b_0^m u_m(t) \\ &a_1^m = 2.8,\ a_2^m = 4.0,\ b_1^m = 4.0 \end{aligned} \tag{8.74}$$

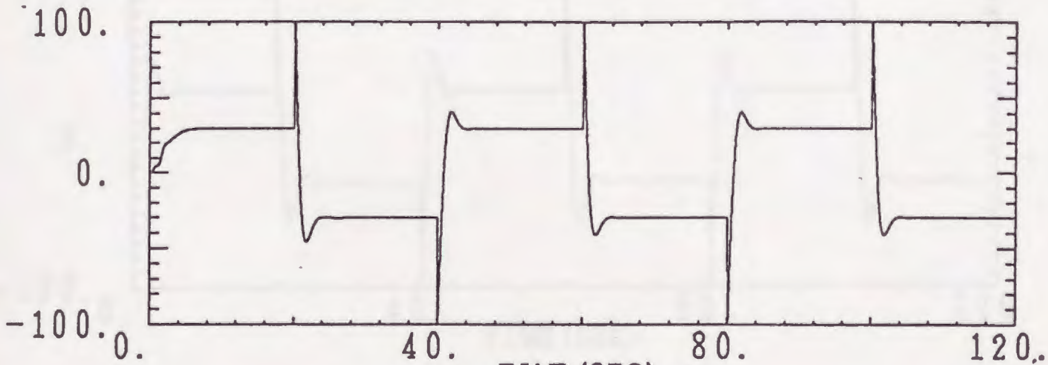
And the command input $u_m(t)$ is a square wave with a time period of 40 seconds. The amplitude is ± 10.

The low-pass filter is chosen to be

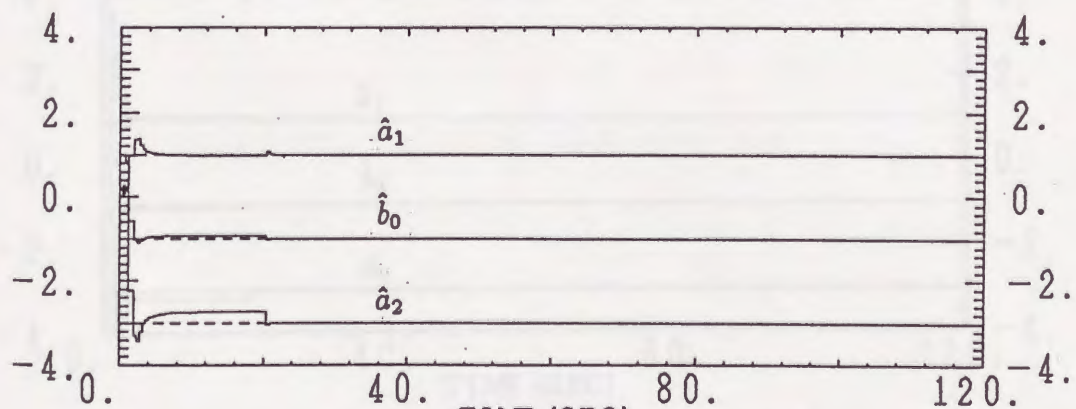
$$F(p) = \frac{1}{(\tau p + 1)^2}, \qquad \tau = 0.2 \tag{8.75}$$



(a) system output $y(t)$ and command input $u_m(t)$

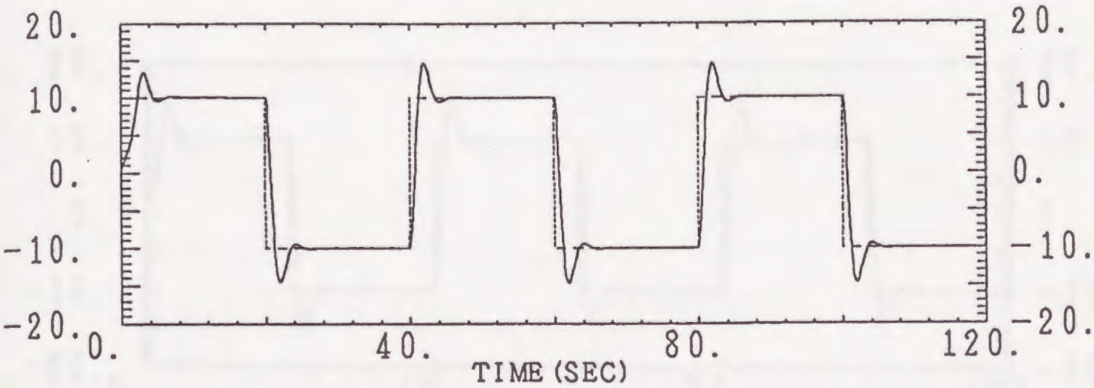


(b) system input $u(t)$

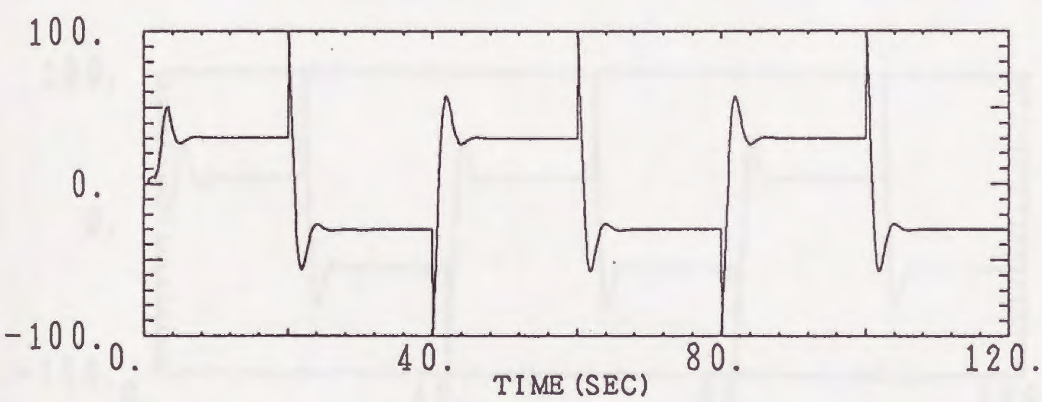


(c) parameter estimates($\hat{a}_1 = 0.9997, \hat{a}_2 = -2.9992, \hat{b}_0 = -1.0000$)

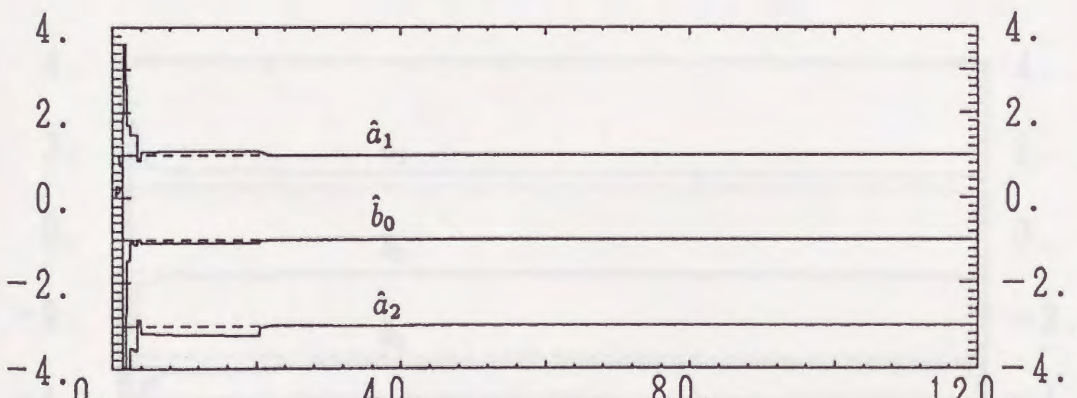
Figure 8.2: Results of BPF method ($T = 0.02,\ T_e = 0.4$).



(a) system output $y(t)$ and command input $u_m(t)$

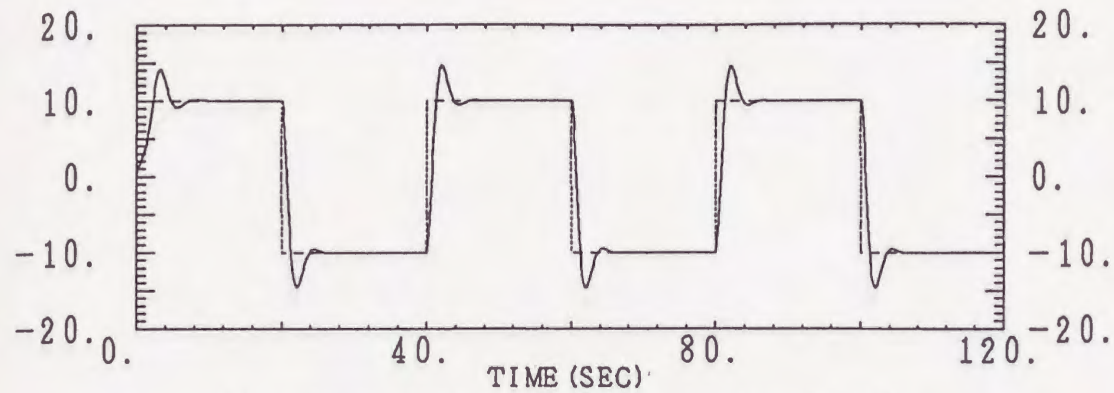


(b) system input $u(t)$

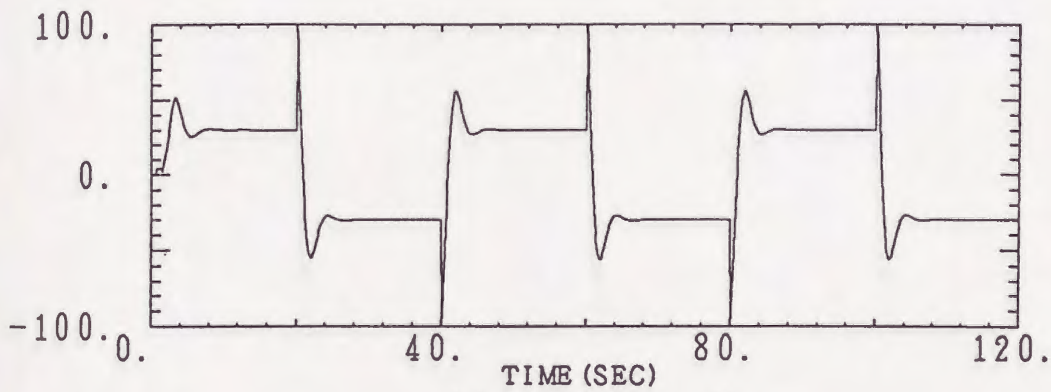


(c) parameter estimates($\hat{a}_1 = 0.9974, \hat{a}_2 = -2.9858, \hat{b}_0 = -0.9953$)

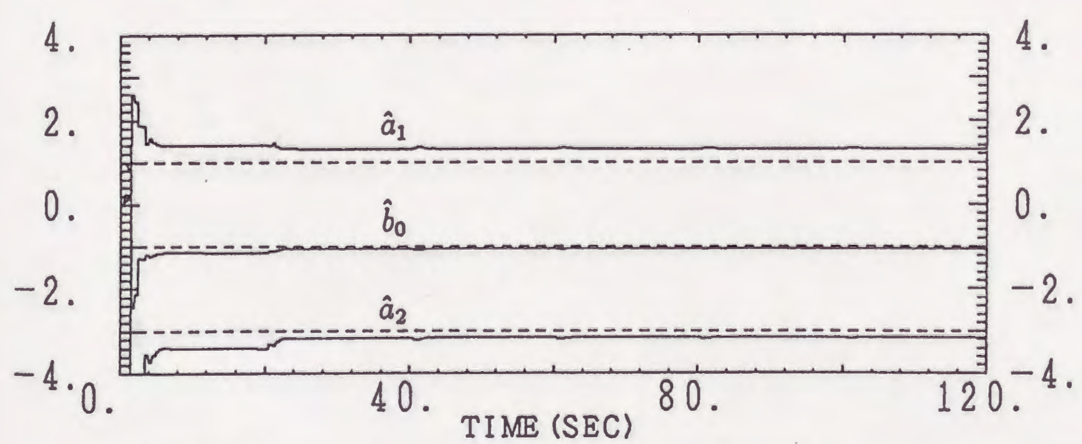
Figure 8.3: Results of BPF method ($T = 0.1,\ T_e = 0.5$).



(a) system output $y(t)$ and command input $u_m(t)$

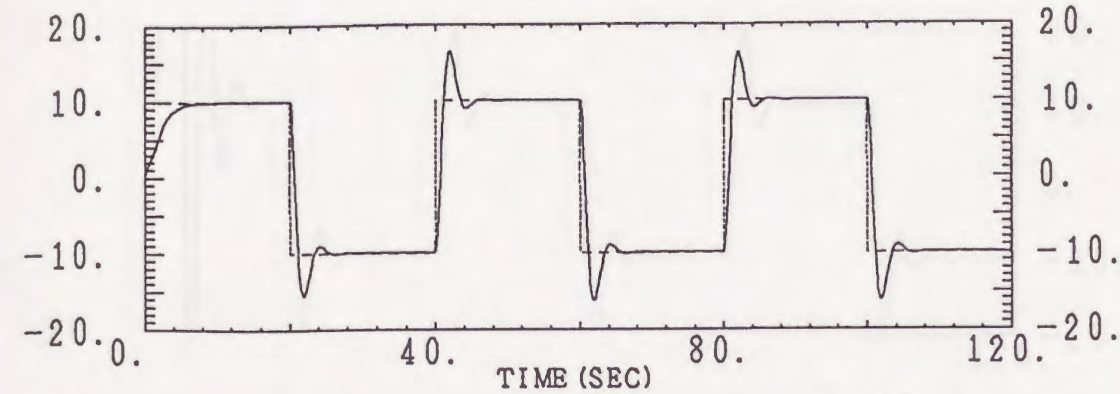


(b) system input $u(t)$

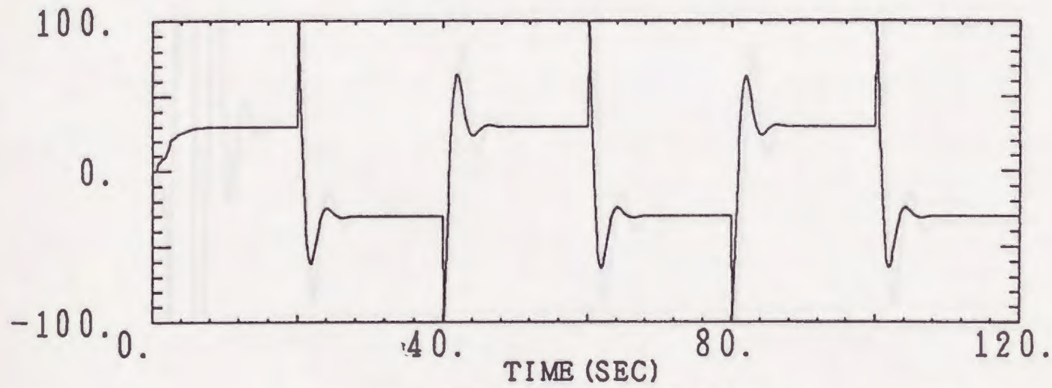


(c) parameter estimates($\hat{a}_1 = 1.3146, \hat{a}_2 = -3.1768, \hat{b}_0 = -1.0577$)

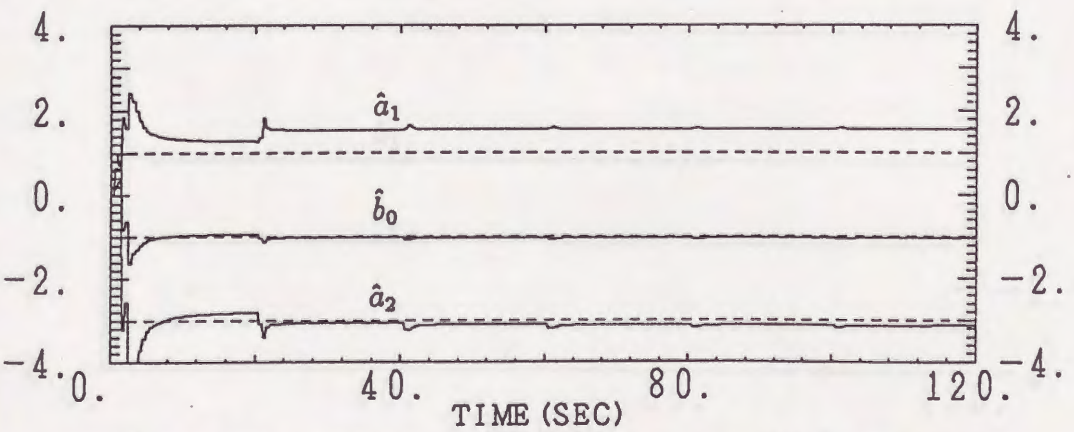
Figure 8.4: Results of trapezoidal method ($T = 0.1$, $T_e = 0.5$).



(a) system output $y(t)$ and command input $u_m(t)$

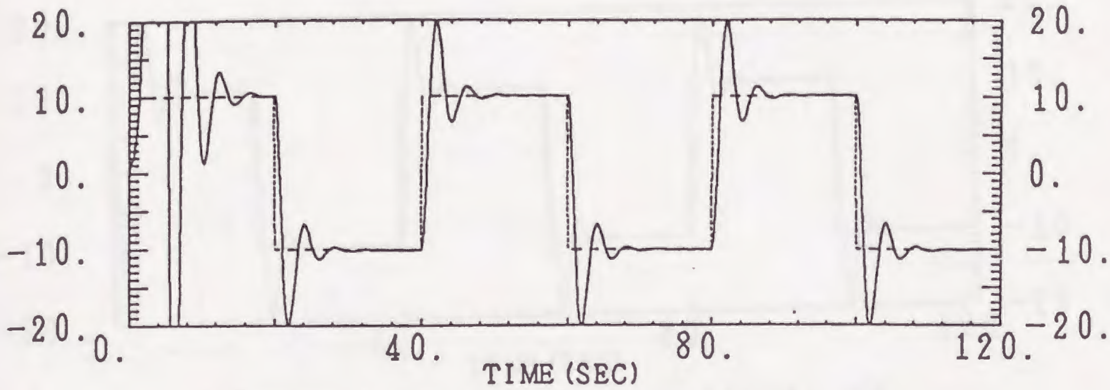


(b) system input $u(t)$

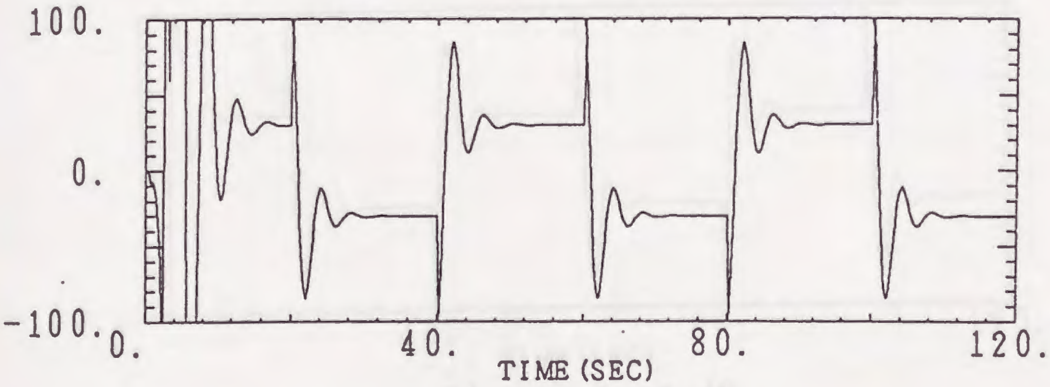


(c) parameter estimates($\hat{a}_1 = 1.5628, \hat{a}_2 = -3.1713, \hat{b}_0 = 1.0542$)

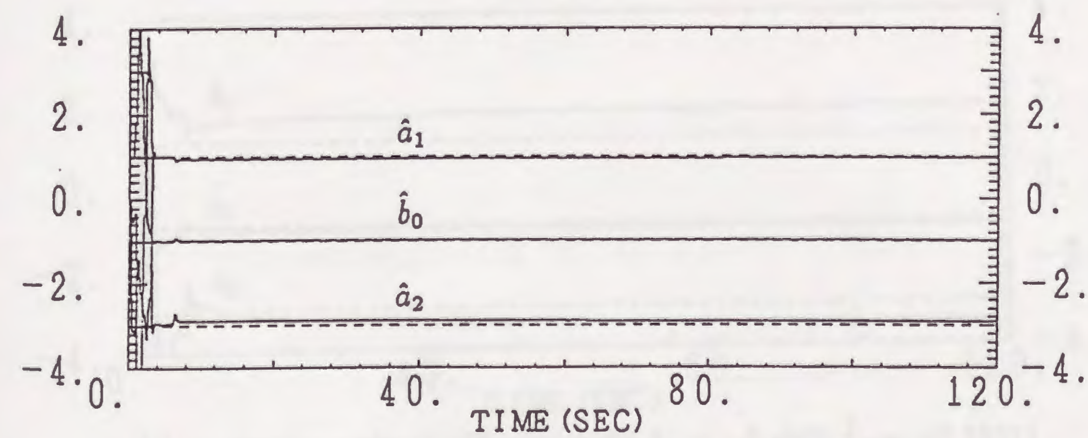
Figure 8.5: Results of delta operator method ($T = 0.1,\ T_e = 0.5$).



(a) system output $y(t)$ and command input $u_m(t)$

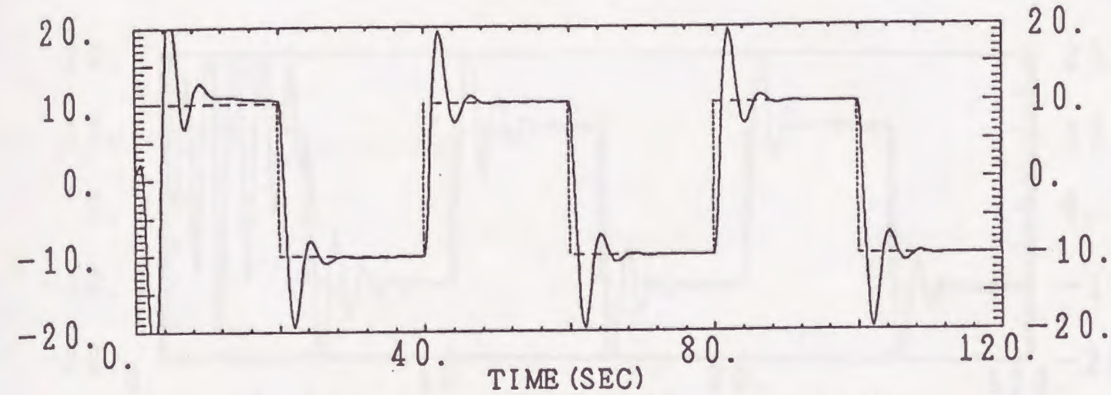


(b) system input $u(t)$



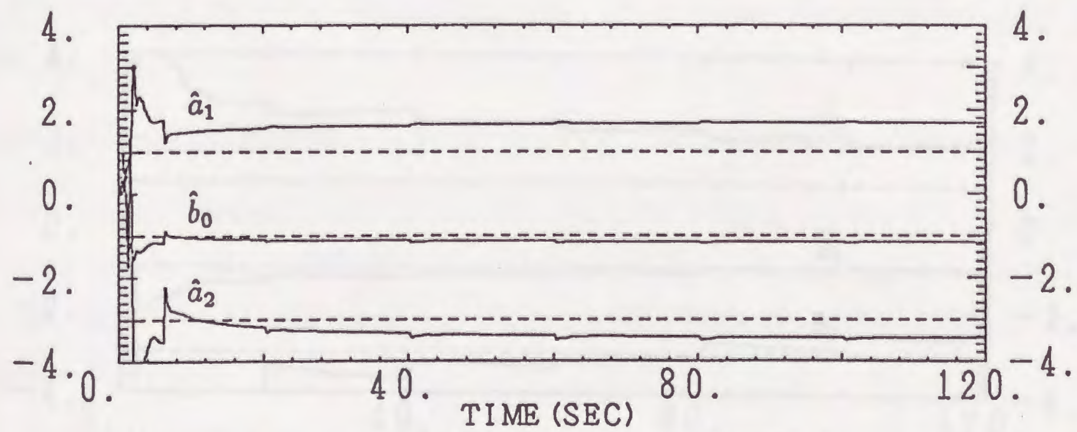
(c) parameter estimates($\hat{a}_1 = 0.9536, \hat{a}_2 = -2.9117, \hat{b}_0 = -0.9732$)

Figure 8.6: Results of BPF method ($T = 0.2,\ T_e = 0.4$).



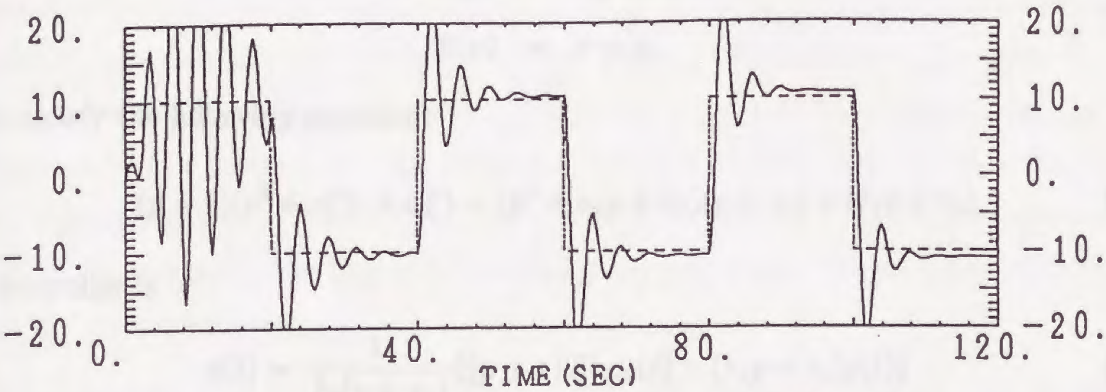
(a) system output $y(t)$ and command input $u_m(t)$

(b) system input $u(t)$

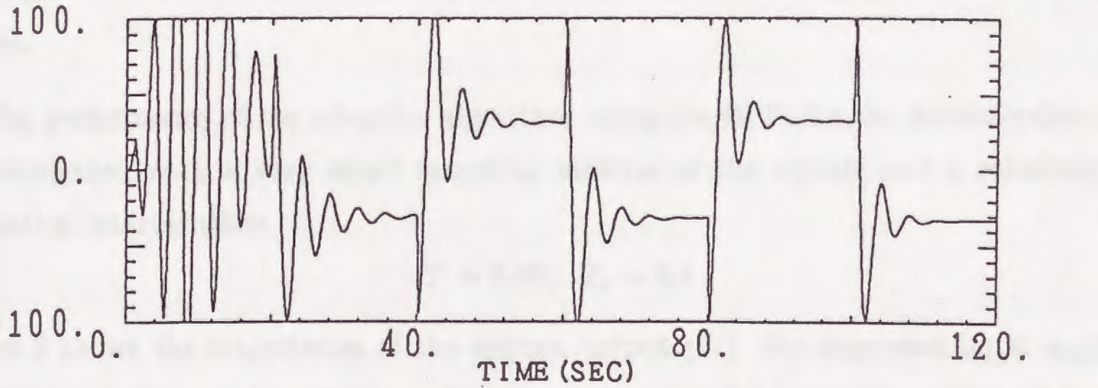


(c) parameter estimates($\hat{a}_1 = 1.6970, \hat{a}_2 = -3.4605, \hat{b}_0 = -1.1537$)

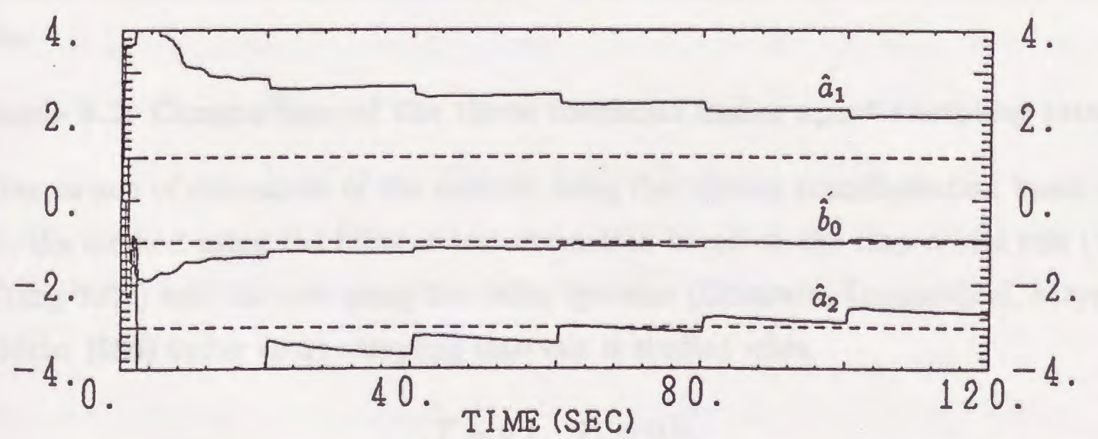
Figure 8.7: Results of trapezoidal method ($T = 0.2$, $T_e = 0.4$).



(a) system output $y(t)$ and command input $u_m(t)$



(b) system input $u(t)$



(c) parameter estimates

Figure 8.8: Results of delta operator method ($T = 0.2$, $T_e = 0.4$).

The polynomial $T(p)$ is chosen to be

$$T(p) = p+1 \tag{8.76}$$

and $R(p)$, $S(p)$ are described as

$$\begin{aligned} R(p) &= r_1 p + r_2 \\ S(p) &= p + s_1 \end{aligned} \tag{8.77}$$

which satisfy the following equation:

$$(p+1)(p^2 + a_1^m p + a_2^m) = (p^2 + a_1 p + a_2)(p + s_1) + (r_1 p + r_2) \tag{8.78}$$

The controller is

$$u(t) = \frac{1}{b_0(p+s_1)}\{(p+1)b_0^m u_m(t) - (r_1 p + r_2)y(t)\} \tag{8.79}$$

Example 8.1: Performance of the BPF method with a very small sampling interval.

The performance of the adaptive algorithm using the BPFs for the second-order system is investigated with a very small sampling interval of the signals and a relatively large estimation interval when

$$T = 0.02, \quad T_e = 0.4 \tag{8.80}$$

Figure 2 shows the trajectories of the system output $y(t)$, the command input $u_m(t)$, the control input $u(t)$ and the system parameter estimates with their true values. It is shown that when a small sampling interval is taken, the adaptive algorithm has very excellent results.

Example 8.2: Comparison of the three methods under apart sampling intervals.

Comparison of robustness of the method using the bilinear transformation based on the BPFs, the method using the bilinear transformation based on the trapezoidal rule (Sagara and Yang 1988) and the one using the delta operator (Goodwin, Lozano-Leal, Mayne and Middelton 1986) under apart sampling intervals is studied when

$$T = 0.1, \quad T_e = 0.5 \tag{8.81}$$

and

$$T = 0.2, \quad T_e = 0.4 \tag{8.82}$$

The results for $T = 0.1, T_e = 0.5$ are shown in Figures 3~5. For the BPF method, the parameter estimates are still accurate. The overshoot of the output $y(t)$ in Figure 3 becomes large compared with the result in Figure 2, however the control performance is still excellent at low frequencies (constant $u_m(t)$). When an apart sampling interval is taken, the truncation error appears significantly at high frequencies (changing $u_m(t)$) but can be neglected at low frequencies (constant $u_m(t)$). In practice, the controlled plants are usually excited by band-limitted low frequency signals, therefore the adaptive algorithm need not require too small a sampling interval considering the limitation of computational time. For the trapezoidal rule method and the delta operator method, although the controlled system output signals are similar to the result by the BPF method (in fact, the overshoot of the delta operator method is the largest of the three), the parameter estimates are greatly biased due to the truncation error.

The results for $T = 0.2, T_e = 0.4$ are shown in Figures 6~8. Clearly, in this case, the truncation error is greater than that in the case of $T = 0.1$. The system outputs for both the trapezoidal rule and the BPF methods are still acceptable at low frequencies while very large overshoot arises in both methods. It is shown that the parameter estimates by the BPF method are still acceptable, while the parameter estimates by the trapezoidal rule method are much greatly biased. For the delta operator method, unfortunately, the parameter estimates do not converge and the system output performance is very poor.

Due to the limitation of simulation, although it is difficult to draw the general conclusions about the three methods for all cases, we can summarize the simulation results as follows:

1 For very small sampling intervals, all the three methods give excellent performances of the adaptive parameter estimation and control algorithm. It seems that selection of the estimation interval T_E is not so serious as that of the sampling interval T.

2 The BPF method is the most satisfactory for the purposes of both parameter estimation and system output control, since the step functions produced by sampling and holding employed in digital control for continuous systems can be precisely expressed by the BPF coefficients.

3 The trapezoidal rule method has similar control performance to the BPF method. However, the parameter estimation error by this method is much greater than that by the BPF method when an apart sampling interval is taken. This is because that by the trapezoidal rule, we can only obtain sampled data of continuous signals at sam-

pling time instants, thus the ZOH input signal of a digital control system cannot be expressed so precisely as the BPF method.

4 The delta operator method which has been studied by some authors (Goodwin, Lozano-Leal, Mayne and Middelton 1986, Janecki 1988) is the best known of the three. However, as mentioned previously, since the delta operator cannot give so accurate approximations as the bilinear transformation and hence produces more truncation error, the robustness under apart sampling intervals is the poorest of the three.

8.7 Conclusion

In this chapter, the techniques of the implementation of multi-rate adaptive control of continuous systems have been discussed.

Approximated models of a continuous system by the bilinear transformation and the delta operator are discussed and it is pointed out that the bilinear transformation based on the BPFs is the most satisfactory one in digital control systems, where the step functions produced by sampling and holding can be precisely expressed by the BPFs. The relation between the BPF model and the common ZOH sampled model is discussed and it is shown that the BPF model can be thought as the Páde approximation of the ZOH sampling of the controlled system model, whereas the approximated discrete-time model obtained by the bilinear transformation based on the trapezoidal rule neglects the fact that the input control signal $u(t)$ is constant between the sampling instants, if a ZOH is used.

Simulation experiments have been carried out with an unstable continuous system. It is shown that our algorithm is useful not only for adaptive control of continuous systems but also for parameter estimation of unstable continuous systems. Comparison of the BPF method with the trapezoidal rule method and the famous delta operator method has been taken through numerical examples. It is concluded that the BPF method is the most satisfactory of the three while the robustness to the trucation error of the delta operator method is the poorest.

Chapter 9

Conclusions

In this dissertation, the methods of identification and adaptive control for continuous-time systems using digital signal processing techniques have been studied in a unified sense. The major results and concluding remarks are summarized as follows.

1: In chapter **2**, the popular integral-equation approach to identification of continuous-time systems is reviewed in a unified sense. A general form of the IIF employed in the conventional integral-equation method concerning the OFs and the numerical integrating rules is formulated and it is emphasized that the troubling initial condition problem arises due to the cancellation of the differential operators. Motivated by this fact, a new calculation procedure of the multiple integrations of the signal derivatives termed NIIF is proposed for which the initial conditions need not be identified as unknown parameters. Therefore complexity of the identification algorithms can be greatly reduced compared with the conventional methods. Effects of the measurement noise in integral-equation approach is also investigated. It is found that the noise reducing effects of the IIF and the proposed NIIF are similar since the frequency responses are same. Since the IIF and the NIIF can be viewed as a kind of unstable IIR filters which have multiple poles on the unit circle, the equation error in the integral-equation due to the noise increases with the time and this can make the LS estimates diverge.

2: In chapter **3**, a unified approach to direct recursive identification techniques of continuous systems from sampled input-output data using digital low-pass filters is discussed. Using a pre-designed digital low-pass filter, a discrete-time estimation model in continouou-time system parameters is constructed easily. Thus the system parameters can be identified directly by recursive identification algorithms. It is concluded through numerical results that if the filter is designed so that its pass-band matches that of the system under study closely and thus the noise effects are sufficiently reduced, accurate estimates can be obtained by recursive identification algorithms such as the LS method and the IV method. It is also pointed out that some well-known distinct methods are unified as either the IIR or the FIR filtering approach. Some new comments to the initial condition problem which is unclear in the literature are given. It is found that the NIIF method can be viewed as a special case of the SVF method. Therefore, when designing the IIR filters in continuous-time domain, it is not necessary to let the pre-filter has fast damped characteristic so that the initial conditions decay as soon as possible as suggested in some other previous works. It is pointed that when using the IIR filtering approach, one may be troubled by the initial conditions if he neglects the fact that usually the commulative law does not hold in the linear pre-processing procedure.

3: In chapter 4 some recursive identification algorithms for continuous systems using an adaptive procedure are discussed. Using the estimated denominator of the transfer function of the discrete-time model obtained by the bilinear transformation to construct the adaptive IIR filters which are introduced to avoid direct approximations of differentiations from sampled data, an approximated discrete-time estimation model with continuous system parameters is derived. With filtered inputs and delayed filtered outputs as instrumental variables, some kinds of recursive IV identification algorithms are proposed to obtain consistent estimates in the presence of noise. The proposed identification algorithms have close relations to the standard recursive identification algorithms for common discrete-time systems. The results of this chapter show that the continuous-time system parameters can be identified in very similar ways to those for common discrete-time systems. The continuous-time system identification requires a digital filtering procedure to avoid direct approximation of differentiations from sampled data while the discrete-time system identification is usually based directly on a linear regression model composed of delayed sampled input-output data.

4: In chapter 5, the problem of identification of continuous systems is considered when both the discrete input and output measurements are contaminated by white noises. It is found that in the presence of input measurement noise, the pass-band of the digital low-pass filters should be chosen such that it includes the main frequencies of both the system input and output signals in some range. When the noise effects cannot be neglected, the BCLS method combined with a digital filter is applied to obtain the consistent estimate, which compensates the bias of the LS estimate with the estimates of the noise variances . And it is pointed out that if the input signals are persistently exciting, the BCLS algorithm is stable. Both classes of filters (FIR filter and IIR filter) are employed. The FIR filters can be applied to the BCLS method directly, whereas the IIR filters require some approximations. And numerical examples show that the BCLS method combined with a digital pre-filter yields very accurate parameter estimates when both the discrete input and output measurements are contaminated by white noises.

5: In chapter **6**, the BCIV method for identification of continuous systems is proposed, in the case where the discrete input measurement is corrupted by a white noise and the discrete output measurement is corrupted by a noise which may be coloured. The continuous system is identified through the discrete-time estimation model derived in chapter **4** using the adaptive procedure. The effects of the output noise is avoided by the IV method with filtered inputs and delayed filtered outputs as instrumental variables and the bias of the IV estimate due to the input noise is eliminated by the proposed BCIV method. Although general conclusions on the stability of the method have not been given, Empirical numerical studies tell that during the identification process, if we monitor the stability of the algorithm and contract the estimates within the stable region, the method gives excellent estimates in typical cases.

6: In chapter **7**, a new approach using two-dimensional filtering techniques to recursive parameter identification of second-order distributed parameter systems in the presence of measurement noise under unknown initial condition and boundary condition is proposed. The LS method is still efficient in the presence of low measurement noise if the filter parameters are designed so that the noise effects are reduced sufficiently. Using filtered input data as instrumental variables, an IV method is also presented to obtain consistent estimates when the digital low-pass filters are not designed appropriately or the output data is corrupted by high measurement noise.

7: In chapter **8**, comparison of the discretization methods for the adaptive system using the BPFs, the trapezoidal integrating rule and the well-known delta operator is first discussed through theoretical analysis and simulation study. It is shown that the bilinear transformation gives more accurate approximations than the delta operator.

And it is shown that the approximated discrete-time model obtained by the bilinear transformation based on the BPFs can be viewed as the Páde approximation of the ZOH sampling of the controlled system model, whereas the approximated discrete-time model obtained by the bilinear transformation based on the trapezoidal rule neglects the fact that the input control signal $u(t)$ is constant between the sampling instants, when a ZOH is used. Therefore the BPF method is the most effective one of the three for digital systems. Then the implementation techniques of multi-rate indirect model reference adaptive control for continuous systems purely using digital computers are described. To reduce the computational burden, the algorithm is implemented in a multi-rate manner with a small sampling interval of the system signals and a relatively large parameter estimation interval. Comparison of the discretization methods for the adaptive system using the BPFs, the trapezoidal integrating rule and the delta operator is discussed through simulation study. It is shown that the BPF method is the most effective one since the BPFs give excellent approximations of the signals of the digital controlled continuous system.

References

ÅSTRÖM, K. J., 1969, On the choice of sampling rates in identification of time series. *Information Sciences*, **1**, 273-278.

ÅSTRÖM, K. J., and EYKHOFF, P., 1971, System identification–a survey. *Automatica*, **7**, 123-167.

ÅSTRÖM, K. J., HAGANDER, P., and STERNBY, J., 1984, Zeros of sampled systems. *Automatica*, **20**, 31-38.

ÅSTRÖM, K. J., and WITTENMARK, B., 1984, *Computer Controlled Systems–Theory and Design* (Prentice-Hall, Inc., Englewood Cliffs).

ANDERSON, B. D. O., 1985, Identification of scalar errors-in-variables models with dynamics. *Automatica*, **21**, 709-716.

ANDERSON, B. D. O., and DEISTLER, M., 1984, Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, **15**, 1-13.

ANDERSON, B. D. O., and DEISTLER, M., 1987, Dynamic errors-in-variables systems with three variables. *Automatica*, **23**, 611-616.

CHAO, Y. C., CHEN, C. L., and HUANG, H. P., 1987, Recursive parameter estimation of transfer function matrix models via Simpson's Integrating rules. *International Journal of Systems Science*, **18**, 901-911.

CHEN, W. L., and LEE, C. L., 1982, On the convergence of the block-pulse series solution of a linear time-invariant system. *International Journal of Systems Science*, **13**, 491-498.

CHENG,B., and HSU,N.S., 1982, Single-input single-output system identification via block-pulse functions. *International Journal of Systems Science*, **13**, 697-702.

CRITTENDEN, J. L., MULHOLLAND, R. J., HILL, J., and MARTINEZ, E. F., 1983, Sampling and data analysis properties of prony's method applied to model identification. *International Journal of Systems Science*, **14**, 571-584.

DEISTLER, M., and ANDERSON, B. D. O., 1989, Linear dynamic errors-in-variable models. *Journal of Econometrics*, **41**, 39-63.

EGARDT, B., 1979, Unificationn of some continuous-time adaptive control schemes. *I.E.E.E. Transactions on Automatic Control*, **24**, 588-592.

EITELBERG, E., 1988, Continuous-time system representation with exact macro difference expressions. *International Journal of Control*, **47**, 1207-1212.

EYKHOFF, P., 1974, *System Identification: Parameter and State Estimation* (Wiley, London).

FENG, C. B., and ZHENG, W. X., 1988, On-line modified least-squares parameter estimation of linear systems with input-output data polluted by measurement noises. *Preprints of the 8th IFAC/IFORS Sympo. on Identification and System Parameter Estimation*, 1189-1194(Beijing).

FERNANDO, K. V., and NICHOLSON, H. 1985, Identification of linear systems with input and output noise: the Koopmans-Levin method. *Proceedings of the Institution of Electrical Engineers*, **132**, Pt.D., 30-36.

FRIELANDER, B., 1982, A modified prefilter for some recursive parameter estimation algorithms. *I.E.E.E. Transactions on Automatic Control*, **27**, 232-235.

GAWTHROP, P. J., 1980, Hybrid self-tuning control. *Proceedings of the Institution of Electrical Engineers*, **D-127**, 229-236.

GAWTHROP, P. J., 1982, A continuous-time approach to discrete-time self-tuning control. *Optimal Control: Applications & Methods.* **3**, 399-414.

GAWTHROP, P. J., 1987, *Continuous-Time Self-Tuning Control*, Volume 1-Design (Wiley, London).

GOODWIN, G. C., LOZANO-LEAL, R., MAYNE, D. Q., and MIDDLETON, R. H., 1986, Rapprochement between continuous and discrete model reference adaptive control. *Automatica*, **22**, 199-207.

HABERLAND, B. L., and RAO, S. S., 1973, Discrete-time models: bilinear transform and ramp approximation equivalence. *I.E.E.E. Transactions on Audio Electroacoustics*, **21**, 382-384.

HARA, K., YAMAMOTO, T., and OUGITA, K., 1988, Parameter identification of distributed parameter systems using spline functions. *International Journal of Systems Science*, **19**, 49-64.

HANSELMANN, H., 1987, Implementation of digital controllers–a survey. *Automatica*, **23**, 7-32.

HAYKIN, S. S., 1972, A unified treatment of recursive digital filtering. *I.E.E.E. Transactions on Automatic Control*, **17**, 113-116.

HORI, N., NIKIFORUK, N., and KANAI, K., 1988, On a discrete-time system expressed in the Euler operator. *Proc. of American Control Conference*, 873-877, Atlanta.

HORI, N., MORI, T.,and NIKIFORUK, P. N., 1989, A new perspective for discrete-time models of a continuous-time system. *Proc. of American Control Conference*, 1659-1662, Pittsburgh.

HOUPIS, C. H., 1985, Refined design method for sampled systems: the pseudo-continuous-time (PCT) control system design. *Proceedings of the Institution of Electrical Engineers*, **D-132**, 69-74.

HSU, N. S., and CHENG, B., 1981, Analysis and optimal control of time-varying systems via block-pulse functions. *International Journal of Control*, **33**, 1107-1122.

HSU, N. S., and CHENG, B., 1982, Identification of non-linear distribuited systems via block-pulse functions. *International Journal of Control*, **36**, 281-291.

HOMSSI, L., TITLI, A., and DESPUJOLS, A., 1991, Continuous-time process identification: Comparison of eight methods and prticular aspects. *Preprints of the 9th IFAC/IFORS Sympo. on Identification and System Parameter Estimation*, Budapest, 1634-1642.

HORNG, I. R., CHOU, J. H., and TSAI,C. H., 1986, Analysis and identification of linear distributed system via Chebyshev series. *International Journal of Systems Science*, **17**, 1089-1095.

HUANG, H. P., CHEN, C. L., and CHAO,Y.C., 1987, Augmented hybrid method for continuous process identification from sampled data with coloured noise. *International Journal of Control*, **46**, 1373-1390.

HUANG, H. P., CHEN, C. L., and CHAO, Y. C., 1988, Identification of MIMO continuous-time models by indirect method. *International Journal of Systems Science*, **19**, 1281-1297.

HWANG, C., and GUO, T. Y., 1984, Parameter identification orthogonal shifted Legendre polynomials. *Journal of Franklin Institute*, **318**, 59-69.

HWANG, C., and SHIH, Y. P., 1982, Parameter identification via Laguerre polynomials. *International Journal of Systems Science*, **13**, 209-211.

ISHIMARU, N., HANAZAKI, I., and AKIZUKI, K., 1988, System identification with the least square method using backward difference of sampled data. *Transations of I.E.E. of Japan*, **c-108**, 298-303 (in Japanese).

JAMES, P. N., SOUTER, P., and DIXON, D. C, 1972, Suboptimal estimation of the parameters of discrete systems in the presence of correlated noise. *Electronics Letters*, **8**, 411-412.

JANECKI, D., 1988, Model reference adaptive control using delta operator. *I.E.E.E. Transactions on Automatic Control*, **33**, 771-775.

JHA, A. N., and ZAMAN, S., 1985, Identification of linear distributed systems using Laguerre operational matrices. *International Journal of Systems Science*, **16**, 761-767.

JIANG, N., and JIANG, J., 1988, Parameter or non-parameter identification of distributed parameter systems via block-pulse functions. *International Journal of Systems Science*, **19**, 1039-1045.

JORDAN, J. R., JALALI-NAINI, S. A., and MACKIE, R. D. L., 1990, System identification with Hermite modulating functions. *Proceedings of the Institution of Electrical Engineers*, **D-137**, 87-93.

KALMAN, R. E., 1983, Identifiability and Modeling in econometrics; Developments in Statistics. *Developments in Statistics*, **4**, 97-136 (Academic Press, edited by P.R. Krishnaiah).

KAWAJI, S., 1983, Block-pulse series analysis of linear systems incorporating observers. *International Journal of Control*, **37**, 1113-1120.

KAYA, Y., YAMAMURA, S., 1962, A self adaptive system with a PID controller. *A.I.E.E. Trans. App. Ind.*, **80**, 378-384.

KRAUS, F. and SCHAUFELBERGER, W., 1990, Identification with block pulse functions, modulating functions and differential operators. *International Journal of Control*, **51**, 931-942.

KRISHNA,H., 1988, Computational aspects of the bilinear transformation based algorithm for s-plane to z-plane mapping. *I.E.E.E. Transactions on Automatic Control*, **33**, 1086-1087.

KUO, B. C., 1980, *Digital Control Systems*, (Holt, Reinhart and Winston).

LEE, T. T., and CHANG, Y. F., 1986, Analysis and Identification of linear distributed systems via double general orthogonal polynomials. *International Journal of Control*, **44**, 395-405.

LEVADI, V. S., 1964, Parameter estimation of linear systems in the presence of noise. *Proceeding of the 1964 International Conference on Microwaves, Circuit Theory, and Information Theory*, 85-86, Tokyo.

LIU, C. C., and SHIH, Y. P. , 1985, System analysis, parameter estimation and optimal regulator design of linear systems via Jaccobi series. *International Journal of Control*, **42**, 211-224.

LJUNG, L., SÖDERSTRÖM, T., and GUSTAVSSON, I., 1975, Counter examples to general convergence of a commonly used recursive identification method. *I.E.E.E. Transactions on Automatic Control*, **20**, 643-652.

LJUNG, L., 1977a, On positive real transfer functions and the convergence of some recursive schemes. *I.E.E.E. Transactions on Automatic Control*, **22**, 539-551.

LJUNG, L., 1977b, Analysis of recursive stochastic algorithms. *I.E.E.E. Transactions on Automatic Control*, **22**, 551-575.

LJUNG, L., and SÖDERSTRÖM, T., 1983, *Theory and practice of recursive identification* (Cambridge, Mass: MIT Press).

MATHEW, A. V., and FAIRMAN, F. W., 1974, Transfer function matrix identification. *I.E.E.E. Transactions on Circuits and Systems*, **21**, 584-588.

MIZUNO, N., and FUJII, S., 1991, Adaptive control for continuous-time systems by using approximate discrete-time model. *Transations of the Society of Instrument and Control Engineers*, **27**, 678-685 (in Japanese).

MOHAN, B. M., and DATTA, K. B., 1987, Identification of linear time-invariant distributed parameter systems via shifted Legedre polynomials. *Proc. IMACWS/IFAC Int. Sympo. on Modelling and Simulation of Distributed Parameter Systems*, Hiroshima, Japan, 35-41.

MOHAN, B. M., and DATTA, K. B., 1989, Identification via Forurier series for a class of lumped and distributed parameter systems. *I.E.E.E. Transactions on Circuits and Systems*, **36**, 1454-1457.

MOHAN, B. M., and DATTA, K. B., 1991, Linear time-invariant distributed parameter system identification via orthogonal functions. *Automatica*, **27**, 409-412.

MULHOLLAND, R. J., and WEIDNER, R. J., 1980, Sampling strategies for the identification of compartment models. *International Journal of Systems Science*, **11**, 377-387.

NG, T. S., and GOODWIN, G. C., 1976, On optimal choice of sampling strategies for linear system identification. *International Journal of Control*, **23**, 459-475.

PALANISAMY, K. R., and BHATTACHARYA, D. K., 1981, System identification via block-pulse func- tions. *International Journal of Systems Science*, **12**, 643-647.

PARASKEVOPOULOS, P. N., 1985, Legendre series approach to identification and analysis of linear systems. *I.E.E.E. Transactions on Automatic Control*, **30**, 585-589.

PATRA, A., and RAO, G. P., 1989, Continuous-time approach to self tuning control: algorithm, implementation and assessment. *Proceedings of the Institution of Electrical Engineers*, **D-136**, 334-340.

PEARSON, A. E., and LEE, F. C., 1985a, On the identification of polynomial input-output differential systems. *I.E.E.E. Transactions on Automatic Control*, **30**, 778-782.

PEARSON, A. E., and LEE, F. C., 1985b, Parameter identification of linear differential systems via Forurier based modulating functions. *Control-Theory and Advanced Control*, **1**, 239-266.

RANGANATHAN, V., JHA, A. N., and RAJAMANI, V. S., 1984, Identification of linear distributed systems via Laguerre polynomials *International Journal of Systems Science*, **15**, 1101-1106.

RANGANATHAN, V., JHA, A. N., and RAJAMANI, V. S., 1986, Identification of non-linear distributed systems via a Laguerre polynomial approach. *International Journal of Systems Science*, **17**, 241-249.

RANGANATHAN, V., JHA, A. N., and RAJAMANI, V. S., 1987, Identification of distributed linear time varying systems via Laguerre polynomials. *International Journal of Systems Science*, **18**, 681-685.

RAO, G. P., and PALANISAMY, K. R., 1981, Improved algorithm for parameter identification in continuous systems via Walsh functions. *Proceedings of the Institution of Electrical Engineers*, **D-130**, 9-16.

RAO, G. P., and SRINIVASAN, T., 1978, Analysis and synthesis of dynamic systems containing time-delays. *Proceedings of the Institution of Electrical Engineers*, **D-125**, 1064-1068.

RAO,G.P., and SRINIVASAN,T., 1980, Multi-dimensional block-pulse functions and their use in the study of distributed parameter systems. *International Journal of Systems Science*, **11**, 689-708.

ROWE, I. H., 1970, A bootstrap method for the statistical estimation of model parameters. *International Journal of Control*, **12**, 721-738.

SAGARA, S., EGUCHI, M., and WADA, K., 1985a, A choice of the lower bound of sampling period for continuous system identification using discrete model. *Transations of the Society of Instrument and Control Engineers*, **21**, 13-20 (in Japanese).

SAGARA, S., EGUCHI, M., and WADA, K., 1985b, Optimal sampling period choice for system identification based on covariance matrix of pole estimation. *Transations of the Society of Instrument and Control Engineers*, **21**, 1150-1155 (in Japanese).

SAGARA, S., and WADA, K., 1977, On-line modified least-squares parameter estimation of linear discrete dynamic systems. *International Journal of Control*, **25**, 329-345.

SAGARA, S., and YANG, Z. J., 1988, Implementation of model reference adaptive control for continuous systems via bilinear transformation. *Proceedings of the 27th SICE Annual Conference Vol.1*, Narashino, Japan (in Japanese), 469-470.

SAGARA, S., YANG, Z. J., and WADA, K., 1990, Consistent parameter estimation for continuous systems via IIR digital filters. *Transations of the Society of Instrument and Control Engineers*, **26**, 39-45 (in Japanese).

SAGARA, S., YANG, Z. J., and WADA, K., 1991a, Digital filtering approach to identification of continuous systems. *Systems and Control–Topics in theory and applications* (Mita Press, Tokyo, Japan, edited by T. Ono and F. Kozin), 319-340.

SAGARA, S., YANG, Z. J., and WADA, K., 1991b, Identification of continuous systems using digital low-pass filters. *International Journal of Systems Science*, **22**, 1159-1176.

SAGARA, S., YANG, Z. J., and WADA, K., 1991c, Recursive identification algorithms for continuous systems using an adaptive procedure. *International Journal of Control*, **53**, 391-409.

SAGARA, S., YANG, Z. J., and WADA, K., 1991d, Identification of continuous systems from noisy sampled input-output data. *Preprints of the 9th IFAC/IFORS Sympo. on Identification and System Parameter Estimation*, Budapest, 1622-1627.

SAGARA, S., YANG, Z. J., and WADA, K., 1991e, Parameter identification of distributed parameter systems in the presence of measurement noise. *International Journal of Systems Science*, **22**, 1391-1401.

SAGARA, S., YANG, Z. J., and WADA, K., 1991f, Implementation of multi-rate model reference adaptive control for continuous systems via bilinear transformation based on block-pulse functions. *International Journal of Systems Science*, **22**, 1757-1774.

SAGARA, S., YUAN, K., and WADA, K., 1988a, Order determination and parameter estimation of continuous-time systems. *Preprints of the 8th IFAC/IFORS Sympo. on Identification and System Parameter Estimation*, Beijing, 739-744.

SAGARA, S., YUAN, K., and WADA, K., 1988b, Application of block-pulse functions to model reference adaptive control systems. *Transations of the Society of Instrument and Control Engineers*, **24**, 377-384 (in Japanese).

Sagara, S., and Z.Y. Zhao 1990, Numerical integration approach to on-line identification of continuous-time systems. *Automatica*, **26**, 63-74.

Sagara, S., and Z.Y. Zhao 1990, Identification of system parameters in distributed parameter systems. *11th IFAC World Congress*, Tallinn, USSR.

SAHA, D. C., and RAO, G. P., 1982, A general algorithm for parameter identification in lumped continuous systems–the Poisson moment functional approach. *I.E.E.E. Transactions on Automatic Control*, **27**, 223-225.

SAHA, D. C., and MANDAL, S. K., 1990, Recursive least-squares parameter estimation in SISO systems via Poisson moment functionals. Part 1. Open-loop systems. *International Journal of Systems Science*, **21**, 1205-1216.

SARKAR, T. K., RADHAKRISHNA, N., and CHEN, H., 1986, Survey of various z-domain to s-domain transformations. *I.E.E.E. Transactions on Instrumentation and Measurement*, **26**, 712-717.

SCHOUKENS, J., 1990, Modeling of Continuous time systems using a discrete time representation. *Automatica*, **26**, 5790-583.

SCHOUKENS, J., PINTELON, R., and RENNEBOOG, J., 1988, A maximum likelihood estimator for linear and nonlinear systems–a practical application of estimation techniques in measurement problems. *I.E.E.E. Transactions on Instrumentation and Measurement*, **37**, 10-17.

SHIEH, L. S., YEUNG, C. K., and MCINNIS, B. C., 1978, Solution of state-space equations via block-pulse functions. *International Journal of Control*, **28**, 383-392.

SINBROT, M., 1957, On the analysis of linear and nonlinear systems. *Transactions of ASME*, **79**, 547-552.

SINHA, N. K., 1972, Estimation of transfunction of continuous system frpm sampled data. *I.E.E. Proc.*, **119**, 612-614

SINHA, N. K., and LASTMAN, G. J., 1981, Transformation algorithm for identification of continuous-time multivariable systems from discrete data. *Electronic letters*, **17**, 779-780.

SINHA, N. K., and LASTMAN, G. J., 1982, Identification of continuous-time multivariable systems from discrete data. *International Journal of Control*, **35**, 117-126.

SINHA, N. K., and ZHOU, Q. J., 1985, Choice of the sampling interval for the identification of continuous-time systems from samples of input/output data. *Proceedings of the Institution of Electrical Engineers*, **D-132**, 263-267.

SINHA, N. K., and ZHOU, Q. J., 1983, Discrete-time approximation of multivariable continuous-time systems. *Proceedings of the Institution of Electrical Engineers*, **D-130**, 103-110.

SINHA, M. S. P., RAJAMANI, V. S., and SINHA, A. K., 1980, Identification of non-linear distributed system using Walsh functions. *International Journal of Control*, **32**, 669-676.

SÖDERSTRÖM, T., 1981, Identification of stochastic linear systems in presence of input noise. *Automatica*, **17**, 713-725.

SÖDERSTRÖM, T., LJUNG, L., and GUSTAVSSON, I., 1978, A theoretical analysis of recursive identification methods. *Automatica*, **14**, 231-244.

SÖDERSTRÖM, T., and STOICA, P., 1981, Comparison of some instrumental variable methods-Consistency and accuracy aspects. *Automatica*, **17**, 101-115.

Söderström, T., and Stoica P., 1983, *Instrumental variable Methods for System Identification* (Springer Verlag, Berlin).

Söderström, T., and Stoica, P., 1988, On system idenfication techniques for adaptive filtering. *I.E.E.E. Transactions on Circuits and Systems*, **30**, 457-461.

Söderström, T., and Stoica, P., 1989, *System Identification* (Prentice-Hall).

Steiglitz, K., and McBride, L. E., 1965, A technique for the identification of linear systems. *I.E.E.E. Transactions on Automatic Control*, **10**, 461-464.

Stoica, P., 1980, Comparison of two iterative estimation procedures. *Electronics letters*, **16**, 754-756.

Stoica, P., Eykhoff, P., Janssen, P., and Söderström, T., 1987, Model structure selection by cross validation. *International Journal of Control*, **43**, 1841-1878.

Stoica, P., and Nehorai, A., 1987, On the uniqueness of prediction error methods for systems with noisy input-output data. *Automatica*, **23**, 541-543.

Stoica, P., and Söderström, T., 1981, The Steiglitz-McBride algorithm revisited–convergence analysis and accuracy aspects. *I.E.E.E. Transactions on Automatic Control*, **26**, 712-717.

Stoica, P., and Söderström, T., 1981, Asymptotic behaviour of some bootstrap estimators. *International Journal of Control*, **33**, 433-454.

Stoica, P., and Söderström, T., 1982, Bias correction in least-squares identification. *International Journal of Control*, **35**, 449-457.

Unbehauen, H., and Rao, G. P., 1990, Continuous-time approaches to system identification–a survey. *Automatica*, **26**, 23-35.

Wada, K., and Eguchi, M. 1986, Estimation of pulse transfer function model via bias compensated least-squares method in the presence of input and output noises. *SICE 9-th Dynamical System Theory Symposium*, Kumamoto, Japan, 315-318 (in Japanese).

Wang, M. L., Yang, S. Y., and Chang, R. Y. 1987, New approach for parameter identification via generalized orthogonal polynomials. *International Journal of Systems Science*, **18**, 569-579.

Whitefield, A. H., and Messali, N., 1987, Integral-equation to system identification. *International Journal of Control*, **45**, 1431-1445.

WONG, K. Y., and POLAK, E., 1967, Identification of linear discrete time systems using the instrumental variable approach. *I.E.E.E. Transactions on Automatic Control*, **12**, 707-718.

YANG, Z. J., 1989, *Digital Filtering Approach to Parameter Estimation of Continuous Systems*. Master's thesis, Department of Electrical Engineering, Faculty of Engineering, Kyushu University (in Japanese).

YANG, Z. J., SAGARA, S., and WADA, K., 1991, System parameter identification from noisy sampled input-output data. *Proceedings of the 30th SICE Annual Conference Vol. 2*, Yonezawa, Japan, 981-984.

YOUNG, P., 1970, An instrumental variable method for real-time identification of a noisy process. *Automatica*, **6**, 271-287.

YOUNG, P., 1976, Some observations on instrumental variable methods of time-series analysis. *International Journal of Control*, **23**, 593-612.

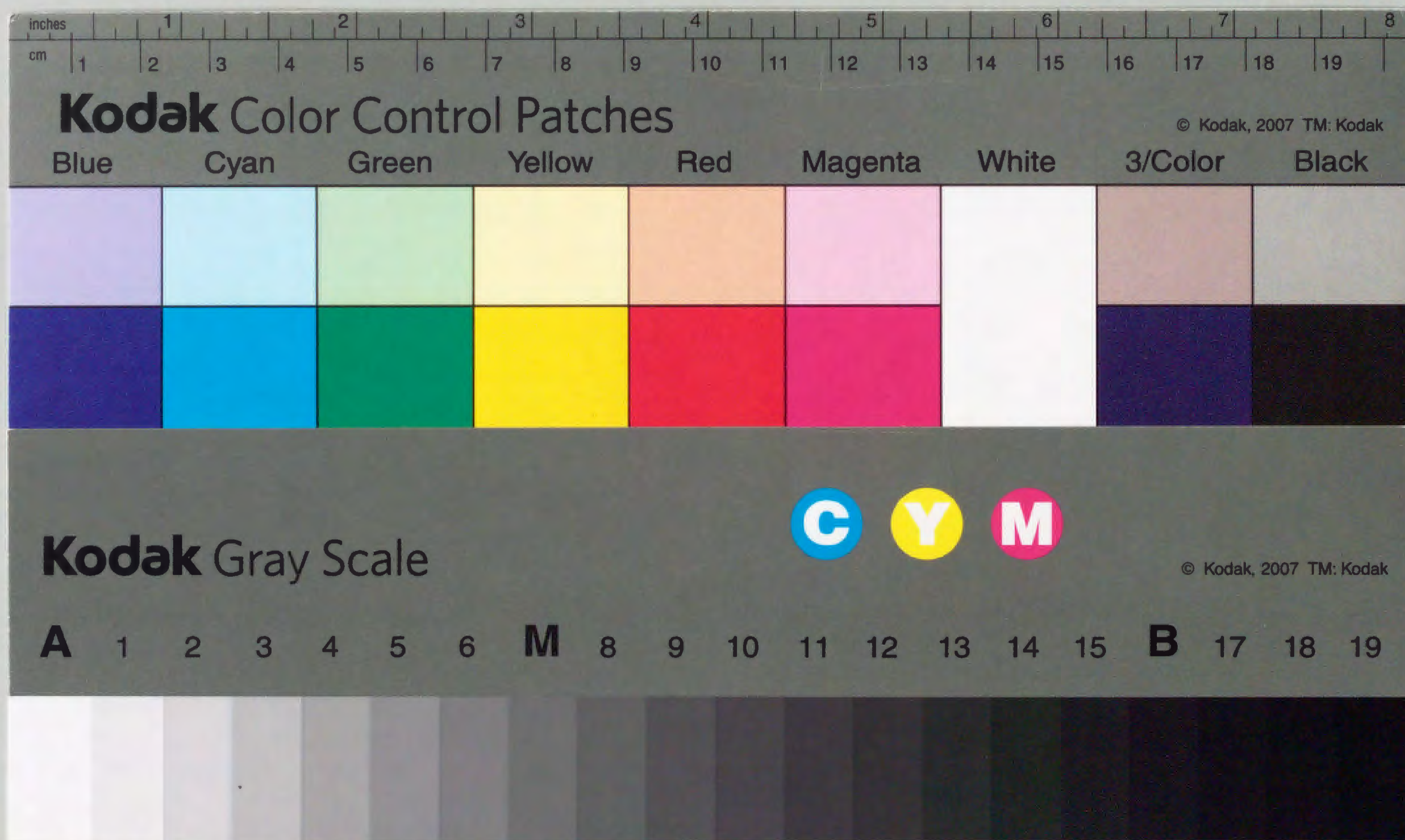
YOUNG, P., 1981, Parameter estimation for continuous-time models–a survey. *Automatica*, **17**, 23-39.

YOUNG, P., and JAKEMAN, A., 1979, Refined instrumental variable methods of recursive time-series analysis. Part I. Single input, single output systems. *International Journal of Control*, **29**, 1-30.

YOUNG, P., and JAKEMAN, A., 1980, Refined instrumental variable methods of recursive time-series analysis. Part III. Extensions. *International Journal of Control*, **31**, 741-764.

ZHENG, W. X., and FENG, C. B., 1989, Unbiased parameter estimation of linear systems in the presence of input and output noise. *International Journal of Adaptive Control and Signal Processing*, **3**, 231-251.

ZHU, J. M., and LU, Y. Z., 1988, Application of single-step method of block-pulse functions to the optimal control of linear distributed parameter system. *International Journal of Systems Science*, **19**, 2459-2472.


inches
cm
Kodak Color Control Patches
© Kodak, 2007 TM: Kodak
Blue
Cyan
Green
Yellow
Red
Magenta
White
3/Color
Black
Kodak Gray Scale
C
Y
M
© Kodak, 2007 TM: Kodak
A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19