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 realization with finite covariance data based on “LQ decomposition” in a Hilbert space generated by a stationary second order process under the assumption that the covariance matrix has a decomposition FHF^T, where F, H ε R^{n×n}, and a finite-dimension realization with Fε R^{n×n}.

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

\[ y(t) := \begin{bmatrix} y_{t} & y_{t+1} & y_{t+2} & \cdots \end{bmatrix} \in \mathbb{R}^{p×\infty}. \]

We also define a vector space spanned by all finite linear combinations of \{y(t)\} as

\[ \mathcal{Y} := \left\{ \sum_{k} a_k^{T} y_k \mid a_k \in \mathbb{R}^p, k = 0, \pm 1, \cdots \right\}. \]

For the elements \( a^{T}y_{t} \) and \( b^{T}y_{j} \in \mathcal{Y} \), we define an inner product by

\[ \langle a^{T}y_{t}, b^{T}y_{j} \rangle_{\mathcal{Y}} := \lim_{\nu \to \infty} \frac{1}{\nu} \sum_{k=t_0}^{t_0+\nu-1} a^{T}y_{k+i}b^{T}y_{k+j} \]

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

Let $U$ be a Hilbert subspace of $Y^\infty$, and the orthogonal projection of $\eta$ in $Y^\infty$ onto the space $U$ be denoted by $\hat{E}_U(\eta | U)$. Also, let the row space spanned by a matrix $U$ be expressed as $\text{span}(U)$. If $\langle U, U \rangle$ has an inverse, the orthogonal projection is written as

$$
\hat{E}_U(\eta | U) := \hat{E}_U(\eta | \text{span}(U)) = \langle \eta, U \rangle U^{-1}. \quad (3)
$$

We extend $Y^\infty$ to $Y^{* \times \infty}$ so that matrices are included as its elements.

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

$$
\begin{bmatrix}
x_{t+1} \\
y_t
\end{bmatrix} =
\begin{bmatrix}
F & W_t \\
H & V_t
\end{bmatrix}
\begin{bmatrix}
x_t \\
v_t
\end{bmatrix}
$$

where $F \in \mathbb{R}^{n \times n}$ and $H \in \mathbb{R}^{p \times n}$ satisfy a decomposition $A_k = HF^{k-1}G$, $x_t \in Y^{n \times \infty}$ is a state matrix, and the elements of tail matrices, $w_t \in Y^{p \times \infty}$ and $v_t \in Y^{p \times \infty}$ are white noises satisfying

$$
\left\langle \begin{bmatrix} w_s \\ v_s \end{bmatrix}, \begin{bmatrix} w_t \\ v_t \end{bmatrix} \right\rangle = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{st}
$$

with $R > 0$.

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

$$
\dot{x}_{t+1} = F \dot{x}_t + \hat{\Gamma}_t (y_t - H \hat{x}_t), \quad x_0 = 0
$$

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

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

$$
\begin{bmatrix}
\dot{x}_{t+1} \\
y_t
\end{bmatrix} =
\begin{bmatrix}
F & \hat{\Gamma}_t \\
H & I
\end{bmatrix}
\begin{bmatrix}
\dot{x}_t \\
v_t
\end{bmatrix}, \quad \dot{x}_0 = 0 \quad (4)
$$

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

In this paper, we assume that a set of exact but finite covariance data $\{A_0, A_1, A_2, \cdots, A_{2\tau-1}\}$ is available with $\tau > n$; this is equivalent to the fact that a finite number of tail matrices $y_t \in Y^{p \times \infty}$, $t = 0, 1, \cdots, 2\tau - 1$ are given. Under this assumption, the problem is to give a finite-interval realization of $y_t$ by “LQ decomposition” in a Hilbert space and provide a method of computing the system matrices $F, H, \hat{\Gamma}_t$ and $\hat{\Gamma}_t$ 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 $y_t \in Y^{p \times \infty}$, $t = 0, 1, \cdots, 2\tau - 1$, we define data matrices as

$$
Y_t := \begin{bmatrix}
y_{t-1} \\
y_{t-2} \\
\vdots \\
y_1 \\
y_0
\end{bmatrix}, \quad Y_t^+ := \begin{bmatrix}
y_t \\
y_{t+1} \\
\vdots \\
y_{2\tau-2} \\
y_{2\tau-1}
\end{bmatrix} \quad (5)
$$

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

$$
Z_{-s} := \begin{bmatrix}
z_{-s} \\
z_{-s+1} \\
\vdots \\
z_{-2\tau+2} \\
z_{-2\tau+1}
\end{bmatrix}, \quad Z_{+s} := \begin{bmatrix}
z_{s+1} \\
z_{s+2} \\
\vdots \\
z_{\tau} \\
z_0
\end{bmatrix} \quad (6)
$$
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 $Y^- = Z^-\tau$ and $Y^+ = Z^+\tau$, where the former are termed the past data matrices, while the latter the future data matrices.

Moreover, we define covariance matrices

$$
\Phi_t := (Y_t^-, Y_t^+) = \begin{bmatrix}
A_0 & A_1 & A_2 & \cdots & A_{t-1} \\
A_1^T & A_0 & A_1 & \cdots & A_{t-2} \\
A_2^T & A_1^T & A_0 & \cdots & A_{t-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A_{t-1}^T & A_{t-2}^T & A_{t-3}^T & \cdots & A_0
\end{bmatrix},
$$

$$
\Psi_t := (Z^+_{t\tau}, Z^-_{t\tau}) = \begin{bmatrix}
A_0 & A_1^T & A_2^T & \cdots & A_{t-1}^T \\
A_1 & A_0 & A_1 & \cdots & A_{t-2} \\
A_2 & A_1 & A_0 & \cdots & A_{t-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A_{t-1} & A_{t-2} & A_{t-3} & \cdots & A_0
\end{bmatrix},
$$

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

$$
\mathcal{H}_\tau = (Y^+_{\tau\tau}, Y^-_{\tau\tau}) = \begin{bmatrix}
A_1 & A_2 & A_3 & \cdots & A_\tau \\
A_2 & A_3 & A_4 & \cdots & A_{\tau+1} \\
A_3 & A_4 & A_5 & \cdots & A_{\tau+2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A_\tau & A_{\tau+1} & A_{\tau+2} & \cdots & A_{2\tau-1}
\end{bmatrix} = (Z^+_{\tau\tau}, Z^-_{\tau\tau}).
$$

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])

$$
\Psi_{\tau \tau}^{-1/2} \mathcal{H}_\tau \Phi_{\tau \tau}^{-1/2} = [U_\tau \ V_\tau] \begin{bmatrix}
\Sigma_\tau & 0 \\
0 & \Sigma_\tau
\end{bmatrix} \begin{bmatrix}
V_\tau^T \\
V_\tau^T\Sigma_\tau V_\tau
\end{bmatrix}
$$

where $\text{rank} \, \Sigma_\tau = n$, and $U_\tau^T U_\tau = I_n$, $V_\tau^T V_\tau = I_n$. Hence, we get

$$
\mathcal{H}_\tau = \Psi_{\tau \tau}^{-1/2} U_\tau \Sigma_\tau V_\tau^T \Phi_{\tau \tau}^{-1/2}.
$$

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

$$
\mathcal{O}_\tau := \Psi_{\tau \tau}^{-1/2} U_\tau \Sigma_\tau \Phi_{\tau \tau}^{-1/2},
$$

$$
\mathcal{C}_\tau := \Sigma_\tau \Phi_{\tau \tau}^{-1/2} V_\tau^T
$$

with $\text{rank} \, \mathcal{O}_\tau = n$, $\text{rank} \, \mathcal{C}_\tau = n$, 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

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

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

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

Let $(\hat{A}, \hat{B}, \hat{C})$ 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 $\hat{A} = Q^{-1}_\tau A Q_\tau$, $\hat{B} = Q^{-1}_\tau B$, and $\hat{C} = C Q_\tau$, where $Q_\tau \in \mathbb{R}^{n \times n}$ is a non-singular transform [5], so that we have $A_k = C A_{k-1} B$, $k = 1, 2, \cdots, 2\tau - 1$. The triplet $(A, B, C)$ obtained above is a finite-interval stochastically balanced realization which is minimal and dependent on $\tau$ [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 $t = 1, \cdots, 2\tau - 1$

$$
\hat{v}_t := y_t - \hat{E}_\tau(y_t | Y^-_t)
$$

with the initial condition $\hat{v}_0 := y_0$.

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

$$
\langle \hat{v}_i, \hat{v}_j \rangle = \hat{R}_{ij} \delta_{ij}, \quad i, j = 0, 1, \cdots, 2\tau - 1
$$

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

$$
\langle y_i, \hat{v}_j \rangle = 0, \quad 0 \leq i < j \leq 2\tau - 1.
$$

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

$$
\hat{L}_{i,j} := \langle y_i, \hat{v}_j \rangle = \hat{R}_{ij}^{-1}, \quad 0 \leq j \leq i \leq 2\tau - 1.
$$
An explicit form of \( \hat{L}_{i,j} \in \mathbb{R}^{p \times p} \) for \( j \leq i \leq 2\tau - 1, \ 0 \leq j \leq \tau - 1 \) is provided later in (24).

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

\[
\hat{L}_{\tau} := \begin{bmatrix}
\hat{L}_{\tau-1,r-1} & \hat{L}_{\tau-1,r-2} & \cdots & \hat{L}_{\tau-1,0} \\
0 & \hat{L}_{\tau,0} \\
\hat{L}_{\tau+1,r-1} & \hat{L}_{\tau+1,r-2} & \cdots & \hat{L}_{\tau+1,0} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{L}_{2\tau-1,r-1} & \hat{L}_{2\tau-1,r-2} & \cdots & \hat{L}_{2\tau-1,0}
\end{bmatrix},
\]

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

\[
\hat{V}_t^- = \begin{bmatrix}
\hat{v}_{t-1} \\
\vdots \\
\hat{v}_0 
\end{bmatrix}, \quad \hat{V}_t^+ = \begin{bmatrix}
\hat{v}_t \\
\vdots \\
\hat{v}_{2\tau-1}
\end{bmatrix}
\]

and covariance matrices:

\[
\hat{R}_
abla := \langle \hat{V}_\tau^-, \hat{V}_\tau^- \rangle, \quad \hat{R}_\nabla := \langle \hat{V}_\tau^+, \hat{V}_\tau^+ \rangle.
\]

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

\[
\begin{bmatrix}
Y_t^- \\
Y_t^+
\end{bmatrix} = \begin{bmatrix}
\hat{L}_{\tau}^- & \hat{L}_{\tau}^+ \\
\hat{S}_{\tau} & \hat{S}_{\tau}^+
\end{bmatrix} \begin{bmatrix}
\hat{V}_t^- \\
\hat{V}_t^+
\end{bmatrix}
\]

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

\[
\begin{bmatrix}
\hat{V}_t^- \\
\hat{V}_t^+
\end{bmatrix} = \begin{bmatrix}
\hat{R}_t^- & 0 \\
0 & \hat{R}_t^+
\end{bmatrix}.
\]

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

\[
\hat{E}_{\nabla} (Y_t^+ | Y_t^-) = \hat{S}_{\tau} \hat{V}_t^-.
\]

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

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

\[
\hat{P}_t := C_t \Phi_t^{-1} C_t^T, \quad t = 1, \cdots, \tau
\]

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

\[
C_t = \begin{bmatrix}
B & AB & \cdots & A^{t-1}B
\end{bmatrix}, \quad t = 1, \cdots, \tau
\]

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

**Proposition 1** ([8]) The matrix \( \hat{P}_t \) satisfies the following discrete-time Riccati equation with \( P_0 = 0 \)

\[
\hat{P}_{t+1} = A \hat{P}_t A^T + (B - A \hat{P}_t C^T)(A_0 - C \hat{P}_t C^T)^{-1}(B - A \hat{P}_t C^T)^T
\]

for \( t = 0, 1, 2, \cdots, \tau - 1 \). In terms of the solution \( \hat{P}_t \) of Riccati equation, we define matrices

\[
\hat{R}_t := A_0 - C \hat{P}_t C^T, \quad \hat{K}_t := (B - A \hat{P}_t C^T)(A_0 - C \hat{P}_t C^T)^{-1},
\]

for \( t = 0, 1, \cdots, \tau - 1 \). Also, define (18)

\[
\hat{x}_t := C_t \Phi_t^{-1} Y_t^-.
\]

Then, we can prove the following lemma.

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

\[
\begin{bmatrix}
\hat{x}_{t+1} \\
y_t
\end{bmatrix} = \begin{bmatrix}
A & \hat{K}_t \\
C & I
\end{bmatrix} \begin{bmatrix}
\hat{x}_t \\
y_t
\end{bmatrix}, \quad \hat{x}_0 = 0
\]

where

\[
\langle \hat{v}_t, \hat{v}_s \rangle = \hat{R}_t \delta_{ts}, \quad t, s = 0, 1, \cdots, \tau - 1
\]

and where \( \langle \hat{v}_t, \hat{x}_s \rangle = 0, \ t \geq s \). Moreover, the orthogonal projection of \( Y_t^+ \) onto \( Y_t^- \) is given by the state matrix \( \hat{x}_\tau \) as follows

\[
\hat{E}_{\nabla} (Y_t^+ | Y_t^-) = \hat{S}_{\tau} \hat{V}_\tau^-.
\]

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

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

\[
\hat{L}_{i,j} := \begin{cases}
I_p & (i = j = 0, 1, \cdots, \tau - 1) \\
CA_{t-j-1} \hat{K}_j & (0 \leq j \leq \tau - 1)
\end{cases}
\]

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 $\mathcal{S}_t$ in Theorem 1.

Lemma 2 The block matrix $\hat{S}_t$ has rank $n$, and satisfies
$$\hat{S}_t = \mathcal{O}_t \mathcal{F}_t$$
where $\mathcal{O}_t$ is the extended observability matrix, and $\mathcal{F}_t$ is defined by
$$\mathcal{F}_t := [ \hat{K}_{t-2} A \hat{K}_{t-2} \cdots A^{T-1} \hat{K}_0 ] .$$

Theorem 3 Given $\hat{S}_t$, $\hat{R}_t$ and $\Phi_{-t}$, we compute the weighted SVD:
$$\psi^{-\frac{1}{2}} \hat{S}_t (\hat{R}_t)^{\frac{1}{2}} = \hat{U} \hat{Y} \hat{V}^T, \quad \hat{Y} \in \mathbb{R}^{n \times n}. \quad (25)$$
Then, the matrix $\mathcal{O}_t$ and $\mathcal{F}_t$ are given by
$$\mathcal{O}_t = \psi^{\frac{1}{2}} \hat{U} \hat{\Sigma}^{\frac{1}{2}} \hat{v}, \quad \mathcal{F}_t = \hat{\Sigma}^{\frac{1}{2}} \hat{v}^T (\hat{R}_t)^{-\frac{1}{2}} \quad (26)$$
where $\hat{\Sigma}^{\frac{1}{2}}$ is diagonal.

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

In terms of $\hat{L}_{i,j}$ of (24), define the matrix
$$\hat{T}_t := \begin{bmatrix}
\hat{L}_{t-1} & \hat{L}_{t-2} & \cdots & \hat{L}_{0} \\
\hat{L}_{t-1} & \hat{L}_{t-2} & \cdots & \hat{L}_{0} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{L}_{t} & \hat{L}_{t-1} & \cdots & \hat{L}_{0}
\end{bmatrix},$$
where $\hat{T}_t \in \mathbb{R}^{p \times tp}$. We obtain $\hat{K}_t$ in (23) as follows.

Lemma 3 Define the non-stationary gains as
$$\hat{K}_t := [ \hat{K}_{t-2} \hat{K}_{t-3} \cdots \hat{K}_0 ]. \quad (27)$$
Then, we have the decomposition $\hat{T}_t = \mathcal{O}_t \hat{K}_t$, and hence the non-stationary gains are computed by
$$\hat{K}_t = \mathcal{O}_t^T \hat{T}_t \quad (28)$$
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 $Y_t^-$ and $Y_t^+$, we compute $\hat{V}_t^-$, $\hat{V}_t^+$ and $\hat{S}_t$ by (17), and then compute the covariance matrix
$$\hat{R}_t = \text{block-diag}(\hat{R}_{t-1}, \hat{R}_{t-2}, \cdots, \hat{R}_0). \quad (29)$$

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

Step 3: Compute $A$ and $C$ by
$$\mathcal{O}_t (1 : p(\tau - 1), :) A = \mathcal{O}_t (p + 1 : p\tau, :) \quad C = \mathcal{O}_t (1 : p, :). \quad (29)$$

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

The system (23) with matrices $A, C, \hat{K}_t$ and $\hat{R}_t$ is convergent to $R_{\infty}$, respectively. Thus, we see that the use of quadruple $(A, C, \hat{K}_t, \hat{R}_t)$ is most natural for approximating a stationary process $y_t$ instead of $(A, C, \hat{K}_t, \hat{R}_t)$.

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 $\tau$, and that $\hat{R}_{t-1}$ and $\hat{K}_{t-1}$ in (20) and (21) converge to $R_{\infty}$ and $K_{\infty}$ for $\tau \rightarrow \infty$, respectively. Thus, we see that the use of quadruple $(A, C, \hat{K}_t, \hat{R}_t)$ is most natural for approximating a stationary process $y_t$ instead of $(A, C, \hat{K}_t, \hat{R}_t)$.

Usually, in real system identification, we have a finite string of observed time series $\{y_0, y_1, \cdots, y_{2\tau - 1}\}$ with $\tau$ sufficiently large, where we approximate covariance matrices as $A_{i-j} \approx \frac{1}{2\tau} \sum_{k=-\tau}^{\tau} y_{t+i} y_{t+j}$. For $t = 0, 1, \cdots, 2\tau - 1$, define
$$y_t := [ y_t, y_{t+1}, \cdots, y_{t+\tau-1} ] \in \mathbb{R}^{p \times \nu}.$$ Define bilinear form as $\langle y_i, y_j \rangle := \frac{1}{\nu} y_i y_j^T$ so that we approximate $A_{i-j}$ by $\langle y_i, y_j \rangle$. Also define $Y_t^-$ and $Y_t^+$ as in (5) where we assume that the positivity condition is satisfied for observed data:
$$\begin{bmatrix} Y_t^- \ Y_t^+ \end{bmatrix} \begin{bmatrix} Y_t^- \ Y_t^+ \end{bmatrix}^T > 0.$$ Subspace Identification Method

Step 1: Compute the following decomposition
$$\begin{bmatrix} Y_t^- \ Y_t^+ \end{bmatrix} = \begin{bmatrix} \hat{L}_t \hat{0} \end{bmatrix} \begin{bmatrix} \hat{V}_t^- \hat{V}_t^+ \end{bmatrix} \begin{bmatrix} \hat{S}_t \hat{L}_t^T \end{bmatrix}. \quad (30)$$
where $\hat{L}_\tau$, $\hat{L}_\tau^+$ and $\hat{S}_\tau$ are described as

\[
\hat{L}_\tau = \begin{bmatrix}
L_{\tau-1,\tau-1} & L_{\tau-1,\tau-2} & \cdots & L_{\tau-1,0} \\
L_{\tau-2,\tau-2} & L_{\tau-2,0} & \cdots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & L_{0,0}
\end{bmatrix},
\]

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

\[
\hat{S}_\tau = \begin{bmatrix}
L_{\tau,\tau-1} & L_{\tau,\tau-2} & \cdots & L_{\tau,0} \\
L_{\tau+1,\tau-1} & L_{\tau+1,\tau-2} & \cdots & L_{\tau+1,0} \\
\vdots & \ddots & \ddots & \ddots \\
L_{2\tau-1,\tau-1} & L_{2\tau-1,\tau-2} & \cdots & L_{2\tau-1,0}
\end{bmatrix},
\]

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

\[
\hat{R}_\tau = \text{block-diag}(\hat{R}_{\tau-1}, \hat{R}_{\tau-2}, \cdots, \hat{R}_0),
\]

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

Step 2: Define $\Psi_r := (Y_1^+, Y_1^+)_{\tau}$ and compute the weighted SVD of $\hat{S}_\tau$ as

\[
\Psi_r^T \hat{S}_\tau (\hat{R}_\tau^+) = [V_1 \quad V_2] \begin{bmatrix}
\Sigma_1 & 0 \\
0 & \Sigma_2
\end{bmatrix} \begin{bmatrix}
V_1 \\
V_2
\end{bmatrix}^T
= U_1 \Sigma_1 V_1^T.
\]

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

\[
\mathcal{O}_r = \Psi_r^T U_1 \Sigma_1^+, \quad \hat{\mathcal{F}}_r = \Sigma_1^+ V_1^T (\hat{R}_\tau)^{-\frac{1}{2}}.
\]

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

\[
\mathcal{O}_r(1 : (\tau-1)p,:) \hat{A} = \mathcal{O}_r(p+1 : \tau p,:),
\]

\[
\hat{C} = \mathcal{O}_r(1 : p,:),
\]

\[
\hat{K}_{\tau-1} = \hat{\mathcal{F}}_r(:, 1 : p),
\]

\[
\hat{R}_{\tau-1} = \hat{R}_\tau(1 : p, 1 : p).
\]

We see that the system

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

with $E\hat{v}(s)\hat{v}(t)^T = \hat{R}_{\tau-1}\delta_{st}$ is an approximation for the balanced stochastic realization of a stationary process $\hat{y}(t)$ 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 $\{A_0, A_1, \cdots, A_{2\tau-1}\}$, 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 $(A, C, K_t, R_t)$, $t = 0, 1, \cdots, \tau-1$ 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


