A Study on Order Estimation by Subspace-based System Identification

○ Kunihiro OURA * , Kageo AKIZUKI* and Izumi HANAZAKI**

* School of Science and Engineering, Waseda University ,
55N-401B , 3-4-1 , Ohkubo , Shinjuku-ku , Tokyo , 169-8555 , JAPAN
Tel: +81-3-5286-3373 ; Fax: +81-3-3207-7143 ; E-mail: kuni@akizuki.elec.waseda.ac.jp

** College of Science and Engineering, Department of Systems and Engineering,
Tokyo Denki University , Ishizaka , Hatoyama, Hiki-gun , Saitama , 350-0394 , JAPAN
Tel: +81-492-96-2911 ; Fax: +81-492-96-6185 ; E-mail: hana@k.dendai.ac.jp

Abstract

In this paper, we consider the problem of order estimation by subspace-based system identification. There are many methods that can be used for order estimation, such as final prediction error (FPE) and Akaike's information criterion (AIC). But they are sometimes difficult to use because of their assumptions that the system is corrupted by Gaussian white noise, and they cannot be used for multi-variable cases for their computational complexity. From these backgrounds in consideration, we think of using subspace-based state-space system identification (4SID) for order estimation. Subspace methods are well known for its usefulness in identification of multi-variable systems, and it is said that the order of the system can be estimated by checking the singular values. We point out the difficulty in evaluation of the singular values when additive noise is effective, and propose a technique to estimate the order by introducing block matrices. Numerical examples for several plants will show usefulness of our method.

1 Introduction

In identification experiments, selection of model order is very important to have successful results. There are many estimation methods, but most of them based on the model of transfer function form, whereas there is state space representation form which is used for subspace methods.

Subspace methods are well-known for their characteristics that, (1) they need no assumptions on the model structure, (2) they are easily extended to the multivariable systems, (3) they use numerically stable algorithm. It is said that, in subspace methods, the order of the model is easily selected by using singular value

decomposition (SVD). Therefore in this paper, using the order estimation technique by subspace-based system identification, we select the transfer function model structure. In the first part, we use state-space model for observed input-output, and estimate the order of state-space model. Then we transform the state-space model to transfer function form. So that our final purpose of this paper is to develop an estimation method of transfer function model by subspace methods. There are some works on the subject (e.g. [1]), but in previous work, orders were estimated by GRAB (Gramian based) rank test. But we use many data for estimation, considering the properties of singular values.

2 Problem statement

Let us assume that a set of input-output data are generated by the plant described by (1).

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) + \mathbf{v}(k) \end{cases} \quad (1)$$

where $\mathbf{x}(k)$ is an n_s -dimensional state vector, $\mathbf{u}(k)$ is an m_s -dimensional input vector and $\mathbf{y}(k)$ is an l_s -dimensional output vector. $\mathbf{v}(k)$ is assumed to be a stationary random process uncorrelated with $\mathbf{u}(k)$. It is guaranteed that the system is controllable and observable.

It is easy to show that this model can be transformed to its transfer function form (2) in SISO case, where polynomials $A(z^{-1}), B(z^{-1})$ are given by (3).

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})} u(k) + v(k) \quad (2)$$

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n} \\ B(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n} \end{cases} \quad (3)$$

Then in this paper, we assume that the plant is depicted by fig.1, where $v(k)$ is not assumed to be white noise, and the number of data is denoted by N .

Figure 1: The plant model of this paper

The purpose of the paper is to consider the estimation method of order n (not n_s) from input-output data $\{u(k), y(k) : k = 1 \cdots N\}$ by using state space model.

3 Order estimation in subspace-methods

Let us review how the order can be selected in subspace methods. We think of two cases, (1) noise free output and (2) noisy output, respectively.

3.1 Estimation method when output is noise-free

When the output is noise-free, that is $v(k) = 0$, state-space model (1) is given by (4).

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases} \quad (4)$$

So that if we define the Hankel matrix as follows,

$$Y = \begin{bmatrix} y(k) & \cdots & y(k+N-1) \\ y(k+1) & \cdots & y(k+N) \\ \vdots & \ddots & \vdots \\ y(k+i-1) & \cdots & y(k+N+i-2) \end{bmatrix} \quad (5)$$

$$U = \begin{bmatrix} u(k) & \cdots & u(k+N-1) \\ u(k+1) & \cdots & u(k+N) \\ \vdots & \ddots & \vdots \\ u(k+i-1) & \cdots & u(k+N+i-2) \end{bmatrix} \quad (6)$$

and if we use the following notation,

$$X = \begin{bmatrix} x(k) \\ x(k+1) \\ \vdots \\ x(k+N-1) \end{bmatrix}^T, \Gamma_i = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix} \quad (7)$$

$$H_i = \begin{bmatrix} D & 0 \\ CB & D \\ \vdots & \ddots & D \\ CA^{i-1}B & \cdots & CB & D \end{bmatrix} \quad (8)$$

we have (9) for the matrix representation of (4).

$$Y = \Gamma_i X + H_i U \quad (9)$$

The RQ-factorization of the data-matrix pair (10) retrieve matrix with column space equal to the column space of extended observability matrix Γ_i , that is R_{22} spans the column space equal to Γ_i .

$$\begin{bmatrix} U \\ Y \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \quad (10)$$

So that by computing the SVD of the matrix R_{22} as (11), the order of state space model is given by (12).

$$R_{22} = [U_n \quad U_n^\perp] \begin{bmatrix} S_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_n^T \\ (V_n^\perp)^T \end{bmatrix} \quad (11)$$

$$n_s = \text{rank}(S_n) \quad (12)$$

3.2 Estimation method when output is noisy

In the case that there are additive noise to the output, we have (13) instead of (4) for state space model,

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) + v(k) \end{cases} \quad (13)$$

and we have (14) instead of (9) for the matrix representation of (13), where V is independent from U that is (15).

$$Y = \Gamma_i X + H_i U + V \quad (14)$$

$$p \lim UV^T = 0 \quad (15)$$

Again, by RQ-factorization of the data-matrix pair (16), we have matrices $R_{11}, \tilde{R}_{21}, \tilde{R}_{22}$.

$$\begin{bmatrix} U \\ Y \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ \tilde{R}_{21} & \tilde{R}_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \quad (16)$$

If N is large, $\tilde{R}_{22}$ turns to (17), so that if $v(k)$ is white noise, we can use the same procedure, that is SVD of $\tilde{R}_{22}$.

$$\frac{1}{N} \tilde{R}_{22} \tilde{R}_{22}^T \rightarrow \frac{1}{N} R_{22} R_{22}^T + R_{vv} \quad (17)$$

But if $v(k)$ is colored noise, we should introduce the instrumental variable (IV) matrix Φ which satisfies (18).

$$p \lim \frac{1}{N} \Phi V^T = 0 \quad (18)$$

In this case, RQ-factorization turns to (19), and $R_{22} R_{32}^T$ spans the column space equal to Γ_i .

$$\begin{bmatrix} U \\ Y \\ \Phi \end{bmatrix} = \begin{bmatrix} R_{11} & & 0 \\ R_{21} & R_{22} & \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} \quad (19)$$

Then the order of the state space model can be estimated by SVD of $R_{22}R_{32}^T$ as (20).

$$R_{22}R_{32}^T = \begin{bmatrix} U_n & U_n^\perp \end{bmatrix} \begin{bmatrix} S_n & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} V_n^T \\ (V_n^\perp)^T \end{bmatrix} \quad (20)$$

Here, we should discriminate a difference between S_n and S_2 . It depends on the size of singular values, and it can be done by visual inspection. After then, we have the order as (21), which is just the same as (12).

$$n_s = \text{rank}(S_n) \quad (21)$$

3.3 Transformation to transfer function form

After we have estimated the order of state space model, we must transform the model to transfer function form. As we know, state space model of order n_s can be transformed to (22) (in SISO case, to (23)).

$$y(k) = \{C_T(zI - A_T)^{-1}B_T + D\}u(k) + v(k) \quad (22)$$

$$y(k) = \{c_T(zI - A_T)^{-1}b_T + d\}u(k) + v(k) \quad (23)$$

In (22) and (23), matrices A_T, B_T, C_T, D are computed from the following set of equations [2].

$$U_n^{(1)}A_T = U_n^{(2)} \quad (24)$$

$$C_T = U_n(1:l_s,:) \quad (25)$$

$$\begin{bmatrix} \Xi(:,1:m_s) \\ \vdots \\ \Xi(:,(i-1)m_s+1:im_s) \end{bmatrix} = K \begin{bmatrix} D \\ B_T \end{bmatrix} \quad (26)$$

where,

$$U_n = \begin{bmatrix} U_n^{(1)} \\ U_n^{(2)} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_i \end{bmatrix} \quad (27)$$

$$\Xi = (U_n^{(1)})^T R_{21}R_{11}^{-1} \quad (28)$$

$$(U_n^\perp)^T = [\psi_1 \ \psi_2 \ \cdots \ \psi_i] \quad (29)$$

$$K = \begin{bmatrix} \psi_1 & \psi_2 & \cdots & \psi_i \\ \psi_2 & \psi_3 & \ddots & \\ \vdots & \ddots & & \\ \psi_i & & & 0 \end{bmatrix} \begin{bmatrix} I_{l_s} & 0 \\ 0 & U_n^{(1)} \end{bmatrix} \quad (30)$$

So that order of the transfer function form n can be estimated by that of the state space model. Especially, in SISO case we have $n = n_s$.

4 Order estimation using block matrices

4.1 Introduction of block matrices

As we see in the preceding section, the order of the model can be estimated by subspace methods. However, if the additive noise $v(k)$ has much effects on the system, discrimination between S_n and S_2 by visual inspection becomes very difficult. It is difficult to estimate the order from singular values. Then we should use a lot of data for SVD, in order to discriminate them. Because singular values of the plant are nearly in proportion to the data length (number of data), whereas other singular values are not related to it.

Therefore, if it is evaluated that noise is effective to the plant, we should use large N to estimate the order of the plant. But when N is large, it is impossible to apply subspace methods due to heavy computational load, so that we propose a technique to apply them by using block matrices.

First, let us divide Hankel matrices of size $(i \times N)$ into k block matrices of size $(i \times M)$, as (31).

$$Y = \begin{bmatrix} y(i+1) & y(i+2) & \cdots & y(i+N) \\ y(i+2) & y(i+3) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ y(2i) & \cdots & \cdots & y(2i+N-1) \end{bmatrix} \\ = \begin{bmatrix} y(1+i) & \cdots & y(i+M) \\ \vdots & \cdots & \vdots \\ y(2i) & \cdots & y(2i+M-1) \end{bmatrix} \\ \cdots \begin{bmatrix} y(i+(k-1)M+1) & \cdots & y(i+N) \\ \vdots & \cdots & \vdots \\ y(2i+(k-1)M) & \cdots & y(2i+N-1) \end{bmatrix} \quad (31)$$

If we use past inputs for instrumental variable matrix in (19), RQ-factorization is given as (32).

$$\begin{bmatrix} U_{i+I,i,N} \\ Y_{i+I,i,N} \\ U_{I,i,N} \end{bmatrix} = \begin{bmatrix} R_{11} & & 0 \\ R_{21} & R_{22} & \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} \quad (32)$$

where we use the notation as (33).

$$Y_{i+I,i,N} = \begin{bmatrix} y(i+1) & y(i+2) & \cdots & y(i+N) \\ y(i+2) & y(i+3) & \cdots & y(i+N+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(2i) & y(2i+1) & \cdots & y(2i+N-1) \end{bmatrix} \quad (33)$$

Then, by introducing the block matrices for each Hankel matrix as (34), RQ-factorization for each block matrices are defined, as (35).

$$\begin{aligned}
U_{i+I,i,N} &= [U_{i+I,i,M}, U_{M+i+I,i,M}, \dots, U_{(k-I)M+i+I,i,M}] \\
Y_{i+I,i,N} &= [Y_{i+I,i,M}, Y_{M+i+I,i,M}, \dots, Y_{(k-I)M+i+I,i,M}] \\
U_{I,i,N} &= [U_{I,i,M}, U_{M+i+I,i,M}, \dots, U_{(k-I)M+i+I,i,M}] \quad (34)
\end{aligned}$$

$$\begin{bmatrix} U_{(j-I)M+i+I,i,M} \\ Y_{(j-I)M+i+I,i,M} \\ U_{(j-I)M+i+I,i,M} \end{bmatrix} = \begin{bmatrix} R_{11}^j & 0 \\ R_{21}^j & R_{22}^j \\ R_{31}^j & R_{32}^j & R_{33}^j \end{bmatrix} \begin{bmatrix} Q_1^j \\ Q_2^j \\ Q_3^j \end{bmatrix} \quad (j = 1, \dots, k) \quad (35)$$

So that we have to calculate $R_{22}R_{32}^T$ from matrices $(R_{11}^1, R_{21}^1, \dots, R_{31}^1), \dots, (R_{11}^k, R_{21}^k, \dots, R_{31}^k)$ that are given by sets of block matrices.

4.2 Derivation of an equation for $R_{22}R_{32}^T$

We are going to derive an equation, using block matrices. From RQ-factorization of (32), we have (36).

$$\begin{aligned}
Y_{i+I,i,N}U_{i+I,i,N}^T &= R_{21}R_{11}^T \\
Y_{i+I,i,N}U_{I,i,N}^T &= R_{21}R_{31}^T + R_{22}R_{32}^T \quad (36) \\
U_{i+I,i,N}U_{I,i,N}^T &= R_{11}R_{31}^T
\end{aligned}$$

As R_{11} is full-rank matrix, it is invertible. Then we have (37) from the 1st and 3rd equations of (36).

$$R_{21}R_{31}^T = Y_{i+I,i,N}U_{i+I,i,N}^T(R_{11}R_{11}^T)^{-1}U_{i+I,i,N}U_{I,i,N}^T \quad (37)$$

It is clear from (32) that $R_{11}R_{11}^T = U_{i+I,i,N}U_{i+I,i,N}^T$, then

$$R_{21}R_{31}^T = Y_{i+I,i,N}U_{i+I,i,N}^T(U_{i+I,i,N}U_{i+I,i,N}^T)^{-1}U_{i+I,i,N}U_{I,i,N}^T \quad (38)$$

So that we have an equation (39) from (36) and (38).

$$\begin{aligned}
R_{22}R_{32}^T &= Y_{i+I,i,N}U_{I,i,N}^T \\
&\quad - Y_{i+I,i,N}U_{i+I,i,N}^T(U_{i+I,i,N}U_{i+I,i,N}^T)^{-1}U_{i+I,i,N}U_{I,i,N}^T \quad (39)
\end{aligned}$$

If we introduce a block matrices of the form (34), the first term of (39) is clearly given by (40).

$$\begin{aligned}
Y_{i+I,i,N}U_{I,i,N}^T &= Y_{i+I,i,M}U_{I,i,M}^T + Y_{M+i+I,i,M}U_{M+i+I,i,M}^T \\
&\quad + \dots + Y_{(k-I)M+i+I,i,M}U_{(k-I)M+i+I,i,M}^T \quad (40)
\end{aligned}$$

It is easy to show that the first term of (40) can be given by (41), considering the RQ-factorization (35).

$$Y_{i+I,i,M}U_{I,i,M}^T = R_{21}^1R_{31}^{1T} + R_{22}^1R_{32}^{1T} \quad (41)$$

By using the same notation as (41), (40) becomes (42).

$$\begin{aligned}
Y_{i+I,i,N}U_{I,i,N}^T &= (R_{21}^1R_{31}^{1T} + R_{22}^1R_{32}^{1T}) + \\
&\quad \dots + (R_{21}^kR_{31}^{kT} + R_{22}^kR_{32}^{kT}) \quad (42)
\end{aligned}$$

Similarly, we have (43) for the second term of (40).

$$\begin{aligned}
Y_{i+I,i,N}U_{i+I,i,N}^T &= R_{21}^1R_{11}^{1T} + R_{21}^2R_{11}^{2T} + \dots + R_{21}^kR_{11}^{kT} \\
U_{i+I,i,N}U_{I,i,N}^T &= R_{11}^1R_{31}^{1T} + R_{11}^2R_{31}^{2T} + \dots + R_{11}^kR_{31}^{kT} \\
U_{i+I,i,N}U_{i+I,i,N}^T &= R_{11}^1R_{11}^{1T} + R_{11}^2R_{11}^{2T} + \dots + R_{11}^kR_{11}^{kT} \quad (43)
\end{aligned}$$

By substitution of (42) and (43) for (39), we have (44).

$$\begin{aligned}
R_{22}R_{32}^T &= \left\{ \sum_{j=1}^k (R_{21}^jR_{31}^{jT} + R_{22}^jR_{32}^{jT}) \right\} \\
&\quad - \left(\sum_{j=1}^k R_{21}^jR_{11}^{jT} \right) \times \left(\sum_{j=1}^k R_{11}^jR_{11}^{jT} \right)^{-1} \times \left(\sum_{j=1}^k R_{11}^jR_{31}^{jT} \right) \quad (44)
\end{aligned}$$

Then we can calculate $R_{22}R_{32}^T$ from block matrices.

5 Numerical example

Several numerical examples help us to assure usefulness of the proposed technique. The model of the plant is given, in transfer function form as (45), where the input $u(k)$ is white noise with variance $\sigma_u^2 = 1.0$. We only treat the SISO case in this section, but extension to MIMO case is straightforward.

$$\begin{aligned}
A(z^{-1}) &= 1 + z^{-1} + 0.8273z^{-2} \\
&\quad + 0.9624z^{-3} + 0.5814z^{-4} \\
B(z^{-1}) &= 1 - 0.68z^{-1} + 0.18z^{-2} + 0.16z^{-3} \quad (45)
\end{aligned}$$

(1) Case of noise-free

Singular values for the data $N = 1000$ are depicted in fig2.

Figure2: Singular values

We can estimate easily, that the order $n = 4$.

(2) Case of Additive white noise

When white noise with variance $\sigma^2 = 2.0$ is added to the plant, singular values for the data $N = 1000$ are depicted in fig3.

Figure3: Singular values

It is hard to estimate the order from visual inspection. Then we use block matrices to calculate singular values with many data $N = 30000$. Results are depicted in fig4.

Figure4: Singular values

From fig4, we can estimate the order as $n = 4$, which shows usefulness of our technique.

(3) Case of Additive colored noise

When colored noise (46) with variance $\sigma_w^2 = 0.5$ is added to the plant, singular values for the data $N = 1000$ are depicted in fig5.

$$v(k) = \frac{1 + 0.5z^{-1}}{1 - z^{-1} + 0.5z^{-2}} w(k) \quad (46)$$

Figure5: Singular values

Again, it is hard to estimate the order from fig5. Then we use block matrices to calculate singular values with many data $N = 30000$. Results are depicted in fig6.

Figure6: Singular values

From fig6, we can estimate the order as $n = 4$.

6 Conclusion

We considered an order estimation method by subspace-based system identification. When the additive noise was effective to the plant, it was difficult to discriminate the 'true' singular values from others in SVD. In order to make their difference clear, we proposed our technique which used block matrices. Some numerical examples showed usefulness of our method.

References

- [1] J.Sorelius, T.Soderstrom, P.Stoica and M.Cedervall, "Order estimation method for subspace-based system identification", *Proceedings of the IFAC Symposium on System Identification* (Kitakyu, Japan), pp.1131-1136, 1997
- [2] M.Verhaegen and P.Dewilde, "Subspace model identification Part.1: The output-error state-space model identification class of algorithms", *Int.J.Control*, Vol. 56, No.5, pp.1187-1210, 1992
- [3] M.Verhaegen and P.Dewilde, "Subspace model identification Part.2: Analysis of the elementary output-error state-space model identification algorithm", *Int.J.Control*, Vol. 56, No.5, pp.1211-1241, 1992
- [4] M.Verhaegen, "Subspace model identification Part.3: Analysis of the ordinary output-error state-space model identification algorithm", *Int.J.Control*, Vol. 58, No.3, pp.555-586, 1993
- [5] P.V.Overschee and B.D.Moor "Subspace Algorithms for the stochastic identification problem", *Automatica*, Vol.29, No.3, pp.649-660, 1993
- [6] P.V.Overschee and B.D.Moor "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems", *Automatica*, Vol.30, No.1, pp.75-93, 1994