STOCHASTIC REALIZATION ON A FINITE INTERVAL VIA "LQ DECOMPOSITION" IN HILBERT SPACE

Hideyuki Tanaka***, Tohru Katayama***

* Department of Applied Mathematics and Physics
Graduate School of Informatics, Kyoto University
Kyoto 606-8501, Japan
e-mail: {htanaka, katayama}@amp.i.kyoto-u.ac.jp

Keywords: Stochastic realization; Non-stationary realization; LQ-decomposition; Hilbert space; Stochastic subspace identification.

Abstract

In this paper, we consider a stochastic realization problem with finite covariance data based on "LQ decomposition" in a Hilbert space, and re-derive a non-stationary finite-interval realization ([4, 5]). We develop a new algorithm of computing system matrices of the finite-interval realization by LQ decomposition, followed by the SVD of a certain block matrix. Also, a stochastic subspace identification based on a finite time-series data is briefly discussed.

1 Introduction

Stochastic realization problem is to find a set of Markov models that generate a given covariance matrices of a stationary random process [3, 1]. It is well known that stochastic realization theory is an underlying principle for stochastic subspace identification methods [7, 8], in which Van Overschee and De Moor developed a subspace algorithm based on the non-stationary Kalman filter.

Lindquist and Picci [4, 5] have analyzed state space identification algorithms in the light of geometric theory of stochastic realization. In fact, they have discussed the state space modeling of the time-series data by separating three different cases: **(i)** an infinite complete covariance sequence is available, **(ii)** a finite complete covariance data is available, and **(iii)** a finite string of time-series data is available; especially, for the second case, they have derived a non-stationary finite-interval realization of a stationary process.

Recently, in [6], we have re-derived a balanced stochastic realization of Desai *et al.* [2] based on "LQ decomposition" in a Hilbert space generated by a stationary second order process under the assumption **(i)**, and briefly discussed a subspace identification method. In this paper, along the line of [6], we consider a stochastic realization problem on a finite interval [4, 5], thereby extending the result of [6] to the case where **(ii)** finite covariance data are available; we derive a non-stationary finite-interval realization of a stationary process by using "LQ decomposition" in a Hilbert space. The result is useful for studying a subspace identification method that estimates system matrices that produce a positive covariance sequence.

Due to space limitation, proofs of theorems and lemmas are omitted.

2 Problem Statement

Consider a second-order stationary process $\{y_t,\ t = 0, \pm 1, \cdots\}$, where y_t is a p-dimensional non-deterministic process with mean zero and covariance matrices

$$\Lambda_k = \mathrm{E}\left(y_{t+k} y_t^T\right), \quad k = 0, \pm 1, \pm 2, \cdots \tag{1}$$

where a set of covariance matrices $\{\Lambda_k,\ k = 0, \pm 1, \cdots\}$ is a positive real sequence in the sense that $\sum_{i,j} u_i^T \Lambda_{i-j} u_j > 0,\ u_i \not\equiv 0$. We assume that there exists a finite dimensional realization for y, so that the covariance matrix has a decomposition $\Lambda_k = HF^{k-1}G,\ k = 1, 2, \cdots$, where (F, G, H) is a minimal realization with $F \in \mathbb{R}^{n \times n}$.

According to [4, 5], we define the tail matrix by

$$\boldsymbol{y}_t := \begin{bmatrix} y_t & y_{t+1} & y_{t+2} & \cdots \end{bmatrix} \in \mathbb{R}^{p \times \infty}.$$

We also define a vector space spanned by all finite linear combinations of $\{\boldsymbol{y}_t\}$ as

$$\mathcal{Y}^\infty := \left\{ \sum a_k^T \boldsymbol{y}_k \mid a_k \in \mathbb{R}^p,\ k = 0, \pm 1, \cdots \right\}.$$

For the elements $a^T \boldsymbol{y}_i$ and $b^T \boldsymbol{y}_j \in \mathcal{Y}^\infty$, we define an inner product by

$$\begin{aligned} \langle a^T \boldsymbol{y}_i, b^T \boldsymbol{y}_j \rangle_{\frac{I}{\infty}} &:= \lim_{\nu \to \infty} \frac{1}{\nu} \sum_{k=t_0}^{t_0+\nu-1} a^T \boldsymbol{y}_{k+i} \boldsymbol{y}_{k+j}^T b \\ &= a^T \Lambda_{i-j} b \end{aligned} \tag{2}$$

where the right hand side is independent of t_0, because y is stationary. By completing the vector space $\mathcal{Y}^\infty$ with respect to convergence in the norm induced by the inner product (2), we get a Hilbert space, which is also written as $\mathcal{Y}^\infty$.

Let $\mathcal{U}$ be a Hilbert subspace of $\mathcal{Y}^\infty$, and the orthogonal projection of $\boldsymbol{\eta} \in \mathcal{Y}^\infty$ onto the space $\mathcal{U}$ be denoted by $\hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{\eta} \mid \mathcal{U})$. Also, let the row space spanned by a matrix U be expressed as $\mathrm{span}(U)$. If $\langle U, U\rangle_{\frac{I}{\infty}}$ has an inverse, the orthogonal projection is written as

$$\begin{aligned}\hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{\eta} \mid U) &:= \hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{\eta} \mid \mathrm{span}(U)) \\ &= \langle \boldsymbol{\eta}, U\rangle_{\frac{I}{\infty}} \langle U, U\rangle^{-1}_{\frac{I}{\infty}} U. \qquad (3)\end{aligned}$$

We extend $\mathcal{Y}^\infty$ to $\mathcal{Y}^{\bullet\times\infty}$ so that matrices are included as its elements[1].

We assume that the data are generated by a linear system and described by

$$\begin{bmatrix} \boldsymbol{x}_{t+1} \\ \boldsymbol{y}_t \end{bmatrix} = \begin{bmatrix} F \\ H \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} \boldsymbol{w}_t \\ \boldsymbol{v}_t \end{bmatrix}$$

where $F \in \mathbb{R}^{n\times n}$ and $H \in \mathbb{R}^{p\times n}$ satisfy a decomposition $\Lambda_k = HF^{k-1}G$, $\boldsymbol{x}_t \in \mathcal{Y}^{n\times\infty}$ is a state matrix, and the elements of tail matrices, $\boldsymbol{w}_t \in \mathcal{Y}^{n\times\infty}$ and $\boldsymbol{v}_t \in \mathcal{Y}^{p\times\infty}$ are white noises satisfying

$$\left\langle \begin{bmatrix} \boldsymbol{w}_s \\ \boldsymbol{v}_s \end{bmatrix}, \begin{bmatrix} \boldsymbol{w}_t \\ \boldsymbol{v}_t \end{bmatrix} \right\rangle_{\frac{I}{\infty}} = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{st}$$

with $R > 0$.

Given finite data $\boldsymbol{y}_t \in \mathcal{Y}^{p\times\infty}$, $t = 0, 1, \cdots, 2\tau - 1$ with $\tau > n$, Lindquist and Picci [4, 5] have derived a finite-interval realization for $\boldsymbol{y}_t$, which is given by the following (transient) Kalman filter with zero initial conditions

$$\hat{\boldsymbol{x}}_{t+1} = F\hat{\boldsymbol{x}}_t + \hat{\Gamma}_t(\boldsymbol{y}_t - H\hat{\boldsymbol{x}}_t), \quad \hat{\boldsymbol{x}}_0 = 0$$

where $\hat{\boldsymbol{x}}_t \in \mathcal{Y}^{n\times\infty}$ is the estimation of the state matrix $\boldsymbol{x}_t \in \mathcal{Y}^{n\times\infty}$, $\hat{\Gamma}_t$ is the forward non-stationary Kalman gain.

By using the non-stationary forward Kalman filter, it has been shown that the tail matrices $\boldsymbol{y}_t$, $t = 0, 1, \cdots, \tau-1$, satisfy the following time-varying system

$$\begin{bmatrix} \hat{\boldsymbol{x}}_{t+1} \\ \boldsymbol{y}_t \end{bmatrix} = \begin{bmatrix} F & \hat{\Gamma}_t \\ H & I \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}}_t \\ \hat{\boldsymbol{v}}_t \end{bmatrix}, \quad \hat{\boldsymbol{x}}_0 = 0 \qquad (4)$$

where $\hat{\boldsymbol{v}}_t$ is the forward (transient) innovation process defined by $\hat{\boldsymbol{v}}_t := \boldsymbol{y}_t - C\hat{\boldsymbol{x}}_t$.

In this paper, we assume that a set of exact but finite covariance data $\{\Lambda_0, \Lambda_1, \Lambda_2, \cdots, \Lambda_{2\tau-1}\}$ is available with $\tau > n$; this is equivalent to the fact that a finite number of tail matrices $\boldsymbol{y}_t \in \mathcal{Y}^{p\times\infty}$, $t = 0, 1, \cdots, 2\tau - 1$ are given. Under this assumption, the problem is to give a finite-interval realization of $\boldsymbol{y}_t$ by "LQ decomposition" in a Hilbert space and provide a method of computing the system matrices F, H, $\hat{\Gamma}_t$ and $\hat{R}_t$ in (4) for $t = 0, 1, \cdots, \tau - 1$.

3 LQ Decomposition of Data Matrix

In this section, after providing some notations, we review a finite-interval realization derived from the CCA, and then compute the LQ decomposition of a given data matrix with the help of the finite-interval realization.

3.1 Covariance matrices

In terms of tail matrices $\boldsymbol{y}_t \in \mathcal{Y}^{p\times\infty}$, $t = 0, 1, \cdots, 2\tau - 1$, we define data matrices as

$$Y_t^- := \begin{bmatrix} \boldsymbol{y}_{t-1} \\ \boldsymbol{y}_{t-2} \\ \vdots \\ \boldsymbol{y}_1 \\ \boldsymbol{y}_0 \end{bmatrix}, \quad Y_t^+ := \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{y}_{t+1} \\ \vdots \\ \boldsymbol{y}_{2\tau-2} \\ \boldsymbol{y}_{2\tau-1} \end{bmatrix} \qquad (5)$$

for $t = 1, \cdots, 2\tau - 1$. For notational convenience, we define the reversed tail matrices by $\boldsymbol{\zeta}_{-s} := \boldsymbol{y}_{-s+2\tau-1}$ for $s = 0, 1, \cdots, 2\tau - 1$, and

$$Z_{-s}^- := \begin{bmatrix} \boldsymbol{\zeta}_{-s} \\ \boldsymbol{\zeta}_{-s-1} \\ \vdots \\ \boldsymbol{\zeta}_{-2\tau+2} \\ \boldsymbol{\zeta}_{-2\tau+1} \end{bmatrix}, \quad Z_{-s}^+ := \begin{bmatrix} \boldsymbol{\zeta}_{-s+1} \\ \boldsymbol{\zeta}_{-s+2} \\ \vdots \\ \boldsymbol{\zeta}_{-1} \\ \boldsymbol{\zeta}_0 \end{bmatrix} \qquad (6)$$

[1] We define $\mathcal{Y}^{p\times\infty}$ as

$$\mathcal{Y}^{p\times\infty} := \left\{ \begin{bmatrix} \boldsymbol{\eta}_1^T & \boldsymbol{\eta}_2^T & \cdots & \boldsymbol{\eta}_p^T \end{bmatrix}^T \middle| \, \boldsymbol{\eta}_k \in \mathcal{Y}^\infty \right\}.$$

For given $\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1^T & \cdots & \boldsymbol{\alpha}_p^T \end{bmatrix}^T \in \mathcal{Y}^{p\times\infty}$, we define the orthogonal projection of $\boldsymbol{\alpha}$ onto the space $\mathrm{span}(U)$ as

$$\hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{\alpha} \mid U) := \begin{bmatrix} \hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{\alpha}_1 \mid U) \\ \vdots \\ \hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{\alpha}_p \mid U) \end{bmatrix}.$$

It should be noted that a bilinear form $\langle \cdot, \cdot \rangle_\infty$ is described as

$$\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle_{\frac{I}{\infty}} := \begin{bmatrix} \langle \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1 \rangle_{\frac{I}{\infty}} & \cdots & \langle \boldsymbol{\alpha}_1, \boldsymbol{\beta}_q \rangle_{\frac{I}{\infty}} \\ \vdots & & \vdots \\ \langle \boldsymbol{\alpha}_p, \boldsymbol{\beta}_1 \rangle_{\frac{I}{\infty}} & \cdots & \langle \boldsymbol{\alpha}_p, \boldsymbol{\beta}_q \rangle_{\frac{I}{\infty}} \end{bmatrix}$$

for $\boldsymbol{\alpha} \in \mathcal{Y}^{p\times\infty}$ and $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^T & \cdots & \boldsymbol{\beta}_q^T \end{bmatrix}^T \in \mathcal{Y}^{q\times\infty}$, and the orthogonal projection $\boldsymbol{\eta} \in \mathcal{Y}^{\bullet\times\infty}$ onto $\mathrm{span}(U)$ is also calculated as in (3).

for $s = 1, \cdots, 2\tau - 1$. It may be noted that for $t = s = \tau$, all the data matrices have the same number of rows with $Y_\tau^- = Z_{-\tau}^-$ and $Y_\tau^+ = Z_{-\tau}^+$, where the former are termed the past data matrices, while the latter the future data matrices.

Moreover, we define covariance matrices

$$\begin{aligned}
\Phi_t &:= \langle Y_t^-, Y_t^- \rangle_{\frac{I}{\infty}} \\
&= \begin{bmatrix} \Lambda_0 & \Lambda_1 & \Lambda_2 & \cdots & \Lambda_{t-1} \\ \Lambda_1^T & \Lambda_0 & \Lambda_1 & \cdots & \Lambda_{t-2} \\ \Lambda_2^T & \Lambda_1^T & \Lambda_0 & \cdots & \Lambda_{t-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Lambda_{t-1}^T & \Lambda_{t-2}^T & \Lambda_{t-3}^T & \cdots & \Lambda_0 \end{bmatrix}, \quad (7) \\
\Psi_{-t} &:= \langle Z_{-t}^+, Z_{-t}^+ \rangle_{\frac{I}{\infty}} \\
&= \begin{bmatrix} \Lambda_0 & \Lambda_1^T & \Lambda_2^T & \cdots & \Lambda_{t-1}^T \\ \Lambda_1 & \Lambda_0 & \Lambda_1^T & \cdots & \Lambda_{t-2}^T \\ \Lambda_2 & \Lambda_1 & \Lambda_0 & \cdots & \Lambda_{t-3}^T \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Lambda_{t-1} & \Lambda_{t-2} & \Lambda_{t-3} & \cdots & \Lambda_0 \end{bmatrix}, \quad (8)
\end{aligned}$$

for $t = 1, \cdots, 2\tau$ and the block Hankel matrix

$$\begin{aligned}
\mathcal{H}_\tau &= \langle Y_\tau^+, Y_\tau^- \rangle_{\frac{I}{\infty}} \\
&= \begin{bmatrix} \Lambda_1 & \Lambda_2 & \Lambda_3 & \cdots & \Lambda_\tau \\ \Lambda_2 & \Lambda_3 & \Lambda_4 & \cdots & \Lambda_{\tau+1} \\ \Lambda_3 & \Lambda_4 & \Lambda_5 & \cdots & \Lambda_{\tau+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Lambda_\tau & \Lambda_{\tau+1} & \Lambda_{\tau+2} & \cdots & \Lambda_{2\tau-1} \end{bmatrix} \quad (9) \\
&= \langle Z_{-\tau}^+, Z_{-\tau}^- \rangle_{\frac{I}{\infty}}.
\end{aligned}$$

It should be noted that the covariance matrices of (7) and (8) are defined for $t = 1, \cdots, 2\tau$, but the block Hankel matrix (9), the covariance matrix of the future and the past, is defined for $\mathcal{H}_\tau$ only.

3.2 Canonical correlation analysis

As usual, we compute the canonical decomposition, or the weighted SVD, of the block Hankel matrix $\mathcal{H}_\tau$ as ([4, 5])

$$\begin{aligned}
\Psi_{-\tau}^{-\frac{1}{2}} \mathcal{H}_\tau \Phi_\tau^{-\frac{T}{2}} &= [\bar{U}_\tau \ \ \tilde{U}_\tau] \begin{bmatrix} \bar{\Sigma}_\tau & 0 \\ 0 & \tilde{\Sigma}_\tau \end{bmatrix} \begin{bmatrix} \bar{V}_\tau^T \\ \tilde{V}_\tau^T \end{bmatrix} \\
&= \bar{U}_\tau \bar{\Sigma}_\tau \bar{V}_\tau^T, \quad \bar{\Sigma}_\tau \in \mathbb{R}^{n\times n}
\end{aligned}$$

where $\operatorname{rank} \bar{\Sigma}_\tau = n$, and $\bar{U}_\tau^T \bar{U}_\tau = I_n$, $\bar{V}_\tau^T \bar{V}_\tau = I_n$. Hence, we get

$$\mathcal{H}_\tau = \Psi_{-\tau}^{\frac{1}{2}} \bar{U}_\tau \bar{\Sigma}_\tau \bar{V}_\tau^T \Phi_\tau^{\frac{T}{2}}.$$

It therefore follows that the extended observability matrix $\mathcal{O}_\tau$ and the extended reachability matrix $\mathcal{C}_\tau$ are respectively given by

$$\begin{aligned}
\mathcal{O}_\tau &:= \Psi_{-\tau}^{\frac{1}{2}} \bar{U}_\tau \bar{\Sigma}_\tau^{\frac{1}{2}}, \\
\mathcal{C}_\tau &:= \bar{\Sigma}_\tau^{\frac{1}{2}} \bar{V}_\tau^T \Phi_\tau^{\frac{T}{2}}
\end{aligned}$$

with $\operatorname{rank} \mathcal{O}_\tau = n$, $\operatorname{rank} \mathcal{C}_\tau = n$, and hence we have $\mathcal{H}_\tau = \mathcal{O}_\tau \mathcal{C}_\tau$.

From the assumption about the covariance data, there exist matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{p\times n}$ such that

$$\mathcal{O}_\tau = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\tau-1} \end{bmatrix}, \quad (10)$$

$$\mathcal{C}_\tau = \begin{bmatrix} B & AB & \cdots & A^{\tau-1}B \end{bmatrix} \quad (11)$$

where it should be noted that matrices A, B and C are dependent on τ.

Let ($\bar{A}$, $\bar{B}$, $\bar{C}$) be a stochastically balanced realization obtained by the infinite covariance data [2, 5], namely with $\tau \to \infty$. Then, it follows that (A, B, C) in (10) and (11) satisfies the relation $\bar{A} = Q_\tau^{-1} A Q_\tau$, $\bar{B} = Q_\tau^{-1} B$ and $\bar{C} = C Q_\tau$ where $Q_\tau \in \mathbb{R}^{n\times n}$ is a non-singular transform [5], so that we have $\Lambda_k = CA^{k-1}B$, $k = 1, 2, \cdots, 2\tau - 1$. The triplet (A, B, C) obtained above is a finite-interval stochastically balanced realization which is minimal and dependent on τ [5].

3.3 LQ decomposition in a Hilbert space

We describe a stochastic realization in terms of a (transient) innovation process [4, 5], and then provide an "LQ decomposition" of a data matrix in a Hilbert space.

Define the variables for $t = 1, \cdots, 2\tau - 1$

$$\hat{\boldsymbol{v}}_t := \boldsymbol{y}_t - \hat{\mathrm{E}}_{\frac{I}{\infty}}(\boldsymbol{y}_t \,|\, Y_t^-) \quad (12)$$

with the initial condition $\hat{\boldsymbol{v}}_0 := \boldsymbol{y}_0$.

Lemma 1 *The process $\hat{\boldsymbol{v}}_j$ defined by (12) is a white noise satisfying*

$$\langle \hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_j \rangle_{\frac{I}{\infty}} = \hat{R}_j \delta_{ij}, \quad i, j = 0, 1, \cdots, 2\tau - 1 \quad (13)$$

where $\hat{R}_j > 0$, $j = 0, 1, \cdots, 2\tau - 1$, and

$$\langle \boldsymbol{y}_i, \hat{\boldsymbol{v}}_j \rangle_{\frac{I}{\infty}} = 0, \quad 0 \le i < j \le 2\tau - 1. \quad (14)$$

Define $\hat{L}_{i,j}$ as

$$\hat{L}_{i,j} := \langle \boldsymbol{y}_i, \hat{\boldsymbol{v}}_j \rangle_{\frac{I}{\infty}} \hat{R}_j^{-1}, \quad 0 \le j \le i \le 2\tau - 1. \quad (15)$$

An explicit form of $\hat{L}_{i,j} \in \mathbb{R}^{p\times p}$ for $j \le i \le 2\tau - 1,\ 0 \le j \le \tau - 1$ is provided later in (24).

In terms of $\hat{L}_{i,j}$ of (15), we define

$$\hat{\mathcal{L}}_\tau^- := \begin{bmatrix} \hat{L}_{\tau-1,\tau-1} & \hat{L}_{\tau-1,\tau-2} & \cdots & \hat{L}_{\tau-1,0} \\ & \hat{L}_{\tau-2,\tau-2} & \cdots & \hat{L}_{\tau-2,0} \\ & & \ddots & \vdots \\ 0 & & & \hat{L}_{0,0} \end{bmatrix},$$

$$\hat{\mathcal{L}}_\tau^+ := \begin{bmatrix} \hat{L}_{\tau,\tau} & & & 0 \\ \hat{L}_{\tau+1,\tau} & \hat{L}_{\tau+1,\tau+1} & & \\ \vdots & \vdots & \ddots & \\ \hat{L}_{2\tau-1,\tau} & \hat{L}_{2\tau-1,\tau+1} & \cdots & \hat{L}_{2\tau-1,2\tau-1} \end{bmatrix},$$

$$\hat{\mathcal{S}}_\tau := \begin{bmatrix} \hat{L}_{\tau,\tau-1} & \hat{L}_{\tau,\tau-2} & \cdots & \hat{L}_{\tau,0} \\ \hat{L}_{\tau+1,\tau-1} & \hat{L}_{\tau+1,\tau-2} & \cdots & \hat{L}_{\tau+1,0} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{L}_{2\tau-1,\tau-1} & \hat{L}_{2\tau-1,\tau-2} & \cdots & \hat{L}_{2\tau-1,0} \end{bmatrix},$$

where $\hat{\mathcal{L}}_\tau^-$, $\hat{\mathcal{L}}_\tau^+$, $\hat{\mathcal{S}}_\tau \in \mathbb{R}^{\tau p\times\tau p}$. Moreover, we define

$$\hat{V}_t^- = \begin{bmatrix} \hat{\boldsymbol{v}}_{t-1} \\ \hat{\boldsymbol{v}}_{t-2} \\ \vdots \\ \hat{\boldsymbol{v}}_1 \\ \hat{\boldsymbol{v}}_0 \end{bmatrix}, \quad \hat{V}_t^+ = \begin{bmatrix} \hat{\boldsymbol{v}}_t \\ \hat{\boldsymbol{v}}_{t+1} \\ \vdots \\ \hat{\boldsymbol{v}}_{2\tau-2} \\ \hat{\boldsymbol{v}}_{2\tau-1} \end{bmatrix} \tag{16}$$

and covariance matrices:

$$\hat{\mathcal{R}}_\tau^- := \langle \hat{V}_\tau^-, \hat{V}_\tau^- \rangle_{\frac{I}{\infty}}, \quad \hat{\mathcal{R}}_\tau^+ := \langle \hat{V}_\tau^+, \hat{V}_\tau^+ \rangle_{\frac{I}{\infty}}.$$

Theorem 1 *The past Y_τ^- and the future Y_τ^+ of (5) are decomposed as*

$$\begin{bmatrix} Y_\tau^- \\ Y_\tau^+ \end{bmatrix} = \begin{bmatrix} \hat{\mathcal{L}}_\tau^- & 0 \\ \hat{\mathcal{S}}_\tau & \hat{\mathcal{L}}_\tau^+ \end{bmatrix} \begin{bmatrix} \hat{V}_\tau^- \\ \hat{V}_\tau^+ \end{bmatrix} \tag{17}$$

where $\hat{V}_\tau^-$ and $\hat{V}_\tau^+$ are given by (16) and satisfy

$$\left\langle \begin{bmatrix} \hat{V}_\tau^- \\ \hat{V}_\tau^+ \end{bmatrix}, \begin{bmatrix} \hat{V}_\tau^- \\ \hat{V}_\tau^+ \end{bmatrix} \right\rangle_{\frac{I}{\infty}} = \begin{bmatrix} \hat{\mathcal{R}}_\tau^- & 0 \\ 0 & \hat{\mathcal{R}}_\tau^+ \end{bmatrix}. \tag{18}$$

Moreover, the orthogonal projection of the future onto the past is written as

$$\hat{\mathrm{E}}_{\frac{I}{\infty}}(Y_\tau^+ \mid Y_\tau^-) = \hat{\mathcal{S}}_\tau \hat{V}_\tau^-.$$

It can be shown that the decomposition of (17) is performed by an "LQ decomposition" in the Hilbert space.

Now we evaluate the terms $\hat{\mathcal{L}}_\tau^-$ and $\hat{\mathcal{S}}_\tau$ in (17) [2]. To this end, we define [8]

$$\hat{P}_t := \mathcal{C}_t \Phi_t^{-1} \mathcal{C}_t^T, \quad t = 1, \cdots, \tau \tag{19}$$

[2]The matrix $\hat{\mathcal{L}}_\tau^+$ in (17) is irrelevant for the latter development.

with $\hat{P}_0 := 0$, where it should be noted that $\mathcal{C}_t$ in (19) is a truncated extended reachability matrix defined as

$$\mathcal{C}_t = \begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix}, \quad t = 1, \cdots, \tau$$

by using A and B [see (11)].

Proposition 1 *([8]) The matrix $\hat{P}_t$ satisfies the following discrete-time Riccati equation with $\hat{P}_0 = 0$*

$$\hat{P}_{t+1} = A\hat{P}_t A^T + (B - A\hat{P}_t C^T)(\Lambda_0 - C\hat{P}_t C^T)^{-1}(B - A\hat{P}_t C^T)^T$$

for $t = 0, 1, 2, \cdots, \tau - 1$.

In terms of the solution $\hat{P}_t$ of Riccati equation, we define matrices

$$\hat{R}_t := \Lambda_0 - C\hat{P}_t C^T, \tag{20}$$

$$\hat{K}_t := (B - A\hat{P}_t C^T)(\Lambda_0 - C\hat{P}_t C^T)^{-1}, \tag{21}$$

for $t = 0, 1, \cdots, \tau - 1$. Also, define ([8])

$$\hat{\boldsymbol{x}}_t := \mathcal{C}_t \Phi_t^{-1} Y_t^-. \tag{22}$$

Then, we can prove the following lemma.

Proposition 2 *([4, 5]) The tail matrix $\boldsymbol{y}_t \in \mathcal{Y}^{p\times\infty}$ $(t = 0, 1, \cdots, \tau - 1)$ is realized by the following time-varying system*

$$\begin{bmatrix} \hat{\boldsymbol{x}}_{t+1} \\ \boldsymbol{y}_t \end{bmatrix} = \begin{bmatrix} A & \hat{K}_t \\ C & I \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}}_t \\ \hat{\boldsymbol{v}}_t \end{bmatrix}, \quad \hat{\boldsymbol{x}}_0 = 0 \tag{23}$$

where

$$\langle \hat{\boldsymbol{v}}_t, \hat{\boldsymbol{v}}_s \rangle_{\frac{I}{\infty}} = \hat{R}_t \delta_{ts}, \quad t, s = 0, 1, \cdots, \tau - 1$$

and where $\langle \hat{\boldsymbol{v}}_t, \hat{\boldsymbol{x}}_s \rangle_{\frac{I}{\infty}} = 0$, $t \ge s$. Moreover, the orthogonal projection of Y_τ^+ onto Y_τ^- is given by the state matrix $\hat{\boldsymbol{x}}_\tau$ as follows

$$\hat{\mathrm{E}}_{\frac{I}{\infty}}(Y_\tau^+ \mid Y_\tau^-) = \mathcal{O}_\tau \hat{\boldsymbol{x}}_\tau.$$

By using the above finite-interval realization, we compute the matrices $\hat{L}_{ij}$ defined in (15).

Theorem 2 *The matrices $\hat{L}_{ij} \in \mathbb{R}^{p\times p}$ defined in (15) are given by*

$$\hat{L}_{i,j} := \begin{cases} I_p & (i = j = 0, 1, \cdots, \tau - 1) \\ CA^{i-j-1}\hat{K}_j & \begin{pmatrix} j < i \le 2\tau - 1 \\ 0 \le j \le \tau - 1 \end{pmatrix} \end{cases} \tag{24}$$

and where $\hat{K}_j$ is defined by (21).

4 Finite-Interval Realization

We show that the system matrices A, C and $\hat{K}_t$ in (23) are derived by the decomposition of the matrix $\hat{\mathcal{S}}_\tau$ in Theorem 1.

Lemma 2 *The block matrix $\hat{\mathcal{S}}_\tau$ has rank n, and satisfies*

$$\hat{\mathcal{S}}_\tau = \mathcal{O}_\tau \hat{\mathcal{F}}_\tau$$

where $\mathcal{O}_\tau$ is the extended observability matrix, and $\mathcal{F}_\tau$ is defined by

$$\hat{\mathcal{F}}_\tau := \left[\begin{array}{cccc} \hat{K}_{\tau-1} & A\hat{K}_{\tau-2} & \cdots & A^{\tau-1}\hat{K}_0 \end{array} \right].$$

Theorem 3 *Given $\hat{\mathcal{S}}_\tau$, $\hat{\mathcal{R}}_\tau^-$ and $\Phi_{-\tau}$, we compute the weighted SVD:*

$$\Psi_{-\tau}^{-\frac{1}{2}} \hat{\mathcal{S}}_\tau (\hat{\mathcal{R}}_\tau^-)^{\frac{1}{2}} = \acute{U}\acute{\Sigma}\acute{V}^T, \quad \acute{\Sigma} \in \mathbb{R}^{n\times n}. \tag{25}$$

Then, the matrix $\mathcal{O}_\tau$ and $\hat{\mathcal{F}}_\tau$ are given by

$$\mathcal{O}_\tau = \Psi_{-\tau}^{\frac{1}{2}} \acute{U}\acute{\Sigma}^{\frac{1}{2}}, \quad \hat{\mathcal{F}}_\tau = \acute{\Sigma}^{\frac{1}{2}}\acute{V}^T (\hat{\mathcal{R}}_\tau^-)^{-\frac{1}{2}} \tag{26}$$

where $\acute{\Sigma}^{\frac{1}{2}} = \acute{\Sigma}^{\frac{T}{2}}$ is diagonal.

Lemma 2 and Theorem 3 provide the desired decomposition of $\Lambda_k = CA^{k-1}B$ where the extended observability matrix $\mathcal{O}_k$ in (10) is calculated in (26). The SVD of (25) yields a desired decomposition of $\hat{\mathcal{S}}_\tau$, however it is not a block Hankel matrix.

In terms of $\hat{L}_{i,j}$ of (24), define the matrix

$$\hat{\mathcal{T}}_\tau := \left[\begin{array}{cccc} \hat{L}_{\tau,\tau-1} & \hat{L}_{\tau-1,\tau-2} & \cdots & \hat{L}_{1,0} \\ \hat{L}_{\tau+1,\tau-1} & \hat{L}_{\tau,\tau-2} & \cdots & \hat{L}_{2,0} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{L}_{2\tau-1,\tau-1} & \hat{L}_{2\tau-2,\tau-2} & \cdots & \hat{L}_{\tau,0} \end{array} \right],$$

where $\hat{\mathcal{T}}_\tau \in \mathbb{R}^{\tau p\times\tau p}$. We obtain $\hat{K}_t$ in (23) as follows.

Lemma 3 *Define the non-stationary gains as*

$$\hat{\mathcal{K}}_\tau := [\hat{K}_{\tau-1} \;\; \hat{K}_{\tau-2} \;\; \cdots \;\; \hat{K}_0]. \tag{27}$$

Then, we have the decomposition $\hat{\mathcal{T}}_\tau = \mathcal{O}_\tau \hat{\mathcal{K}}_\tau$, and hence the non-stationary gains are computed by

$$\hat{\mathcal{K}}_\tau = \mathcal{O}_\tau^\dagger \hat{\mathcal{T}}_\tau \tag{28}$$

where $(\cdot)^\dagger$ denotes the pseudo-inverse.

Summarizing above results, a finite-interval realization of a stationary process is obtained by the following steps.

Finite-Interval Realization of a Stationary Process

Step 1: Given Y_τ^- and Y_τ^+, we compute $\hat{V}_\tau^-$, $\hat{V}_\tau^+$ and $\hat{\mathcal{S}}_\tau$ by (17), and then compute the covariance matrix

$$\hat{\mathcal{R}}_\tau^- = \text{block-diag}(\hat{R}_{\tau-1},\, \hat{R}_{\tau-2},\, \cdots,\, \hat{R}_0). \tag{29}$$

Step 2: Compute the weighted SVD of (25) and obtain $\mathcal{O}_\tau$ from (26).

Step 3: Compute A and C by

$$\begin{aligned} &\mathcal{O}_\tau(1:p(\tau-1),:)A = \mathcal{O}_\tau(p+1:p\tau,:) \\ &C = \mathcal{O}_\tau(1:p,:). \end{aligned}$$

Step 4: Compute the gain matrices $\hat{K}_t$ and the covariance matrices $\hat{R}_t$, $t = 0, 1, \cdots, \tau-1$ by (28) and (29), respectively.

The system (23) with matrices A, C, $\hat{K}_t$ and $\hat{R}_t$ $(t = 0,\, 1, \cdots,\, \tau-1)$ given above is a forward non-stationary realization of $\boldsymbol{y}_t$ for $t = 0,\, 1, \cdots,\, \tau-1$.

5 Subspace Identification Method

We observe that a triplet $\{A, B, C\}$ derived in Section 4 is a finite-interval stochastically balanced realization at time τ, and that $\hat{R}_{\tau-1}$ and $\hat{K}_{\tau-1}$ in (20) and (21) converge to $\hat{R}_\infty$ and $\hat{K}_\infty$ for $\tau \to \infty$, respectively. Thus, we see that the use of quadruple $(A, C, \hat{K}_{\tau-1}, \hat{R}_{\tau-1})$ is most natural for approximating a stationary process y_t instead of $(A, C, \hat{K}_\infty, \hat{R}_\infty)$.

Usually, in real system identification, we have a finite string of observed time series $\{y_0,\, y_1,\, \cdots,\, y_{\nu+2\tau-2}\}$ with ν and τ sufficiently large, where we approximate covariance matrices as $\Lambda_{i-j} \approx \frac{1}{\nu}\sum_{t=k}^{k+\nu-1} y_{t+i}y_{t+j}^T$.

For $t = 0, 1, \cdots, 2\tau-1$, define

$$\boldsymbol{y}_t := \left[\begin{array}{cccc} y_t & y_{t+1} & \cdots & y_{t+\nu-1} \end{array} \right] \in \mathbb{R}^{p\times\nu}.$$

Define bilinear form as $\langle \boldsymbol{y}_i, \boldsymbol{y}_j \rangle_{\frac{I}{\nu}} := \frac{1}{\nu}\boldsymbol{y}_i\boldsymbol{y}_j^T$ so that we approximate Λ_{i-j} by $\langle \boldsymbol{y}_i, \boldsymbol{y}_j \rangle_{\frac{I}{\nu}}$. Also define Y_τ^- and Y_τ^+ as in (5) where we assume that the positivity condition is satisfied for observed data:

$$\left\langle \left[\begin{array}{c} Y_\tau^- \\ Y_\tau^+ \end{array} \right], \left[\begin{array}{c} Y_\tau^- \\ Y_\tau^+ \end{array} \right] \right\rangle_{\frac{I}{\nu}} > 0.$$

Subspace Identification Method

Step 1: Compute the following decomposition

$$\left[\begin{array}{c} Y_\tau^- \\ Y_\tau^+ \end{array} \right] = \left[\begin{array}{cc} \hat{\mathcal{L}}_\tau^- & 0 \\ \hat{\mathcal{S}}_\tau & \hat{\mathcal{L}}_\tau^+ \end{array} \right] \left[\begin{array}{c} \hat{V}_\tau^- \\ \hat{V}_\tau^+ \end{array} \right] \tag{30}$$

where $\hat{\mathcal{L}}_\tau^-$, $\hat{\mathcal{L}}_\tau^+$ and $\hat{\mathcal{S}}_\tau$ are described as

$$\hat{\mathcal{L}}_\tau^- = \begin{bmatrix} \hat{L}_{\tau-1,\tau-1} & \hat{L}_{\tau-1,\tau-2} & \cdots & \hat{L}_{\tau-1,0} \\ & \hat{L}_{\tau-2,\tau-2} & \cdots & \hat{L}_{\tau-2,0} \\ & & \ddots & \vdots \\ 0 & & & \hat{L}_{0,0} \end{bmatrix},$$

$$\hat{\mathcal{L}}_\tau^+ = \begin{bmatrix} \hat{L}_{\tau,\tau} & & & 0 \\ \hat{L}_{\tau+1,\tau} & \hat{L}_{\tau+1,\tau+1} & & \\ \vdots & \vdots & \ddots & \\ \hat{L}_{2\tau-1,\tau} & \hat{L}_{2\tau-1,\tau+1} & \cdots & \hat{L}_{2\tau-1,2\tau-1} \end{bmatrix},$$

$$\hat{\mathcal{S}}_\tau = \begin{bmatrix} \hat{L}_{\tau,\tau-1} & \hat{L}_{\tau,\tau-2} & \cdots & \hat{L}_{\tau,0} \\ \hat{L}_{\tau+1,\tau-1} & \hat{L}_{\tau+1,\tau-2} & \cdots & \hat{L}_{\tau+1,0} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{L}_{2\tau-1,\tau-1} & \hat{L}_{2\tau-1,\tau-2} & \cdots & \hat{L}_{2\tau-1,0} \end{bmatrix},$$

where $\hat{L}_{i,i} = I$, and where

$$\hat{\mathcal{R}}_\tau^- = \text{block-diag}\,(\hat{R}_{\tau-1}, \hat{R}_{\tau-2}, \cdots, \hat{R}_0),$$
$$\hat{\mathcal{R}}_\tau^+ = \text{block-diag}\,(\hat{R}_\tau, \hat{R}_{\tau+1}, \cdots, \hat{R}_{2\tau-1}).$$

Step 2: Define $\Psi_{-\tau} := \langle Y_\tau^+, Y_\tau^+ \rangle_{\frac{I}{\nu}}$ and compute the weighted SVD of $\hat{\mathcal{S}}_\tau$ as

$$\Psi_{-\tau}^{-\frac{1}{2}} \hat{\mathcal{S}}_\tau (\hat{\mathcal{R}}_\tau^-)^{\frac{1}{2}} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T$$
$$= U_1 \Sigma_1 V_1^T.$$

Step 3: Define $\mathcal{O}_\tau$ and $\hat{\mathcal{F}}_\tau$ as

$$\mathcal{O}_\tau = \Psi_{-\tau}^{\frac{1}{2}} U_1 \Sigma_1^{\frac{1}{2}}, \quad \hat{\mathcal{F}}_\tau = \Sigma_1^{\frac{1}{2}} V_1^T (\hat{R}_\tau^-)^{-\frac{1}{2}}.$$

Step 4: Compute $\hat{A}$, $\hat{C}$, $\hat{K}_{\tau-1}$ and $\hat{R}_{\tau-1}$ as

$$\mathcal{O}_\tau(1:(\tau-1)p, :)\hat{A} = \mathcal{O}_\tau(p+1:\tau p, :),$$
$$\hat{C} = \mathcal{O}_\tau(1:p, :),$$
$$\hat{K}_{\tau-1} = \hat{\mathcal{F}}_\tau(:, 1:p),$$
$$\hat{R}_{\tau-1} = \hat{\mathcal{R}}_\tau^-(1:p, 1:p).$$

We see that the system

$$\begin{bmatrix} \hat{x}(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} \hat{A} & \hat{K}_{\tau-1} \\ \hat{C} & I \end{bmatrix} \begin{bmatrix} \hat{x}(t) \\ \hat{v}(t) \end{bmatrix}$$

with $\mathrm{E}\{\hat{v}(s)\hat{v}(t)^T\} = \hat{R}_{\tau-1}\delta_{st}$ is an approximation for the balanced stochastic realization of a stationary process $y(t)$ for observed data $\{y_0, y_1, \cdots, y_{\nu+2\tau-2}\}$.

6 Conclusions

In this paper, along the line of [6], we have considered a stochastic realization problem on a finite interval by using a Hilbert space approach [4, 5]. To this end, we have also employed the representation of the state and state covariance matrix due to Van Overschee and De Moor [8], which is extended to the present Hilbert space setting.

In summary, given finite covariance data $\{\Lambda_0, \Lambda_1, \cdots, \Lambda_{2\tau-1}\}$, we have re-derived a finite-interval realization algorithm for a stationary process due to [4, 5] based on the LQ decomposition in a Hilbert space, and developed a new method of computing non-stationary system matrices $(A, C, \hat{K}_t, \hat{R}_t)$, $t = 0, 1, \cdots, \tau-1$ by using the SVD of the matrix obtained by the LQ decomposition. Moreover, we have briefly discussed a stochastic subspace identification method.

Acknowledgments This work was supported by Research Project, Grant-in-Aid for Scientific Research, JSPS.KAKENHI 15760316.

References

[1] H. Akaike, "Markovian representation of stochastic processes by canonical variables," *SIAM J. Control*, vol. 13, no. 1, pp. 162–173, 1975.

[2] U. B. Desai, D. Pal and R. D. Kirkpatrick, "A realization approach to stochastic model reduction," *International J. Control*, vol. 42, no. 4, pp. 821–838, 1985.

[3] P. L. Faurre, "Stochastic realization algorithms," In *System Identification: Advances and Case Studies* (R. K. Mehra and D. G. Lainiotis, eds.), Academic, pp. 1–25, 1976.

[4] A. Lindquist and G. Picci, "Geometric methods for state space identification," In *Identification, Adaptation, Learning: The Science of Learning Models from Data* (S. Bittanti and G. Picci, eds.), Springer, pp. 1–69, 1996.

[5] A. Lindquist and G. Picci, "Canonical correlation analysis, approximate covariance extension, and identification of stationary time series," *Automatica*, vol. 32, no. 5, pp. 709–733, 1996.

[6] H. Tanaka and T. Katayama, "A stochastic realization in a Hilbert space based on "LQ decomposition" with application to subspace identification," Proc. 13th IFAC symposium on System Identification (SYSID2003) (to appear).

[7] P. Van Overschee and B. De Moor, "Subspace algorithms for the stochastic identification problem," *Automatica*, vol. 29, no. 3, pp. 649–660, 1993.

[8] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*, Kluwer Academic Pub., 1996.