Proceedings of the 14th
KACC, Octorber 1999

An Interpretation of QR Factorization in Subspace Identification

°Yoshinori TAKEI*, Jun IMAI*, and Kiyoshi WADA*

*Dept. of Electrical and Electronic Systems Engineering, Graduate School of Information Science and Electrical Engineering, Kyushu University
Hakozaki 6-10-1, Higashi-ku, Fukuoka 812-8581, JAPAN
Tel: +81-92-642-3958 Fax: +81-92-642-3939
E-mail:takei@dickie.ees.kyushu-u.ac.jp

Abstract

Subspace-based state space system identification (4SID) methods have been demonstrated to perform well in a number of applications, but the properties of these have not been fully analyzed or understood yet. For applying the methods, no assumptions on structure of realization are needed and any coordinate transformation is allowed for the estimates. This is one reason why many kinds of properties expected for identification procedures have not been clarified yet. We illustrate, by using Schur complement, an interpretation of the R matrix yielded by the QR factorization in the 4SID procedure. The results in this paper can be useful for analysis of properties of parameters obtained by 4SID methods.

1 Introduction

The 4SID methods have attracted much attention because of being essentially suitable for multivariable system identification. The methods directly realize system matrix of state space model from input-output data without intermediate input-output expression such as impulse response or difference equation. The methods are essentially characterized by determination of the extended observability matrix from input-output data by using QR factorization and singular value decomposition. Then we consider the QR factorization on the procedure of the 4SID methods[1],[2], and show an iterpretation using Schur complement[3]. We study for the case of noise-free, white noise and colored noise, respectively.

2 4SID methods

We consider a discrete time linear system represented by

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{aligned} \tag{1}$$

where x_k is an n dimensional state vector, u_k m dimensional input and y_k l dimensional output, respectivly. The system matrices A, B, C, and D have appropriate dimensions. Furthermore, it is assumed that the model is minimal that is, the system is completely reachable and observable.

Let $i > n$, $N \gg i$, and Hankel matrix $U_{k,i,N}$ of $\{u_k\}$ is defined by

$$U_{k,i,N} := \begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+N-1} \\ u_{k+1} & u_{k+2} & \cdots & u_{k+N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k+i-1} & u_{k+i} & \cdots & u_{k+N+i-2} \end{bmatrix}$$

and $Y_{k,i,N}$ is defined in a similar way.

Then, we have

$$Y_{k,i,N} = \Gamma_i X_{k,N} + H_i U_{k,i,N} \tag{2}$$

from equation (1) where

$$\begin{aligned} X_{k,N} &:= [x_k \; x_{k+1} \; \cdots \; x_{k+N-1}] \\ \Gamma_i &:= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix} \end{aligned}$$

$$H_i := \begin{bmatrix} D & & & O \\ CB & D & & \\ \vdots & \ddots & \ddots & \\ CA^{i-2}B & \cdots & CB & D \end{bmatrix}$$

where Γ_i is called the extended observability matrix. We will omit the subscripts for U and Y unless otherwise mentioned.

The 4SID methods directly realize the quadruple of system matrices $[A_T,\ B_T,\ C_T,\ D]$ of state space model from input-output data $\{u_k,\ y_k\}$. The system matrices $[A_T,\ B_T,\ C_T,\ D]$ are defined as

$$[A_T,\ B_T,\ C_T,\ D] = [TAT^{-1},\ TB,\ CT^{-1},\ D]$$

where T is a similarity transformation. Let the input u_k be such that the following condition is satisfied.

$$\operatorname{rank}\begin{bmatrix} U \\ X \end{bmatrix} = mi + n \tag{3}$$

A 4SID procedure can be described as follows.

[*Algorithm* A[1]]

step1 Compute QR factorization of a matrix $\left[U^T\ Y^T\right]^T$, as in

$$\begin{bmatrix} U \\ Y \end{bmatrix} = \begin{bmatrix} R_{11} & O \\ R_{21} & R_{22} \end{bmatrix}\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \tag{4}$$

where R_{11} and R_{22} are lower triangular, and $Q_1Q_1^T = I_{mi}$, $Q_2Q_2^T = I_{li}$, $Q_1Q_2^T = O$.

step2 Compute Singular value decomposition of R_{22} as given in (4) of step1, i.e.,

$$R_{22} = \left[E_n\ E_n^{\perp}\right]\begin{bmatrix} S_n & O \\ O & S_2 \end{bmatrix}\begin{bmatrix} F_n^T \\ (F_n^{\perp})^T \end{bmatrix} \tag{5}$$

Here the dimension of S_n is equal to the one of the system.

step3 Compute the C_T and A_T from E_n as given in (5) of step2

$$C_T = E_n\,(1{:}i,\ {:}) \tag{6}$$

$$E_n^{(1)}A_T = E_n^{(2)} \tag{7}$$

where $E_n\,(1{:}i,\ {:})$ denotes the first l rows of E_n, $E_n^{(1)}$ is the submatrix composed of the first $(i-1)l$ rows of the matrix E_n and $E_n^{(2)}$ is constructed by the last rows in a similar way.

step4 Solve the following equation for the B_T and D.

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_i \end{bmatrix} = \Psi\begin{bmatrix} D \\ B_T \end{bmatrix} \tag{8}$$

where ξ_j, ψ_j, and Ψ are defined by the following relations:

$$[\ \xi_1\ \ \xi_2\ \ \cdots\ \ \xi_i\ ] := (E_n^{\perp})^T R_{21} R_{11}^{-1}$$

$$[\ \psi_1\ \ \psi_2\ \ \cdots\ \ \psi_i\ ] := (E_n^{\perp})^T$$

$$\Psi := \begin{bmatrix} \psi_1 & \psi_2 & \cdots & \psi_i \\ \psi_2 & & \psi_i & \\ \vdots & \psi_i & & \\ \psi_i & & & O \end{bmatrix}\begin{bmatrix} I_l & O \\ O & E_n^{(1)} \end{bmatrix}$$

Here, the size of $\xi_j\ (1 \le j \le i)$ and $\psi_j\ (1 \le j \le i)$ is $(li-n) \times m$ and $(li-n) \times l$, respectively. ⌟

3 An interpreation of QR factorization

In this section, we illustrate a relation between the extended observability matrix and the R matrix yielded by QR factorization in the 4SID procedure, by using the Schur complement.

Definition of the Schur complement[3]

Suppose we partition $A \in \mathbb{R}^{n\times n}$

$$A = \begin{bmatrix} A_{11} & A_{12}^T \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is r-by-r. Asssume that A_{11} is nonsingular. The matrix $S = A_{22} - A_{21}A_{11}^{-1}A_{12}^T$ is called the Schur complement of A_{11} in A.

3.1 Noise-free case

First, for the following matrix constructed from input-output data

$$\begin{bmatrix} U \\ Y \end{bmatrix}[\ U^T\ \ Y^T\ ] = \begin{bmatrix} UU^T & UY^T \\ YU^T & YY^T \end{bmatrix} \tag{9}$$

The Schur complement S_1 of UU^T in (9) is represented by

$$\begin{aligned} S_1 &= Y(I - U^T(UU^T)^{-1}U)Y^T \\ &= Y\Pi_U^{\perp}Y^T \end{aligned} \tag{10}$$

where $\Pi_U = U^T(UU^T)^{-1}U$ and $\Pi_U^{\perp} = I - \Pi_U$. Here S_1 in (10) can be rewritten by

$$\begin{aligned} S_1 &= R_{22}Q_2^TQ_2R_{22}^T \\ &= R_{22}R_{22}^T \end{aligned} \tag{11}$$

using (2) and (4). From results of (10) and (11), we have

$$Y\Pi_U^{\perp}Y^T = R_{22}R_{22}^T \tag{12}$$

Therefore we see that $R_{22}R_{22}^T$ is yielded by computing $Y\Pi_U^{\perp}Y^T$.

3.2 White noise case

It is assumed that the output of the deterministic system is perturbed by the zero-mean white noise v_k of variance σ_v^2. Then the output equation reads

$$z_k = y_k + v_k \tag{13}$$

Here Hankel matrices Z and V of $\{z_k\}$ and $\{v_k\}$ are defined in a similar way as in U, respectively. From equation (13), we have

$$Z = Y + V \tag{14}$$

Hence, Y in the equation (9) is replaced by Z, (9) can be rewritten as

$$\begin{bmatrix} U \\ Z \end{bmatrix} \begin{bmatrix} U^T & Z^T \end{bmatrix} = \begin{bmatrix} UU^T & UZ^T \\ ZU^T & ZZ^T \end{bmatrix} \tag{15}$$

The Schur complement S_2 of UU^T in (15) is represented by

$$\begin{aligned} S_2 &= Z\Pi_U^{\perp}Z^T \\ &= (Y+V)\Pi_U^{\perp}(Y+V)^T. \end{aligned} \tag{16}$$

Here, using a matrix $\tilde{R}_{22}$ yielded by computing the QR factorization of a matrix consisted of U and Z in a similar way as in (4), the equation (16) can be rewritten by

$$S_2 = \tilde{R}_{22}\tilde{R}_{22}^T \tag{17}$$

Thus from results of (16) and (17) we have

$$\tilde{R}_{22}\tilde{R}_{22}^T = (Y+V)\Pi_U^{\perp}(Y+V)^T \tag{18}$$

Based on the assumption that the input u_k and the noise v_k are independent, the following condition is satisfied:

$$\lim_{N\to\infty} \frac{1}{N} VY^T = O \tag{19}$$

Consequently, we have

$$\frac{1}{N}\tilde{R}_{22}\tilde{R}_{22}^T \Rightarrow \frac{1}{N}Y\Pi_U^{\perp}Y^T + \sigma_v^2 I_{li} \tag{20}$$

where the notation A $\Rightarrow$ B means that A asymptotically converges to B if the number of data N tends to infinity.

Therefore the white noise case is asymptotically treated as the noise-free case in a similar way.

3.3 Colored noise case

We assume that the noise v_k defined in (13) is a colored noise. Using a matrix $\hat{R}_{22}$ yielded by computing the QR factorization of a matrix $[U^T\ Z^T]^T$ in a similar way as in the case of the white noise, we have the following as in (20).

$$\frac{1}{N}\hat{R}_{22}\hat{R}_{22}^T \Rightarrow \frac{1}{N}Y\Pi_U^{\perp}Y^T + R_{vv} \tag{21}$$

where R_{vv} is a covariance matrix of v_k. Since v_k is a colored noise, R_{22} is not an identity matrix multiplied by a scalar. Therefore we introduce an instrumental variable Φ, satisfying the following conditions:

$$\lim_{N\to\infty} \frac{1}{N}\Phi V^T = O, \tag{22}$$

$$\text{rank} \begin{bmatrix} U \\ \Phi \end{bmatrix} = mi + p \tag{23}$$

where the size of Φ is $p \times N$ and $p > n$.

A 4SID procedure using the instrumental variable Φ is described as in

[*Algorithm B*[2]]

<u>**step1**</u>$'$ Compute QR factorization of a matrix consisted of input-output data U, Z and an instrumental variable Φ, i.e.

$$\begin{bmatrix} U \\ Y \\ \Phi \end{bmatrix} = \begin{bmatrix} \hat{R}_{11} & & O \\ \hat{R}_{21} & \hat{R}_{22} & \\ \hat{R}_{31} & \hat{R}_{32} & \hat{R}_{33} \end{bmatrix} \begin{bmatrix} \hat{Q}_1 \\ \hat{Q}_2 \\ \hat{Q}_3 \end{bmatrix} \tag{24}$$

<u>**step2**</u>$'$ Compute Singular value decomposition of $\hat{R}_{22}\hat{R}_{32}^T$ as given in (24) of step1$'$, as in

$$\hat{R}_{22}\hat{R}_{32}^T = \begin{bmatrix} E_n & E_n^{\perp} \end{bmatrix} \begin{bmatrix} S_n & O \\ O & S_2 \end{bmatrix} \begin{bmatrix} F_n^T \\ (F_n^{\perp})^T \end{bmatrix} \tag{25}$$

<u>**step3**</u>$'$ Using E_n, $E_n^{\perp}$, $\hat{R}_{11}$, and $\hat{R}_{21}$ yielded in step1$'$ and step2$'$, compute the quadruple of system matrices $[A,\ B,\ C,\ D]$ in a similar way in *Algorithm A.* ⌟

We consider a matrix Z which ia a linear combination of U and Φ. This is represented by

$$\begin{aligned}\hat{Z} &:= L_1 U + L_2 \Phi \\ &= \begin{bmatrix} L_1 & L_2 \end{bmatrix} \begin{bmatrix} U \\ \Phi \end{bmatrix} = L\Omega \end{aligned} \tag{26}$$

where $L := [\ L_1\ L_2\ ]$, $\Omega := [\ U^T\ \Phi^T\ ]^T$. Since the condition (23) is satisfied, $\hat{L}$ which minimize $\|Z - \hat{Z}\|_F^2$ for Φ exists uniquely. Here $\|\cdot\|_F$ denotes the Frobenius norm. Then $\hat{L}$ is represented by

$$\hat{L} = Z\Omega^T(\Omega\Omega^T)^{-1}. \tag{27}$$

From the result of (27), we substitute $\hat{L}$ for L in (26) to obtain

$$\begin{aligned}\hat{Z} &= Z\Omega^T(\Omega\Omega^T)^{-1}\Omega \\ &= Z\Pi_\Omega\end{aligned} \tag{28}$$

where $\Pi_\Omega = \Omega^T(\Omega\Omega^T)^{-1}\Omega$. Here we consider a matrix $\begin{bmatrix} U^T & \hat{Z}^T \end{bmatrix}^T$. From the assumption, $\hat{Z}$ and V are not correlative. Therefore we have

$$\lim_{N\to\infty} \frac{1}{N} V \begin{bmatrix} U^T & \hat{Z}^T \end{bmatrix} = O, \tag{29}$$

Suppose the following matrix as in (9)

$$\begin{aligned}\begin{bmatrix} U \\ \hat{Z} \end{bmatrix} \begin{bmatrix} U^T & \hat{Z}^T \end{bmatrix} &= \begin{bmatrix} U \\ Z \end{bmatrix} \begin{bmatrix} U^T & \hat{Z}^T \end{bmatrix} \\ &= \begin{bmatrix} U \\ \hat{Z} \end{bmatrix} \begin{bmatrix} U^T & Z^T \end{bmatrix} \\ &= \begin{bmatrix} U \\ Z \end{bmatrix} \Pi_\Omega \begin{bmatrix} U^T & Z^T \end{bmatrix} \\ &= \begin{bmatrix} UU^T & UZ^T \\ ZU^T & Z\Pi_\Omega Z^T \end{bmatrix}\end{aligned} \tag{30}$$

where $U\Pi_\Omega = U$. Thus the Schur complement S_3 of UU^T in (30) is represented by

$$\begin{aligned}S_3 &= Z(\Pi_\Omega - U^T(UU^T)^{-1}U)Z^T \\ &= Z(\Pi_\Omega - \Pi_U)Z^T \\ &= \hat{Z}\Pi_U^\perp \hat{Z}^T.\end{aligned} \tag{31}$$

Here

$$\begin{aligned}\Pi_\Omega &= \Omega^T(\Omega\Omega^T)^{-1}\Omega \\ &= U^T(UU^T)^{-1}U \\ &+ \Pi_U^\perp \Phi^T(\Phi\Pi_U^\perp\Phi^T)^{-1}\Phi\Pi_U^\perp.\end{aligned} \tag{32}$$

Then we have

$$S_3 = Z\Pi_U^\perp \Phi^T(\Phi\Pi_U^\perp\Phi^T)^{-1}\Phi\Pi_U^\perp Z^T \tag{33}$$

From the result of (24), the equation (31) can be rewritten as

$$S_3 = \hat{R}_{22}\hat{R}_{32}^T\Lambda^{-1}\hat{R}_{32}\hat{R}_{22}^T \tag{34}$$

where $\Lambda = \hat{R}_{32}\hat{R}_{32}^T + \hat{R}_{33}\hat{R}_{33}^T$. S_3 converges for $N \to \infty$, asymptotically, that is,

$$\begin{aligned}\frac{1}{N^2}S_3 \Rightarrow \Gamma_i \frac{1}{N} X\Pi_U^\perp\Phi^T(\Phi\Pi_U^\perp\Phi^T)^{-1} \\ \times \frac{1}{N}\Phi\Pi_U^\perp X^T\Gamma_i^T\end{aligned} \tag{35}$$

4 Conclusion

In this paper, we have studied the interpretation of the R matrix yielded by QR factorization in the 4SID procedure. Using the Schur complement for the matrix consisted of input-output data, we have shown the other computation of the R matrix. The results can be useful for analysing properties of parameters obtained by 4SID methods.

References

[1] M. Verhaegen and Dewilde: Subspace Model Identification. Part1. the Output-Error State-Space Model Identification Class of Algorithms, *Int. J. Control*, vol56-5, pp1187-1210, 1992

[2] M. Verhaegen: Subspace Model Identification. Part3. analysis of the ordinary output-error state space model identification algorithm, *Int. J. Control*, vol58-3, pp555-586, 1993

[3] G. H. Golub and C. F. Van Loan: Matrix Computations, THE JOHNS HOPKINS UNIVERSITY PRESS, 1991