On Bias-Compensated Least-Squares Methods Using α -filter and β -filter

Masato Ikenoue*, Shunshoku Kanae**, Zi-Jiang Yang** and Kiyoshi Wada**

* Department of Electrical Engineering, Ariake National College of Technology
150, Higashihagio-machi, Omuta, Fukuoka 836-8585, JAPAN
TEL: +81-944-53-8875 FAX: +81-944-53-8875
E-mail: ikenoue@ariake-nct.ac.jp

** Department of Electrical and Electronic Systems Engineering, Kyushu University
6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, JAPAN
TEL: +81-92-642-3958 FAX: +81-92-642-3939
E-mail: {jin, yoh, wada}@ees.kyushu-u.ac.jp

Abstract

In this paper, two bias-compensated least-squares (BCLS) methods (BCLS- α method and BCLS- β method) are proposed for identification of linear discrete-time system in the case where the input measurement is corrupted by white noise and the output measurement is corrupted by colored noise. It is well known that BCLS method is based on compensation of asymptotic bias on the least-squares (LS) estimate by making use of noise variances estimates. The main feature of proposed algorithms is to introduce an auxiliary multivariate estimator using filter (α -filter or β -filter) in order to estimate input noise variance and output noise covariances. Some simulation results indicate that the proposed methods provide good estimates.

1 Introduction

Many identification methods are based on the assumption that input measurement is noise-free. However, this condition is not satisfied in most practical situations. In the presence of input noise, those methods have been shown to give erroneous results. Several methods have been proposed to estimate unknown parameters of linear discrete-time system in the presence of input and output noise. Joint Output (JO) method [1] and Koopmans-Levin (KL) method [2] require *a priori* knowledge about the values of variances or the ratio to measurement noises.

Bias-compensated least-squares (BCLS) method is proposed by Sagara and Wada [3] and it has been extended by Wada *et al.* [4] to the input-output noise case without any *a priori* knowledge of noise variances. BCLS method based on compensation of asymptotic bias on the least-squares (LS) estimate by making use of noise variances estimates can give consistent estimates for unknown parameters of linear discrete-time system in the presence of input and output noises. In recent years, BCLS method has been developed to improve the

estimation accuracy and several algorithms have been proposed [5] [6] [7]. However, these BCLS methods can only apply to the case where both the input noise and the output noise are white.

On the other hand, another method named bias-eliminated least-squares (BELS) method has been proposed by Feng and Zheng [8] in which the different estimation method of asymptotic bias is used and further developed to be the efficient method [9] [10] [11] to treat bias problem in noisy input-output system identification.

In this paper, two BCLS methods (BCLS- α method and BCLS- β method) are proposed for identification of linear discrete-time system in the case where the input measurement is corrupted by white noise and the output measurement is corrupted by colored noise. Since the BCLS method requires the consistent estimate of the unknown input noise variance and the unknown output noise covariances to compensate the bias of the LS estimate, the estimation of the input noise variance and the output noise covariances plays an important role in BCLS method. In order to estimate input noise variance and output noise covariances, an auxiliary multivariate estimator using filter (α -filter or β -filter) is introduced. Some simulation results indicate that the proposed methods provide good estimates.

This paper is organized as follows. In section 2 and section 3, the problem statement is presented and the asymptotic bias of LS estimator is described. In section 4, the BCLS algorithm is derived for estimating unknown parameters of linear discrete-time system in the presence of input and output noises and it can be learned that the unknown variance of white input noise and the unknown covariances of colored output noise must be estimated in order to obtain consistent estimates of parameters. In section 5, two types of algorithms of unknown input noise variance and output noise covariances are derived. The simulation results are presented in section 6 and finally section 7 gives the conclusion.

2 Problem Statement

Consider the parameter estimation problem of single-input single-output linear discrete-time system described as follows:

$$A(q^{-1})y_t = B(q^{-1})u_t \quad (1)$$

where u_t and y_t are the true input and output, q^{-1} is shift operator, $q^{-1}u_t = u_{t-1}$, and the polynomials $A(q^{-1})$ and $B(q^{-1})$ are defined by

$$\begin{aligned} A(q^{-1}) &= 1 + a_1q^{-1} + \cdots + a_nq^{-n} \\ B(q^{-1}) &= b_1q^{-1} + \cdots + b_nq^{-n} \end{aligned}$$

Let z_t and w_t be the noise-corrupted measurements of y_t and u_t , respectively, i.e.

$$z_t = y_t + e_t, \quad w_t = u_t + d_t \quad (2)$$

where e_t and d_t are the output and input measurement noises, respectively. The output noise e_t is assumed to be zero-mean colored noise with unknown covariances

$$r_{ee}(k) = E[e_t e_{t-k}], \quad (k = 0, \pm 1, \pm 2, \dots)$$

and described as

$$e_t = H(q^{-1})\omega_t$$

where ω_t is a zero-mean white noise with unknown variance σ_ω^2 , and $H(q^{-1})$ takes up any form of rational function. On the other hand, the input noise d_t is a zero-mean white noise with unknown variance σ_d^2 . The true input u_t is a zero-mean stationary random process with finite variance, and u_t , d_t and e_t are assumed to be statistically independent of each other.

Substituting (2) into (1) yields

$$A(q^{-1})z_t = B(q^{-1})w_t + v_t \quad (3)$$

where v_t is a composite noise defined by

$$v_t = A(q^{-1})e_t - B(q^{-1})d_t \quad (4)$$

Define some vectors as

$$\begin{aligned} \theta^T &= [a^T, b^T] = [a_1 \cdots a_n, b_1 \cdots b_n] \\ p_t^T &= [-z_t^T, w_t^T] \\ &= [-z_{t-1} \cdots -z_{t-n}, w_{t-1} \cdots w_{t-n}] \\ q_t^T &= [-y_t^T, u_t^T] \\ &= [-y_{t-1} \cdots -y_{t-n}, u_{t-1} \cdots u_{t-n}] \\ r_t^T &= [-e_t^T, d_t^T] \\ &= [-e_{t-1} \cdots -e_{t-n}, d_{t-1} \cdots d_{t-n}] \end{aligned}$$

then (2), (3) and (4) can be written as

$$p_t = q_t + r_t \quad (5)$$

$$z_t = p_t^T \theta + v_t \quad (6)$$

$$v_t = e_t - r_t^T \theta \quad (7)$$

Let the equation error ξ_t for an estimate $\hat{\theta}$ of θ be defined as

$$\begin{aligned} \xi_t &= \hat{A}(q^{-1})z_t - \hat{B}(q^{-1})w_t \\ &= z_t - p_t^T \hat{\theta} \end{aligned}$$

where the polynomials $\hat{A}(q^{-1})$ and $\hat{B}(q^{-1})$ are defined by

$$\begin{aligned} \hat{A}(q^{-1}) &= 1 + \hat{a}_1q^{-1} + \cdots + \hat{a}_nq^{-n} \\ \hat{B}(q^{-1}) &= \hat{b}_1q^{-1} + \cdots + \hat{b}_nq^{-n} \end{aligned}$$

and

$$\hat{\theta}^T = [\hat{a}^T, \hat{b}^T] = [\hat{a}_1 \cdots \hat{a}_n, \hat{b}_1 \cdots \hat{b}_n]$$

The least-squares estimate is given by

$$\hat{\theta}_{LS,N} = \left(\sum_{t=1}^N p_t p_t^T \right)^{-1} \sum_{t=1}^N p_t z_t \quad (8)$$

From the assumption of e_t and d_t , the composite noise v_t defined by (4) is not white. Hence the least-squares estimate $\hat{\theta}_{LS,N}$ has a bias asymptotically. In the next section, the asymptotic bias induced by least-squares estimator is derived.

3 Asymptotic Bias of Least-Squares Estimator

Substituting (6) into (8) yields

$$\hat{\theta}_{LS,N} = \theta + P_N \sum_{t=1}^N p_t v_t$$

where P_N is the product moment matrix as follows:

$$P_N = \left(\sum_{t=1}^N p_t p_t^T \right)^{-1}$$

Taking limit of above equation yields

$$\lim_{N \rightarrow \infty} \hat{\theta}_{LS,N} = \theta + h \quad (9)$$

where h is the asymptotic bias of the least-squares estimate $\hat{\theta}_{LS,N}$ defined as

$$h = R_{pp}^{-1} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N p_t v_t \quad (10)$$

where $R_{pp} = E[p_t p_t^T]$.

Using the assumption of e_t , d_t and (5), (7), it is easily shown that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N p_t v_t &= E[p_t v_t] \\ &= E[(q_t + r_t)(e_t - r_t^T \theta)] \\ &= E[r_t e_t] - E[r_t r_t^T] \theta \\ &= d - D\theta \end{aligned} \quad (11)$$

where

$$\begin{aligned} \mathbf{d} &= E[\mathbf{r}_t \mathbf{e}_t] = \begin{bmatrix} -r_{ee} \\ \mathbf{0}_n \end{bmatrix} \\ \mathbf{D} &= E[\mathbf{r}_t \mathbf{r}_t^T] = \begin{bmatrix} \mathbf{R}_{ee} & \mathbf{O}_n \\ \mathbf{O}_n & \sigma_d^2 \mathbf{I}_n \end{bmatrix} \\ r_{ee} &= E[\mathbf{e}_t \mathbf{e}_t] = [r_{ee}(1) \cdots r_{ee}(n)]^T \\ \mathbf{R}_{ee} &= E[\mathbf{e}_t \mathbf{e}_t^T] = \begin{bmatrix} r_{ee}(0) & r_{ee}(1) & \cdots & r_{ee}(n-1) \\ r_{ee}(1) & r_{ee}(0) & \cdots & r_{ee}(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{ee}(n-1) & r_{ee}(n-2) & \cdots & r_{ee}(0) \end{bmatrix} \end{aligned}$$

$\mathbf{I}_n$ is an $n \times n$ identity matrix, $\mathbf{O}_n$ is an $n \times n$ zero matrix, and $\mathbf{0}_n$ is an $n \times 1$ zero vector. From (9), (10) and (11), the asymptotic bias $\mathbf{h}$ can be expressed as follows:

$$\mathbf{h} = \lim_{N \rightarrow \infty} \hat{\boldsymbol{\theta}}_{LS,N} - \boldsymbol{\theta} = -\mathbf{R}_{pp}^{-1} (\mathbf{D}\boldsymbol{\theta} - \mathbf{d}) .$$

Additionally, $\mathbf{D}\boldsymbol{\theta} - \mathbf{d}$ can be expressed as follows:

$$\begin{aligned} \mathbf{D}\boldsymbol{\theta} - \mathbf{d} &= \begin{bmatrix} \mathbf{R}_{ee} & \mathbf{O}_n \\ \mathbf{O}_n & \sigma_d^2 \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} - \begin{bmatrix} -r_{ee} \\ \mathbf{0}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{R}_{ee}\mathbf{a} + r_{ee} \\ \sigma_d^2 \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 \boldsymbol{\gamma} + \mathbf{Q}_2 \boldsymbol{\gamma} \\ \mathbf{Q}_3 \boldsymbol{\gamma} \end{bmatrix} \\ &= \mathbf{Q} \boldsymbol{\gamma} \end{aligned}$$

where

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} \mathbf{Q}_1 + \mathbf{Q}_2 \\ \mathbf{Q}_3 \end{bmatrix} \in \mathbb{R}^{2n \times (n+2)} \\ \mathbf{Q}_1 &= \mathbf{a} \mathbf{i}_1^T + \sum_{j=1}^{n-1} [(S_n)^j + (S_n^T)^j] \mathbf{a} \mathbf{i}_{j+1}^T \in \mathbb{R}^{n \times (n+2)} \\ \mathbf{Q}_2 &= [\mathbf{0}_n \ \mathbf{I}_n \ \mathbf{0}_n] \in \mathbb{R}^{n \times (n+2)} \\ \mathbf{Q}_3 &= \mathbf{b} \mathbf{i}_{n+2}^T \in \mathbb{R}^{n \times (n+2)} \\ \mathbf{i}_j &= \mathbf{I}_{n+2}(:, j) \in \mathbb{R}^{n+2} \ (j = 1, \dots, n+2) \\ S_n &= \begin{bmatrix} \mathbf{0}_{n-1}^T & 0 \\ \mathbf{I}_{n-1} & \mathbf{0}_{n-1} \end{bmatrix} \in \mathbb{R}^{n \times n} \\ \boldsymbol{\gamma}^T &= [r_{ee}(0), \dots, r_{ee}(n), \sigma_d^2] \in \mathbb{R}^{n+2} . \end{aligned}$$

Thus the asymptotic bias $\mathbf{h}$ can be expressed as follows:

$$\mathbf{h} = -\mathbf{R}_{pp}^{-1} \mathbf{Q} \boldsymbol{\gamma} . \quad (12)$$

4 Bias-Compensated Least-Squares Method

From (9), it can be expected that a consistent estimate of $\boldsymbol{\theta}$ can be obtained by compensating for the asymptotic bias $\mathbf{h}$ in the least-squares estimate $\hat{\boldsymbol{\theta}}_{LS,N}$.

From (12), estimate of the asymptotic bias $\mathbf{h}$ becomes

$$\hat{\mathbf{h}}_N = -N \mathbf{P}_N \mathbf{Q} \boldsymbol{\gamma} .$$

Hence bias-compensated least-squares estimate $\hat{\boldsymbol{\theta}}_{BC,N}$ is given by

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{BC,N} &= \hat{\boldsymbol{\theta}}_{LS,N} - \hat{\mathbf{h}}_N \\ &= \hat{\boldsymbol{\theta}}_{LS,N} + N \mathbf{P}_N \hat{\mathbf{Q}}_{BC,N-1} \boldsymbol{\gamma} \end{aligned} \quad (13)$$

where

$$\begin{aligned} \hat{\mathbf{Q}}_{BC,N} &= \begin{bmatrix} \hat{\mathbf{Q}}_{1,BC,N} + \mathbf{Q}_2 \\ \hat{\mathbf{Q}}_{3,BC,N} \end{bmatrix} \\ \hat{\mathbf{Q}}_{1,BC,N} &= \hat{\mathbf{a}}_{BC,N} \mathbf{i}_1^T + \sum_{j=1}^{n-1} [(S_n)^j + (S_n^T)^j] \hat{\mathbf{a}}_{BC,N} \mathbf{i}_{j+1}^T \\ \hat{\mathbf{Q}}_{3,BC,N} &= \hat{\mathbf{b}}_{BC,N} \mathbf{i}_{n+2}^T \\ \hat{\boldsymbol{\theta}}_{BC,N}^T &= [\hat{\mathbf{a}}_{BC,N}^T, \hat{\mathbf{b}}_{BC,N}^T] \\ &= [\hat{a}_{1,BC,N} \cdots \hat{a}_{n,BC,N}, \hat{b}_{1,BC,N} \cdots \hat{b}_{n,BC,N}] . \end{aligned}$$

Since the noise covariance vector $\boldsymbol{\gamma}$ in (13) is unknown, it is necessary to estimate $\boldsymbol{\gamma}$.

5 Estimation of $\boldsymbol{\gamma}$

Consider the sum of squared equation errors for $\hat{\boldsymbol{\theta}}_{LS,N}$ defined by

$$\begin{aligned} \hat{g}_N &= \sum_{t=1}^N \hat{\xi}_t^2 = \sum_{t=1}^N (z_t - \mathbf{p}_t^T \hat{\boldsymbol{\theta}}_{LS,N})^2 \\ &= \sum_{t=1}^N z_t v_t - \hat{\boldsymbol{\theta}}_{LS,N}^T \sum_{t=1}^N \mathbf{p}_t v_t \end{aligned} \quad (14)$$

where $\hat{\xi}_t$ is the equation error for the least-squares estimate $\hat{\boldsymbol{\theta}}_{LS,N}$ defined by

$$\hat{\xi}_t = z_t - \mathbf{p}_t^T \hat{\boldsymbol{\theta}}_{LS,N} .$$

Taking limit of (14) yields

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N z_t v_t &= E[z_t v_t] = r_{ee}(0) - \mathbf{d}^T \boldsymbol{\theta} \\ &= (\mathbf{i}_1^T + \mathbf{a}^T \mathbf{Q}_2) \boldsymbol{\gamma} \end{aligned}$$

and

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{p}_t v_t = E[\mathbf{p}_t v_t] = \mathbf{d} - \mathbf{D}\boldsymbol{\theta} = -\mathbf{Q} \boldsymbol{\gamma} .$$

Finally,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \hat{g}_N = (\mathbf{i}_1^T + \mathbf{a}^T \mathbf{Q}_2 + \lim_{N \rightarrow \infty} \hat{\boldsymbol{\theta}}_{LS,N}^T \mathbf{Q}) \boldsymbol{\gamma} . \quad (15)$$

Thus the equation (15) with respect to $(n+2) \times 1$ noise covariance vector $\boldsymbol{\gamma}$ can be obtained. If another $m+1$ ($m \geq n$) equations with respect to $\boldsymbol{\gamma}$ can be obtained, $\boldsymbol{\gamma}$ can be obtained by solutions of system of equations.

5.1 Estimation method using α -filter (BCLS- α method)

To obtain another $m + 1$ equations, α -filter defined by the following equation is introduced

$$\alpha(q^{-1}) = \alpha_0 + \dots + \alpha_{n_\alpha} q^{-n_\alpha} = \sum_{i=0}^{n_\alpha} \alpha_i q^{-i}.$$

Define three filtered signals for w_t , u_t and d_t as

$$\tilde{w}_t = \frac{1}{\alpha(q^{-1})} w_t, \quad \tilde{u}_t = \frac{1}{\alpha(q^{-1})} u_t, \quad \tilde{d}_t = \frac{1}{\alpha(q^{-1})} d_t$$

and define three vectors as

$$\begin{aligned} \tilde{\mathbf{w}}_t &= [\tilde{w}_t \ \tilde{w}_{t-1} \ \dots \ \tilde{w}_{t-m}]^T \\ \tilde{\mathbf{u}}_t &= [\tilde{u}_t \ \tilde{u}_{t-1} \ \dots \ \tilde{u}_{t-m}]^T \\ \tilde{\mathbf{d}}_t &= [\tilde{d}_t \ \tilde{d}_{t-1} \ \dots \ \tilde{d}_{t-m}]^T \end{aligned}$$

where $m \geq n$. It follows from (2) that

$$\tilde{\mathbf{w}}_t = \tilde{\mathbf{u}}_t + \tilde{\mathbf{d}}_t.$$

Since d_t is a zero-mean white noise, the filtered signal $\tilde{d}_t$ is an autoregressive (AR) process of order n_α

$$\alpha(q^{-1}) \tilde{d}_t = d_t. \quad (16)$$

Multiplying (16) by d_{t-k} ($k = 0, \pm 1, \pm 2, \dots$), and taking the expectation yield

$$\sum_{j=0}^{n_\alpha} \alpha_j r_{\tilde{d}d}(k-j) = \begin{cases} \sigma_d^2 & , k = 0 \\ 0 & , k = \pm 1, \pm 2, \dots \end{cases} \quad (17)$$

where

$$\begin{aligned} r_{\tilde{d}d}(k) &= E[\tilde{d}_t d_{t-k}], \quad (k = 0, \pm 1, \pm 2, \dots) \\ r_{\tilde{d}d}(k) &= 0, \quad (k < 0). \end{aligned} \quad (18)$$

Let $\psi(k) = r_{\tilde{d}d}(k)/\sigma_d^2$, then (17) and (18) can be written as

$$\begin{aligned} \left[\sum_{j=0}^n \alpha_j (S_{n+1}^T)^j \right] \psi_n(n) &= \begin{bmatrix} \mathbf{0}_n \\ 1 \end{bmatrix} \\ \psi(k) &= 0, \quad (k < 0) \end{aligned}$$

where $\psi_n(n) = [\psi(n) \ \psi(n-1) \ \dots \ \psi(0)]^T$.

Consider an auxiliary multivariate estimate $\hat{\Phi}_N$ defined by the following equation

$$\hat{\Phi}_N = \left(\sum_{t=1}^N \tilde{\mathbf{w}}_t \mathbf{p}_t^T \right) \left(\sum_{t=1}^N \mathbf{p}_t \mathbf{p}_t^T \right)^{-1}.$$

Define $\hat{\mathbf{f}}_N$ by

$$\begin{aligned} \hat{\mathbf{f}}_N &= \sum_{t=1}^N (\tilde{\mathbf{w}}_t - \hat{\Phi}_N \mathbf{p}_t) \hat{\xi}_t \\ &= \sum_{t=1}^N (\tilde{\mathbf{w}}_t - \hat{\Phi}_N \mathbf{p}_t) (z_t - \mathbf{p}_t^T \hat{\theta}_{LS,N}) \\ &= \sum_{t=1}^N (\tilde{\mathbf{w}}_t - \hat{\Phi}_N \mathbf{p}_t) z_t = \sum_{t=1}^N \tilde{\mathbf{w}}_t v_t - \hat{\Phi}_N \sum_{t=1}^N \mathbf{p}_t v_t. \end{aligned} \quad (19)$$

Taking limit of (19) yields

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \tilde{\mathbf{w}}_t v_t &= E[\tilde{\mathbf{w}}_t v_t] \\ &= E[(\tilde{\mathbf{u}}_t + \tilde{\mathbf{d}}_t)(e_t - \mathbf{r}_t^T \boldsymbol{\theta})] \\ &= -E[\tilde{\mathbf{d}}_t \mathbf{d}_t^T] \mathbf{b} \\ &= -\sigma_d^2 \boldsymbol{\Psi} \mathbf{b} = -\boldsymbol{\Psi} \mathbf{Q}_3 \boldsymbol{\gamma} \end{aligned}$$

and

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{p}_t v_t = E[\mathbf{p}_t v_t] = \mathbf{d} - \mathbf{D} \boldsymbol{\theta} = -\mathbf{Q} \boldsymbol{\gamma}$$

where

$$\begin{aligned} \boldsymbol{\Psi} &= [\psi_m(1) \ \psi_m(2) \ \dots \ \psi_m(n)] \\ \psi_m(k) &= [\psi(k) \ \psi(k-1) \ \dots \ \psi(k-m)]^T \\ &= \begin{bmatrix} (S_{n+1}^T)^{n-k} \\ \mathbf{O}_{(m-n) \times (n+1)} \end{bmatrix} \psi_n(n) \\ &\quad , \quad (k = 1, \dots, n) \\ \psi_n(n) &= \left[\sum_{j=0}^n \alpha_j (S_{n+1}^T)^j \right]^{-1} \begin{bmatrix} \mathbf{0}_n \\ 1 \end{bmatrix}. \end{aligned}$$

Finally, another $m + 1$ equations with respect to $\boldsymbol{\gamma}$ can be obtained

$$\lim_{N \rightarrow \infty} \frac{1}{N} \hat{\mathbf{f}}_N = \left(-\boldsymbol{\Psi} \mathbf{Q}_3 + \lim_{N \rightarrow \infty} \hat{\Phi}_N \mathbf{Q} \right) \boldsymbol{\gamma}. \quad (20)$$

5.2 Estimation method using β -filter (BCLS- β method)

To obtain another $m + 1$ equations, β -filter defined by the following equation is introduced

$$\beta(q^{-1}) = \beta_0 + \dots + \beta_{2n_\beta} q^{-2n_\beta} = \sum_{i=0}^{2n_\beta} \beta_i q^{-i}.$$

Define three filtered signals for w_t , u_t and d_t as

$$\tilde{w}_t = \beta(q^{-1}) w_t, \quad \tilde{u}_t = \beta(q^{-1}) u_t, \quad \tilde{d}_t = \beta(q^{-1}) d_t$$

and define three vectors as

$$\begin{aligned} \tilde{\mathbf{w}}_t &= [\tilde{w}_t \ \dots \ \tilde{w}_{t-m}]^T \\ \tilde{\mathbf{u}}_t &= [\tilde{u}_t \ \dots \ \tilde{u}_{t-m}]^T \\ \tilde{\mathbf{d}}_t &= [\tilde{d}_t \ \dots \ \tilde{d}_{t-m}]^T \end{aligned}$$

where $m \geq n$. It follows from (2) that

$$\tilde{\mathbf{w}}_t = \tilde{\mathbf{u}}_t + \tilde{\mathbf{d}}_t.$$

Since d_t is a zero-mean white noise, the filtered signal $\tilde{d}_t$ is a moving average (MA) process of order $2n_\beta$. Multiplying $\tilde{d}_t$ by d_{t-k} ($k = 0, \pm 1, \pm 2, \dots$), and taking the expectation yield

$$r_{\tilde{d}d}(k) = E[\tilde{d}_t d_{t-k}] = \begin{cases} \sigma_d^2 \beta_k & , k = 0, \dots, 2n_\beta \\ 0 & , \text{otherwise} \end{cases} \quad (21)$$

Let $\psi(k) = r_{dd}(k)/\sigma_d^2$, then (21) can be written as

$$\psi(k) = \begin{cases} \beta_k & , k = 0, \dots, 2n_\beta \\ 0 & , \text{otherwise} \end{cases}$$

Consider an auxiliary multivariate estimate $\hat{\Phi}_N$ defined by following equation

$$\hat{\Phi}_N = \left(\sum_{t=1}^N \tilde{w}_t \tilde{p}_t^T \right) \left(\sum_{t=1}^N \tilde{p}_t \tilde{p}_t^T \right)^{-1}$$

where $\tilde{p}_t = p_{t-n_\beta}$. Define $\hat{f}_N$ by

$$\begin{aligned} \hat{f}_N &= \sum_{t=1}^N (\tilde{w}_t - \hat{\Phi}_N \tilde{p}_t) \hat{\xi}_t \\ &= \sum_{t=1}^N (\tilde{w}_t - \hat{\Phi}_N \tilde{p}_t) (\bar{z}_t - \bar{p}_t^T \hat{\theta}_{LS,N}) \\ &= \sum_{t=1}^N (\tilde{w}_t - \hat{\Phi}_N \tilde{p}_t) \bar{z}_t = \sum_{t=1}^N \tilde{w}_t \bar{v}_t - \hat{\Phi}_N \sum_{t=1}^N \tilde{p}_t \bar{v}_t \end{aligned} \quad (22)$$

where $\hat{\xi}_t = \bar{z}_t - \bar{p}_t^T \hat{\theta}_{LS,N}$, $\bar{z}_t = z_{t-n_\beta}$, $\bar{v}_t = v_{t-n_\beta}$

$$\hat{\theta}_{LS,N} = \left(\sum_{t=1}^N \tilde{p}_t \tilde{p}_t^T \right)^{-1} \sum_{t=1}^N \tilde{p}_t \bar{z}_t$$

Taking limit of (22) yields

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \tilde{w}_t \bar{v}_t &= E[\tilde{w}_t v_{t-n_\beta}] \\ &= E[(\tilde{u}_t + \tilde{d}_t)(e_{t-n_\beta} - r_{t-n_\beta}^T \theta)] \\ &= -E[\tilde{d}_t d_{t-n_\beta}^T] b \\ &= -\sigma_d^2 \Psi b = -\Psi Q_3 \gamma \end{aligned}$$

and

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \tilde{p}_t \bar{v}_t &= E[p_{t-n_\beta} v_{t-n_\beta}] = E[p_t v_t] \\ &= d - D\theta = -Q\gamma \end{aligned}$$

where

$$\begin{aligned} \Psi &= [\psi_m(n_\beta+1) \dots \psi_m(n_\beta+n)] \\ \psi_m(k) &= [\psi(k), \dots, \psi(k-m)]^T \\ &, (k = n_\beta+1, \dots, n_\beta+n) \end{aligned}$$

Finally, another $m+1$ equations with respect to γ can be obtained

$$\lim_{N \rightarrow \infty} \frac{1}{N} \hat{f}_N = (-\Psi Q_3 + \lim_{N \rightarrow \infty} \hat{\Phi}_N Q) \gamma \quad (23)$$

5.3 Estimate of γ

From the combination of (15) and (20) or (15) and (23), the estimate of the unknown noise covariance vector γ can be obtained by the solution of the system of equations such as

$$\begin{bmatrix} i_1^T + \hat{a}_{BC,N-1}^T Q_2 + \hat{\theta}_{LS,N}^T \hat{Q}_{BC,N-1} \\ -\Psi \hat{Q}_{3,BC,N-1} + \hat{\Phi}_N \hat{Q}_{BC,N-1} \end{bmatrix} \hat{\gamma}_N = \frac{1}{N} \begin{bmatrix} \hat{g}_N \\ \hat{f}_N \end{bmatrix} \quad (24)$$

In order to have a unique solution of γ , m is set to be $m = n$ so that (24) is composed of $n+2$ equations for $n+2$ unknowns. In the case of $m > n$, (24) represents a system of overdetermined equations. Since the noise covariance vector γ is estimated by using $m-n$ extra equations which are additional informations about the identified system, γ can be estimated more accurately and more quickly. Then it can be expected that the estimation accuracy of the resulting BCLS estimates are improved.

6 Simulation Results

Computer simulations which compare proposed algorithms with LS algorithm have been carried out. Consider the following second-order system:

$$\frac{B(q^{-1})}{A(q^{-1})} = \frac{0.169901q^{-1} + 0.143831q^{-2}}{1 - 1.575157q^{-1} + 0.606531q^{-2}}$$

The noise free input u_t is generated as

$$u_t = 1.0\epsilon_t - 0.3\epsilon_{t-1} + 0.5\epsilon_{t-2} - 0.7\epsilon_{t-3} + 0.9\epsilon_{t-4}$$

where ϵ_t is a zero-mean white noise with unit variance $\sigma_\epsilon^2 = 1$. The white input noise d_t is with variance σ_d^2 , which yields $\text{SNR} = 10 \log_{10}(E[u_t^2]/E[d_t^2]) = 10$ [dB]. The colored output noise e_t is modeled by

$$e_t = \frac{1 - 0.87q^{-1} + 0.57q^{-2}}{1 - 1.0q^{-1} + 0.4q^{-2}} \omega_t$$

where ω_t is a zero-mean white noise with variance σ_ω^2 , which yields $\text{SNR} = 10 \log_{10}(E[y_t^2]/E[e_t^2]) = 10$ [dB]. The filter $\alpha(q^{-1})$ is designed as

$$\begin{aligned} \alpha(q^{-1}) &= (1 - 0.6q^{-1})(1 - 0.75q^{-1})(1 - 0.9q^{-1}) \\ &= 1 - 2.25q^{-1} + 1.665q^{-2} - 0.405q^{-3} \end{aligned}$$

and the filter $\beta(q^{-1})$ is designed as

$$\beta(q^{-1}) = 1 + 2q^{-1} + 2q^{-2} + q^{-3} + 2q^{-4} + 2q^{-5} + q^{-6}$$

Computer simulation for comparison are carried out through $M = 100$ independent runs with a data length $N = 10000$. In particular, the cases of $m = 2, 3$ are examined for the proposed algorithms. Fig. 1 gives a plot of the mean value of the relative error (RE) which is defined by

$$\text{RE} = \frac{1}{M} \sum_{k=1}^M \frac{\|\hat{\theta}_{k,t} - \theta\|}{\|\theta\|}$$

Fig. 1: RE of parameter estimates.

Table 1: Mean values, standard deviations and RE of parameter estimates for $N = 10000$.

Parameter	a_1	a_2	b_1	b_2	RE
LS	-0.5325 ± 0.0083	-0.4015 ± 0.0080	0.1819 ± 0.0064	0.2160 ± 0.0064	0.8529
BCLS- α ($m = 2$)	-1.5749 ± 0.0295	0.6066 ± 0.0263	0.1745 ± 0.0534	0.1455 ± 0.0502	0.0404
BCLS- α ($m = 3$)	-1.5748 ± 0.0297	0.6065 ± 0.0267	0.1738 ± 0.0469	0.1448 ± 0.0446	0.0368
BCLS- β ($m = 2$)	-1.5785 ± 0.0496	0.6112 ± 0.0625	0.1728 ± 0.0281	0.1432 ± 0.0264	0.0455
BCLS- β ($m = 3$)	-1.5777 ± 0.0407	0.6099 ± 0.0476	0.1733 ± 0.0205	0.1436 ± 0.0170	0.0334
True values	-1.5752	0.6065	0.1699	0.1438	

where $\hat{\theta}_{k,t}$ denotes the estimate of θ at time step t in the k th independent run. Table 1 provides the mean values, the standard deviations and RE of the parameter estimates for $N = 10000$.

Simulation results indicate that LS method gives biased results. On the contrary, the proposed methods can give consistent estimate. Especially, the resulting BCLS estimates in the case of $m > n$ are more accurate than those obtained with $m = n$.

7 Conclusion

In this paper, the method of consistent estimation of linear discrete-time system in the presence of input and output noises has been studied. Two BCLS methods which can apply to the case of the white input noise and colored output noise are proposed. The two types of auxiliary multivariate estimators with filter are introduced for estimation of input noise variance and output noise covariances. It is demonstrated that the proposed

two algorithms can give consistent estimates via simulation results.

References

- [1] T. Söderström: Identification of stochastic linear systems in presence of input noise, *Automatica*, Vol.17, No.5, pp.713-725, 1981.
- [2] K. V. Fernando and H. Nicholson: Identification of linear systems with input and output noise: the Koopmans-Levin method, *IEE Proc. Control Theory and Applications*, Vol.132, pp.30-36, 1985.
- [3] S. Sagara, and K. Wada: On-line modified least-squares parameter estimation on linear discrete dynamic systems, *Int. J. Control*, Vol.25, No.3, pp.329-343, 1977.
- [4] K. Wada, M. Eguchi and S. Sagara: Estimation of pulse transfer function via bias-compensated least-squares method in the presence of input and output noise, *Systems Science*, Vol.16, No.3, pp.57-70, 1990.
- [5] M. Eguchi, K. Wada and S. Sagara: Identification of pulse transfer function in the presence of input and output noise, *IFAC international workshop on ACQP'92*, Vol.2, pp.463-470, 1992.
- [6] L. J. Jia, M. Ikenoue, C. Z. Jin and K. Wada: On bias compensated least squares method for noisy input-output system identification, *Proc. of 40th IEEE Conf. on Decision and Control*, pp.3332-3337, 2001.
- [7] M. Ikenoue, S. Kanae, Z. J. Yang, K. Wada: On the parameter estimation of pulse transfer function in the presence of input and output noise, *IEEJ Trans. EIS*, Vol.124, No.4, pp.981-988, 2004 (in Japanese).
- [8] W. X. Zheng and C. B. Feng: Unbiased estimation of linear systems in the presence of input and output noise, *Int. J. Adaptive Control and Signal Processing*, Vol.3, pp.231-251, 1989.
- [9] W. X. Zheng: On least-squares identification of stochastic linear systems with noisy input-output data, *Int. J. Adaptive Control and Signal Processing*, Vol.13, pp.131-143, 1999.
- [10] W. X. Zheng: Unbiased identification of stochastic linear systems from noisy input and output measurements, *Proc. of 39th IEEE Conf. on Decision and Control*, pp.2710-2715, 2000.
- [11] W. X. Zheng: A bias correction method for identification of linear dynamic errors-in-variables models, *IEEE Trans. Automat. Contr.*, Vol.47, No.7, pp.1142-1147, 2002.