Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems – MTNS 2010 • 5–9 July, 2010 • Budapest, Hungary

Bias-Compensated State Space Model Identification

Kenji Ikeda
The University of Tokushima
Tokushima, Japan 770-8506
Email: ikeda@is.tokushima-u.ac.jp

Yoshio Mogami
The University of Tokushima
Tokushima, Japan 774-8506
Email: moga@is.tokushima-u.ac.jp

Takao Shimomura
The University of Tokushima
Tokushima, Japan 770-8506
Email: simomura@is.tokushima-u.ac.jp

***Abstract*—A method of bias compensation in subspace identification method is proposed. The noise is assumed to be colored with 0 mean and is assumed to be uncorrelated with the input. The covariance matrix of the noise is estimated directly from the residuals instead of estimating the noise model. The proposing method becomes an iterative algorithm but it converges with order 2.**

I. Introduction

State space model identification has attracted attention for nearly two decades and has been applied for many practical situations. Many subspace identification methods such as MOESP[1], [2], N4SID[3], CCA[4], ORT[5], etc. reduces the influence of the noise by using instrumental variables. By using the fact that the past input or output are uncorrelated with the future noise, the data matrix is constructed as a combination of the comparatively future input/output data and the comparatively past input or input/output data in order to obtain an unbiased estimate. However, the required persistently exciting (PE) condition will be more difficult to achieve if the number of the rows of the data matrix becomes larger. Even when the PE condition is satisfied, a large number of rows will result in a large conditioning number of the data matrix, and will affect the precision of the estimate.

When the past input is adopted as an instrumental variable, a very tall data matrix will be required in order to project the influence of the noise onto the complement subspace of the range of the input. When the past input/output is adopted as an instrumental variable, not only the plant model but also the noise model must be estimated. This results in an increase of the degree of the estimated model and will cause the difficulty of the determination of the degree. Furthermore, a model reduction will be required in order to obtain the plant model.

In this paper, a new identification method is proposed in which an idea of bias compensation[6] is applied for subspace identification method. The noise covariance which causes an asymptotic bias of the estimate will be estimated from the residuals and will be subtracted from the data matrix. This results in a decrease of the size of the data matrix and the relaxation of the PE condition. The input dependent part of the output is estimated by using a plant model in the calculation of estimate of the noise covariance. Thus, the proposing method becomes an iterative algorithm. It is shown in [7] that possible divergence of the iteration-type bias-compensation algorithm may occur in the case of high noise. Furthermore,

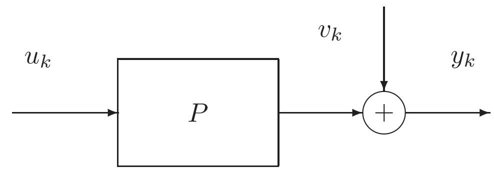

Fig. 1. System to be identified

fast convergence will be desirable in order to reduce the computational cost. Thus, the convergence properties must be analysed for iterative algorithms.

The identification problem will be formulated in section 2, and the subspace identification method is briefly summarized in section 3. Section 4 analyzes the bias of the estimate in the subspace method and proposes a bias compensation method. Stability of the proposed method is also analysed. Section 5 shows a numerical example to illustrate the proposed method and compare the proposed method with PI-MOESP and PO-MOESP. Finally, secton 6 concludes the paper.

Notation:
Colon notation[8] will be adopted. Namely, $X(i_1 : i_2, j_1 : j_2)$ denotes a submatrix consisting of the i_1-th row to i_2-th row and the j_1-th column to the j_2-th column of X. A colon by itself denotes an entire row or column.
$X^\dagger$ denotes a pseudo inverse (Moore-Penrose generalized inverse) [8] of X.

II. Problem Formulation

Consider the following system represented by a state space model:

$$x_{k+1} = Ax_k + Bu_k, \quad (1)$$
$$y_k = Cx_k + Du_k + v_k, \quad (2)$$

where $x_k \in \mathrm{R}^n$, $u_k \in \mathrm{R}^m$, $y_k \in \mathrm{R}^l$, and $v_k \in \mathrm{R}^l$ are the state, the input, the output, and the noise, respectively (Fig.1). $A \in \mathrm{R}^{n\times n}$, $B \in \mathrm{R}^{n\times m}$, $C \in \mathrm{R}^{l\times n}$, and $D \in \mathrm{R}^{l\times m}$ are the system matrices. Let P denote a matrix composed of the system matrices as $P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$.

The following assumptions are made for this system.

(A1) (A, B) is reachable and (A, C) is observable.
(A2) The noise $\{v_k\}$ is a colored noise with 0 mean.

(A3) $\{u_k\}$ is uncorrelated with $\{v_k\}$.

Identification Problem: Estimate the system matrix P from the input and output data $\{u_k, y_k\}$ $(k = 0, \ldots, N+s-2)$ where s is an integer greater than n.

III. Subspace Identification

In this section a subspace identification method is briefly summarized.

From eq. (1), the following matrix I/O equation is obtained for a given ingeter $s > n$:

$$\mathcal{Y}_{0|s-1} = \mathcal{O}_s \mathcal{X}_0 + \mathcal{T}_s \mathcal{U}_{0|s-1} + \mathcal{V}_{0|s-1}, \quad (3)$$

where

$$\mathcal{O}_s = \begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{s-1})^\top \end{bmatrix}^\top, \quad (4)$$

$$\mathcal{T}_s = \begin{bmatrix} D & & & 0 \\ CB & D & & \\ \vdots & & \ddots & \\ CA^{s-2}B & CA^{s-3}B & \cdots & D \end{bmatrix}, \quad (5)$$

$$\mathcal{X}_0 = \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}, \quad (6)$$

$$\mathcal{U}_{0|s-1} = \begin{bmatrix} u_0 & u_1 & \cdots & u_{N-1} \\ u_1 & u_2 & \cdots & u_N \\ \vdots & \vdots & & \vdots \\ u_{s-1} & u_s & \cdots & u_{N+s-2} \end{bmatrix}. \quad (7)$$

$\mathcal{Y}_{0|s-1}$ and $\mathcal{V}_{0|s-1}$ are defined in a silimar way as $\mathcal{U}_{0|s-1}$. The following PE condition is assumed.

(A4) $\operatorname{rank}\left(\begin{bmatrix} \mathcal{X}_0 \\ \mathcal{U}_{0|s-1} \end{bmatrix}\right) = n + ms.$

This assumption (A4) is a sufficient condition that $\mathcal{O}_s$ can be estimated by the following algorithm. It is also an sufficient condition for $\mathcal{U}_{0|s-1}$ to have full row rank.

Block Hankel matrices $\mathcal{U}_{0|s-1}$ and $\mathcal{Y}_{0|s-1}$ is LQ-decomposed as follows:

$$\begin{bmatrix} \mathcal{U}_{0|s-1} \\ \mathcal{Y}_{0|s-1} \end{bmatrix} = LQ^\top = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1^\top \\ Q_2^\top \end{bmatrix}. \quad (8)$$

From the fact that $\mathcal{U}_{0|s-1} = L_{11} Q_1^\top$ and that Q_1 is orthogonal to Q_2,

$$\mathcal{Y}_{0|s-1} \Pi^\perp_{\mathcal{U}_{0|s-1}} = L_{22} Q_2^\top \quad (9)$$

where $\Pi^\perp_{\mathcal{U}_{0|s-1}}$ is a projection matrix to the orthogonal complement of $\operatorname{range}(\mathcal{U}_{0|s-1})$:

$$\Pi^\perp_{\mathcal{U}_{0|s-1}} = I - \mathcal{U}_{0|s-1}^\top (\mathcal{U}_{0|s-1} \mathcal{U}_{0|s-1}^\top)^{-1} \mathcal{U}_{0|s-1}. \quad (10)$$

Thus,

$$L_{22} Q_2^\top = \mathcal{O}_s \mathcal{X}_0 \Pi^\perp_{\mathcal{U}_{0|s-1}} + \mathcal{V}_{0|s-1} \Pi^\perp_{\mathcal{U}_{0|s-1}}. \quad (11)$$

When $\mathcal{V}_{0|s-1} = 0$, it is true that

$$\operatorname{range}(L_{22}) = \operatorname{range}(\mathcal{O}_s). \quad (12)$$

Thus, (A, C) can be estimated from a singular value decomposition of L_{22}:

$$L_{22} = \begin{bmatrix} U_n & U_n^\perp \end{bmatrix} \begin{bmatrix} \Sigma_n & \\ & 0 \end{bmatrix} \begin{bmatrix} V_n^\top \\ (V_n^\perp)^\top \end{bmatrix} \quad (13)$$

$$= U_n \Sigma_n V_n^\top. \quad (14)$$

Because $\operatorname{rank} \mathcal{O}_s = n$, the degree of the system n can be estimated from the size of Σ_n. From eqs. (12) and (14), there exists a nonsingular matrix T such that $U_n = \mathcal{O}_s T$. Thus, $(A_T, C_T) = (T^{-1}AT, CT)$ can be estimated by

$$C_T = CT = U_n(1:l, :), \quad (15)$$

$$A_T = T^{-1}AT = \underline{U_s}^\dagger \overline{U_s}, \quad (16)$$

where

$$\underline{U_s} = U_n(1:(s-1)l, :), \quad \overline{U_s} = U_n(l+1:sl, :).$$

In order to estimate $(B_T, D) = (T^{-1}B, D)$, the following equation is obtained from eqs. (3), (8), (13), and (12):

$$(U_n^\perp)^\top \mathcal{T}_s = (U_n^\perp)^\top L_{21} L_{11}^{-1} - (U_n^\perp)^\top \mathcal{V}_{0|s-1} Q_1 L_{11}^{-1}. \quad (17)$$

The equation above is a linear function of (B_T, D). From the assumption (A3) and

$$\mathcal{V}_{0|s-1} Q_1 L_{11}^{-1} = (\mathcal{V}_{0|s-1} \mathcal{U}_{0|s-1}^\top)(\mathcal{U}_{0|s-1} \mathcal{U}_{0|s-1}^\top)^{-1},$$

the expectation of the second term of the r.h.s. of eq. (17) is 0. Therefore, (B_T, D) can be estimated by

$$\begin{bmatrix} D \\ B_T \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \underline{U_s} \end{bmatrix}^\dagger \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_s \\ \alpha_2 & & \alpha_s & \\ \vdots & \iddots & & \\ \alpha_s & & & 0 \end{bmatrix}^\dagger \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix}, \quad (18)$$

where $\alpha_i \in \mathrm{R}^{(sl-n)\times l}$ and $\beta_j \in \mathrm{R}^{(sl-n)\times m}$ are the block submatrix of the following matrix:

$$(U_n^\perp)^\top = \begin{bmatrix} \alpha_1 & \cdots & \alpha_s \end{bmatrix} \quad (19)$$

$$(U_n^\perp)^\top L_{21} L_{11}^{-1} = \begin{bmatrix} \beta_1 & \cdots & \beta_s \end{bmatrix} \quad (20)$$

If the estimate of (A_T, C_T) is unbiased, bias terms of estimate of (B_T, D) will be negligible under the assumption (A3). In order to obtain an unbiased estimate of (A_T, C_T), instrumental variables are introduced in general. Instead, a bias compensation method is considered in the following section.

IV. Bias Compensation

Asymptotic bias of $L_{22} L_{22}^\top$ will be calculated when $\mathcal{V}_{0|s-1} \neq 0$. From the asuummption (A3),

$$\lim_{N\to\infty} \frac{1}{N} \mathcal{U}_{0|s-1} \mathcal{V}_{0|s-1}^\top = 0, \quad (21)$$

$$\lim_{N\to\infty} \frac{1}{N} \mathcal{X}_0 \mathcal{V}_{0|s-1}^\top = 0. \quad (22)$$

From eq. (11) and an idempotence of $\Pi^\perp_{\mathcal{U}_{0|s-1}}$,

$$L_{22} L_{22}^\top = \mathcal{O}_s \mathcal{X}_0 \Pi^\perp_{\mathcal{U}_{0|s-1}} \mathcal{X}_0^\top \mathcal{O}_s^\top + \mathcal{V}_{0|s-1} \Pi^\perp_{\mathcal{U}_{0|s-1}} \mathcal{V}_{0|s-1}^\top + \mathcal{O}_s \mathcal{X}_0 \Pi^\perp_{\mathcal{U}_{0|s-1}} \mathcal{V}_{0|s-1}^\top + \mathcal{V}_{0|s-1} \Pi^\perp_{\mathcal{U}_{0|s-1}} \mathcal{X}_0^\top \mathcal{O}_s^\top. \quad (23)$$

The limit of the second term of the r.h.s. of the equation above can be calculated by using eqs. (10) and (21) as:

$$\lim_{N\to\infty}\frac{1}{N}\mathcal{V}_{0|s-1}\Pi^{\perp}_{\mathcal{U}_{0|s-1}}\mathcal{V}_{0|s-1}^{\top}=\lim_{N\to\infty}\frac{1}{N}\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}$$
$$-\lim_{N\to\infty}\frac{1}{N}\left(\frac{1}{N}\mathcal{V}_{0|s-1}\mathcal{U}_{0|s-1}^{\top}\right)\left(\frac{1}{N}\mathcal{U}_{0|s-1}\mathcal{U}_{0|s-1}^{\top}\right)^{-1}$$
$$\times\left(\frac{1}{N}\mathcal{U}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\right)^{-1}$$
$$=\lim_{N\to\infty}\frac{1}{N}\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top} \tag{24}$$

The limit of the 3rd and 4th terms of the r.h.s. of eq. (23) can be calculated by using eqs. (10), (21), and (22) as:

$$\lim_{N\to\infty}\frac{1}{N}\mathcal{X}_0\Pi^{\perp}_{\mathcal{U}_{0|s-1}}\mathcal{V}_{0|s-1}^{\top}=\lim_{N\to\infty}\frac{1}{N}\mathcal{X}_0\mathcal{V}_{0|s-1}^{\top}$$
$$-\lim_{N\to\infty}\left(\frac{1}{N}\mathcal{X}_0\mathcal{U}_{0|s-1}^{\top}\right)\left(\frac{1}{N}\mathcal{U}_{0|s-1}\mathcal{U}_{0|s-1}^{\top}\right)^{-1}$$
$$\times\left(\frac{1}{N}\mathcal{U}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\right)=0 \tag{25}$$

Thus, the bias term can be obtained as follows:

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}L_{22}L_{22}^{\top} &= \lim_{N\to\infty}\frac{1}{N}\mathcal{O}_s\mathcal{X}_0\Pi^{\perp}_{\mathcal{U}_{0|s-1}}\mathcal{X}_0^{\top}\mathcal{O}_s^{\top}\\ &+ \lim_{N\to\infty}\frac{1}{N}\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}.\end{aligned} \tag{26}$$

When a plant model $\hat{P}=\left[\begin{array}{c|c}\hat{A}&\hat{B}\\\hline\hat{C}&\hat{D}\end{array}\right]$ is given, an estimate of the noise will be obtained by

$$\hat{x}_{k+1} = \hat{A}\hat{x}_k+\hat{B}u_k, \tag{27}$$
$$\hat{v}_k = y_k-\hat{C}\hat{x}_k-\hat{D}u_k. \tag{28}$$

Define $\hat{\mathcal{V}}_{0|s-1}$ like $\mathcal{U}_{0|s-1}$, then

$$L_{22}L_{22}^{\top}-\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}$$

is an estimate of $\mathcal{O}_s\mathcal{X}_0\Pi^{\perp}_{\mathcal{U}_{0|s-1}}\mathcal{X}_0^{\top}\mathcal{O}_s^{\top}$. Thus, (A_T, C_T) can be estimated from the SVD of $L_{22}L_{22}^{\top}-\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}$ instead of L_{22} in eq. (13).

In order to analyze the estimation error, define $\delta x_k=\hat{x}_k-x_k$, $\delta v_k=\hat{v}_k-v_k$, and $\delta P=\hat{P}-P$, then,

$$\begin{bmatrix}\delta x_{k+1}\\ \delta v_k\end{bmatrix}=\begin{bmatrix}A\\ C\end{bmatrix}\delta x_k-\delta P\begin{bmatrix}\hat{x}_k\\ u_k\end{bmatrix}. \tag{29}$$

Because δv_k is a linear combination of the deterministic signal u_k and $\hat{x}_k$, $\{\delta v_k\}$ is uncorrelated with $\{v_k\}$ and its magnitude is bounded from above by using an appropriate norm of u_k and δv_k and a corresponding norm of δP, *i.e.*, there exists $M_v<\infty$ such that

$$\|\delta v_k\|_{[j,N+j-1]}\le M_v\|\delta P\|\cdot\|u_k\|_{[j,N+j-1]}, \quad \text{for} \quad j=0,1,\ldots,s-1. \tag{30}$$

Decompose $\hat{\mathcal{V}}_{0|s-1}=\mathcal{V}_{0|s-1}+\delta\mathcal{V}_{0|s-1}$, then

$$\begin{aligned}\lim_{N\to\infty}\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top} &= \lim_{N\to\infty}\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\\ &+\lim_{N\to\infty}\delta\mathcal{V}_{0|s-1}\delta\mathcal{V}_{0|s-1}^{\top}.\end{aligned} \tag{31}$$

From this, the magnitude of the error is bounded by $\|\delta P\|^2$, *i.e.*, there exists $M<\infty$ such that

$$\|L_{22}L_{22}^{\top}-\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}-\mathcal{O}_s\mathcal{X}_0\Pi^{\perp}_{\mathcal{U}_{0|s-1}}\mathcal{X}_0^{\top}\mathcal{O}_s^{\top}\|\le M\|\delta P\|^2. \tag{32}$$

In [9], [10], continuity of the SVD is discussed. Namely, a particular SVD of an Hermitian matrix has the differentiability properties. This means there exist $0<M_\Sigma<\infty$, $0<M_U<\infty$ and SVD's:

$$L_{22}L_{22}^{\top}-\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top} = \hat{U}\hat{\Sigma}\hat{U}^{\top}, \tag{33}$$
$$L_{22}L_{22}^{\top}-\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top} = U\Sigma U^{\top} \tag{34}$$

such that

$$\begin{aligned}\|\hat{\Sigma}-\Sigma\| &\le M_\Sigma\|\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}-\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\|\\ &\quad+o(\|\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}-\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\|),\end{aligned} \tag{35}$$
$$\begin{aligned}\|\hat{U}-U\| &\le M_U\|\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}-\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\|\\ &\quad+o(\|\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}-\mathcal{V}_{0|s-1}\mathcal{V}_{0|s-1}^{\top}\|).\end{aligned} \tag{36}$$

Therefore, the magnitude of the error of the plant model estimated by using the SVD of $L_{22}L_{22}^{\top}-\hat{\mathcal{V}}_{0|s-1}\hat{\mathcal{V}}_{0|s-1}^{\top}$ is bounded from above by $\|\delta P\|^2$.

The discussions above suggests the following iteration algorithm:

1) Let $\hat{P}^{(0)}$ be a plant model and let $i=0$.
2) LQ decompose the data matrix as (8).
3) Calculate $\hat{\mathcal{V}}^{(i)}_{0|s-1}$ by using the plant model $\hat{P}^{(i)}$.
4) Calculate $\hat{C}^{(i+1)}$ and $\hat{A}^{(i+1)}$ based on the SVD:
$$L_{22}L_{22}^{\top}-\hat{\mathcal{V}}^{(i)}_{0|s-1}(\hat{\mathcal{V}}^{(i)}_{0|s-1})^{\top}=U_n\Sigma_nU_n^{\top} \tag{37}$$
and eqs. (15), and (16).
5) Calculate $\hat{B}^{(i+1)}$ and $\hat{D}^{(i+1)}$ based on eq. (18).
6) Let $\hat{P}^{(i+1)}=\left[\begin{array}{c|c}\hat{A}^{(i+1)}&\hat{B}^{(i+1)}\\\hline\hat{C}^{(i+1)}&\hat{D}^{(i+1)}\end{array}\right]$.
7) Let $i=i+1$ and go to step 3).

From the analysis above, the iteration algorithm achieves a second order convergence.

Remark 1: Upper bound constant M in eq. (32) depends on the magnitude of the noise. Therefore, the proposed algorithm may diverge when the S/N ratio is small and the initial estimation error is large. A stabilization of the proposed method in such cases is left to the future work.

Remark 2: In the analysis of the bias terms, the first order term of $\delta\mathcal{V}_{0|s-1}$ vanishes as $N\to\infty$. When N is finite, The first order term of $\delta\mathcal{V}_{0|s-1}$ remains and there appears the first order term of $\|\delta P\|$ in the r.h.s. of the upper bound (32). Therefore, the estimation error decreases exponentially when δP becomes small. However, exponential rate of convergence is considered fast when N is large.

V. Numerical Example

In order to illustrate the availability of the proposed method, a numerical example is introduced. Consider the following 3rd

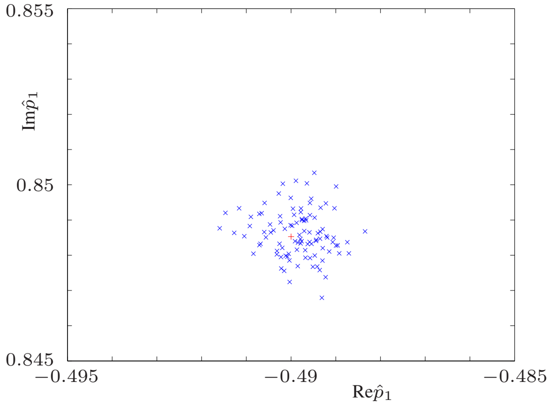

Fig. 2. Estimation result of the proposed method

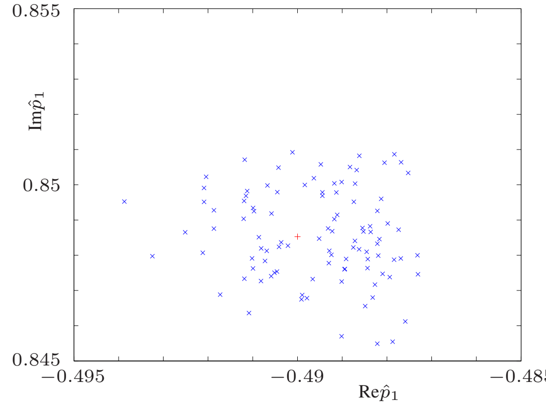

Fig. 3. Estimation result of PI-MOESP

order SISO system:

$$x_{k+1} = \begin{bmatrix} 0.98 & 2 & 0.74 \\ 0 & -0.49 & 1 \\ 0 & -0.72 & -0.49 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 0.85 \end{bmatrix} u_k \quad (38)$$

$$y_k = \begin{bmatrix} 0.57 & 0.72 & 0.27 \end{bmatrix} x_k + v_k. \quad (39)$$

where u_k is a zero mean white noise with covariance 1, v_k is a colored noise generated by

$$v_k = \frac{q^{-1} - 0.5q^{-2}}{1 + 0.5q^{-1} + 0.96q^{-2}} e_k \quad (40)$$

where e_k is a zero mean white noise with covariance 0.1^2. Hundred pairs of I/O data are prepared for estimation and the proposed method is compared with PI-MOESP and MO-MOESP. The estimations are performed with $N = 1001$ and $s = 20$, and the estimates of an eigen value of A, $p_1 = -0.49 + j0.8485$ are compared.

Fig.1 shows an estimation results of the proposed method, while Figs.2 and 3 are the results of PI-MOESP and PO-MOESP, respectively. The degree of the model is 3 for the proposed method and PI-MOESP, while 5 is used for PO-MOESP. All the three methods give unbiased estimate while the covariance of the proposed method is smaller than that of PI-MOESP. PO-MOESP gives the smallest covariance but a model reduction will be required for obtaining the plant model.

VI. Conclusion

Bias-compensated state space identification method is proposed. The noise is assumed to be a zero mean colored noise and is uncorrelated to the input. The bias compensation is based on the analysis of the asymptotic bias of the subspace identification method. The noise covariance is estimated from the plant model and the input signal. Therefore, the proposed method becomes an iterative algorithm. The estimation error is analysed and it is shown that the magnitude of the estimation error at each iteration is bounded from above by the squared norm of the previous model error. Thus, the proposed algorithm achieves a second order convergence.

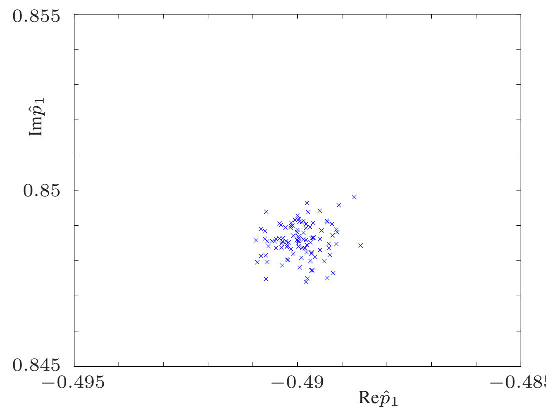

Fig. 4. Estimation result of PO-MOESP

References

[1] M. Verhaegen and P. Dewilde, "Subspace model identification Part 1. The output-error state-space model identification class of algorithms," *International Journal of Control*, vol. 56, pp. 1187–1210, 1992.

[2] ——, "Subspace model identification Part 2. Analysis of the elementary output-error state-space model identification algorithms," *International Journal of Control*, vol. 56, pp. 1211–1241, 1992.

[3] P. V. Overschee and B. D. Moore, "N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems," *Automatica*, vol. 30, pp. 75–93, 1994.

[4] H. Akaike, "Markovian Prepresentation of Stochastic Processed by Canonincal Variables," *SIAM Journal of Control*, vol. 13, no. 1, pp. 162–173, 1975.

[5] T. Katayama and G. Picci, "Realization of Stochastic systems with exogenous inputs and subspace identification methods," *Automatica*, vol. 35, pp. 1635–1652, 1999.

[6] S. Sagara and K. Wada, "On-line modified least-squares parameter estimation of linear discrete dynamic systems," *Int. J. Control*, vol. 25, pp. 329–343, 1977.

[7] T. Söderström, M. Hong, and W. X. Zheng, "Convergence properties of bias-eliminating algorithms for errors-in-variables identification," *Int. J. of Adaptive Control and Signal Processing*, vol. 19, pp. 703–722, 2005.

[8] G. H. Golub and C. F. van Loan, *Matrix Computations (2nd Ed.)*. The Johns Hopkins University Press, 1989.

[9] D. Bauer, M. Deistler, , and W. Scherrer, "Consistency and asymptotic normality of some subspace algorithms for systems without observed inputs," *Automatica*, vol. 35, pp. 1243–1254, 1999.

[10] D. Bauer, "Asymptotic properties of subspace estimators," *Automatica*, vol. 41, pp. 359–376, 2005.