Recursive Implementation of 4SID Prediction for Stochastic Time-varying Systems*

Kentaro Kameyama and Akira Ohsumi

Department of Mechanical and System Engineering,
Graduate School of Science and Technology, Kyoto Institute of Technology,
Matsugasaki, Sakyo, Kyoto 606-8585, Japan

E-mail: ohsumi@ipc.kit.ac.jp; Tel/Fax: +81-75-724-7352

Abstract

In this paper, a recursion of the subspace prediction (SP) algorithm which has been derived recently by the authors is investigated for time-invariant/varying stochastic systems. The proposed algorithm is tested by simulation experiments.

1 Introduction

Nowadays, 4SID-based system identification is recognized to be very efficient to model multivariable systems, including estimations of the system states only from the set of input and output data and achieving a significant level of maturity and acceptability in control system applications; however, most of these seem to point to batch processing. Recursive algorithms are urgently expectant from the viewpoint of lightening computational burden and storage cost.

Furthermore, most of actual phenomena show complex behaviors and their mathematical models are described by time-varying and/or nonlinear equations, and the nonlinear systems are sometimes treated as (high-order) linear time-varying systems from the practical point of view. Recursive algorithms contribute to adapt the 4SID method to the identification of time-varying systems.

On the other hand, in order to obtain accurate mathematical models and to realize efficient control, the estimation and/or prediction of the system are significantly important. Especially, in order to realize the subspace model predictive control the subspace prediction (SP) will be inevitably necessary [1]. The authors have proposed an SP algorithm in the 4SID framework using an idea of the angle between two subspaces of the past and the future [2]. This algorithm is based on the off-line computation, and this prevents the high-speed computation which is required to realize the subspace model predictive control.

Motivated by such a situation, we extend our result of the batch processing SP algorithm to allow on-line application by developing a recursive algorithm.

2 Problem Statement

Suppose that we are given a couple of input and output data sequences \( \{u_k, y_k\} \) and that the output data is generated from the discrete-time time-varying stochastic system:

\[
x_{k+1} = A_k x_k + B_k u_k + w_k \quad (1)
\]

\[
y_k = C_k x_k + D_k u_k + v_k, \quad (2)
\]

where \( u_k \in \mathbb{R}^m, y_k \in \mathbb{R}^l \) and \( x_k \in \mathbb{R}^n \) are input, output and state vectors; \( w_k \in \mathbb{R}^m \) and \( v_k \in \mathbb{R}^l \) are zero-mean white Gaussian sequences with covariance matrices:

\[
E \left\{ \begin{bmatrix} w_k \\ v_k \end{bmatrix} \right\} = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{k\sigma}
\]

(\( \delta_{k\sigma} \): Kronecker delta).

The system order \( n \) is assumed to be known. Let \( k_1, k_2 (k_1 < k_2) \) be two (distinct) time instants, and let \( \mathbf{Y}_n(k_1|k_2 - N + 1) \in \mathbb{R}^{m \times N} \) \((\nu = 1, 2)\) be a block Hankel matrix constructed by arranging (column) output vectors \( y_1(i) = \{y^T(i - \alpha + 1), \cdots, y^T(i)\} \) from \( i = k_2 - N + 1 \) to \( i = k_1 \), where \( \alpha \) is the size of block rows. Similarly, let \( \mathbf{U}_n(k_1|k_2 - N + 1) \in \mathbb{R}^{m \times n} \) be the block Hankel matrix constructed by the input data (e.g., [3], [4], and see [5] for continuous-time systems).

Then, under the quasi-stationarity assumption [6], [7] the input-output algebraic relationships with arguments \( k_1 \) and \( k_2 \) are given, respectively, as

\[
\mathbf{Y}_n(k_1|k_2 - N + 1) = \Gamma_n(k_1) \mathbf{X}_n(k_1|k_2 - N + 1)
\]

\[
+ H_n(k_1) U_n(k_1|k_2 - N + 1)
\]

\[
+ \Sigma_n(k_1) W_n(k_1|k_2 - N + 1) + V_n(k_1|k_2 - N + 1) \quad (\nu = 1, 2), \quad (3)
\]

where \( \Gamma_n(\cdot) \in \mathbb{R}^{\alpha \times n} \) is the extended observability matrix; \( \mathbf{X}_n(\cdot : \cdot) \in \mathbb{R}^{\alpha \times N} \) the matrix constructed by system states; \( W_n(\cdot : \cdot) \in \mathbb{R}^{n \times N} \) and \( V_n(\cdot : \cdot) \in \mathbb{R}^{\alpha \times N} \) are system and observation noise matrices constructed similarly to \( \mathbf{Y}_n(\cdot : \cdot) \); and \( H_n(\cdot), \Sigma_n(\cdot) \) are

---

* Part of this work is supported by the Japan Society for the Promotion of Science (JSPS) under Grant-in-Aid for Scientific Research (B)-16360046.
lower block triangular matrices consisting of system matrices \( \{A_k, B_k, C_k, D_k\} \).

The subspaces spanned by the column vectors of extended observability matrices \( \Gamma_\alpha(k_1) \) and \( \Gamma_\alpha(k_2) \) form a relationship which is described by the concept of angle between subspaces.

Our problem is to derive a recursive algorithm for predicting the future subspace which is spanned by an extended observability matrix \( \Gamma_\alpha(k_3) \) at a future step \( k_3 \) (\( k_3 > k_2 \)) by applying the information about the angle to the past subspace. Hence, the recursive SP problem can be stated as follows: Given a set of input and output data of the unknown linear time-varying system (1)-(2) up to the current time step \( k \), derive a recursive algorithm for predicting the subspace at \( k + \mu \) step ahead (within a similarity transformation).

3 Review of Subspace Prediction Algorithm

Given a set of input and output data up to the current time step \( k_2 \), we are interested in the estimation of all system matrices at some future time, say \( k_3 \), based on the data set. Let \( k_1, k_2 \) and \( k_3 \) be past, current and future times, respectively (\( k_1 < k_2 < k_3 \)).

Under the quasi-stationarity assumption, all system matrices are assumed to change slowly and continuously; so that the extended observability matrices at \( k = k_1 \) and \( k_2 \) can be computed as

\[
\Gamma_\alpha(k_2) = \begin{bmatrix} C_{k_2}^T (C_{k_2} A_{k_2})^T \cdots (C_{k_2} A_{k_2}^{-1})^T \end{bmatrix}^T \quad (\nu = 1, 2). \tag{4}
\]

These matrices are also considered to change slowly, such that, this can be regarded as the rotation and scaling of column vectors \( \gamma_i(k_2) \) \( (i = 1, 2, \ldots, n) \):

\[
\Gamma_\alpha(k_2) = \begin{bmatrix} \gamma_1(k_2) \gamma_2(k_2) \cdots \gamma_n(k_2) \end{bmatrix}.
\]

According to such observation, the angle between signal subspaces \( \Gamma_\alpha(k_1) \) and \( \Gamma_\alpha(k_2) \) is yielded as a result of the rotation of each column vector \( \gamma_i(k_1) \) during the interval \( k_2 - k_1 \) in an \( \alpha \)-dimensional vector space, and the principal (column) vectors of \( \text{span}_{\text{col}} \{\Gamma_\alpha(k_1)\} \) and \( \text{span}_{\text{col}} \{\Gamma_\alpha(k_2)\} \) are calculated by

\[
\Gamma_\alpha(k_1) \{\Gamma_\alpha^T(k_1) \Gamma_\alpha(k_1)\}^{-\frac{1}{2}} \Gamma_\alpha^T(k_1) \Gamma_\alpha(k_2) \{\Gamma_\alpha^T(k_2) \Gamma_\alpha(k_2)\}^{-\frac{1}{2}} = U(k_1) S_{k_1 k_2} V^T(k_2), \tag{5}
\]

where

\[
U(k_1) = [u_1(1) \cdots u_n(1)] \in \mathbb{R}^{\alpha \times n}
\]
\[
V(k_2) = [v_1(1) \cdots v_n(1)] \in \mathbb{R}^{\alpha \times n}
\]

are principal vectors of \( \Gamma_\alpha(k_\nu) \) \( (\nu = 1, 2) \) \cite{2, 10}. The angles \( \{\theta_i(k_2 | k_1)\}_{i=1,2,\ldots,n} \) between principal vectors \( \{u_i(k_1)\} \) and \( \{v_i(k_2)\} \) are related to the first \( n \) singular values \( \{\sigma_i(k_2 | k_1)\} \) with relationship:

\[
\sigma_i(k_2 | k_1) = \cos \theta_i(k_2 | k_1), \tag{6}
\]

and the concept of this relation is depicted in Fig. 1.

Then, the rotation of \( u_i(k_1) \) on the rotation plane defined by both vectors is realized by the following procedure.

First, define the orthonormal basis of the rotation plane on which \( u_i(k_1) \) and \( v_i(k_2) \) lie:

\[
e_i = \frac{u_i(k_1)}{||u_i(k_1)||^2},
\]
\[
e_j = \frac{v_i(k_2) - \langle v_i(k_2), e_i \rangle e_i}{||v_i(k_2) - \langle v_i(k_2), e_i \rangle e_i||^2},
\]

where \( \langle a, b \rangle \) denotes the inner product \( a^T b \) of \( a \) and \( b \). The rest of the vectors \( e_{ij} \) \( (j = 3, \ldots, \alpha \ell) \) of orthonormal basis of \( \mathbb{R}^{\alpha \ell} \) are arbitrarily selected as

\[
e_{ij} e_{jk} = I_{\alpha \ell} \delta_{jk}, \tag{9}
\]

(\( \delta_{jk} \): Kronecker delta; \( I_{\alpha \ell} \): unit matrix of dimension \( \alpha \ell \)). For \( j, k = 1, 2, \ldots, \alpha \ell \), this is, \{\( e_{11, e_{12}, e_{13}, \ldots, e_{1\alpha \ell} \}\} is the set of orthonormal basis; while \{\( e_{21, e_{22} \ell, \ldots, e_{2\alpha \ell} \}\} is the set of orthonormal basis of the rotation plane.

Then, \( u_i(k_1) \) is represented by the orthonormal basis as

\[
u_i(k_1) = a_{i1} e_{11} + a_{i2} e_{12} + a_{i3} e_{13} + \cdots + a_{i\alpha \ell} e_{i\alpha \ell}
\]

\[
= a_{i1} e_{11}, \tag{10}
\]

where \( a_{ij} := \langle u_i(k_1), e_{ij} \rangle \) \( (j = 1, 2, \ldots, \alpha \ell) \) and \( a_{i2} = \cdots = a_{i\alpha \ell} = 0 \). Similarly, \( v_i(k_2) \) is written as

\[
u_i(k_2) = b_{i1} e_{11} + b_{i2} e_{12} + b_{i3} e_{13} + \cdots + b_{i\alpha \ell} e_{i\alpha \ell}
\]

\[
= b_{i1} e_{11} + b_{i2} e_{12}, \tag{11}
\]

where \( b_{ij} := \langle v_i(k_1), e_{ij} \rangle \) \( (j = 1, 2, \ldots, \alpha \ell) \) and \( a_{i3} = \cdots = a_{i\alpha \ell} = 0 \). Then, the relationship between \( a_{ij} \) and \( b_{ij} \) is given as \( (j = 1, 2) \)

\[
a_{i1} \quad b_{i1} \quad \begin{bmatrix} \cos \theta_i & -\sin \theta_i \\ \sin \theta_i & \cos \theta_i \end{bmatrix} \begin{bmatrix} a_{i1} \\ a_{i2} \end{bmatrix}. \tag{12}
\]

![Fig.1. Definition of rotation plane.](image-url)
Substituting (10) and (12) into (11), we have
\[ v_i(k_1) = b_1 e_{i1} + b_2 e_{i2} \]
\[ = e_{i1} b_1 + e_{i2} b_2 \]
\[ = \begin{bmatrix} e_{i1} & e_{i2} \end{bmatrix} \begin{bmatrix} a_{i1} & a_{i2} \\
\cos \theta_i & -\sin \theta_i \\
\sin \theta_i & \cos \theta_i \\
0 & 0 \end{bmatrix} \begin{bmatrix} e_{i1} & e_{i2} \end{bmatrix} \]
\[ = \begin{bmatrix} e_{i1} & e_{i2} \end{bmatrix} \begin{bmatrix} \cos \theta_i & -\sin \theta_i & 0 & 0 \\
\sin \theta_i & \cos \theta_i & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_{i1} & e_{i2} & e_{i3} & \cdots & e_{i,at} \end{bmatrix} \]
\[ = \begin{bmatrix} \cos \theta_i & -\sin \theta_i & 0 & 0 \\
\sin \theta_i & \cos \theta_i & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_{i1} & e_{i2} & e_{i3} & \cdots & e_{i,at} \end{bmatrix} \]
\[ =: R_{\theta_i} u_i(k_1). \]

Here, let \( \hat{v}_i(k_3|k_2) \) be the estimate of \( v_i(k_3) \) based on the data up to the current time \( k_2 \). Then, the relation (13) implies that this estimate can be computed based on \( v_i(k_2) \) by
\[ \hat{v}_i(k_3|k_2) = R_{\theta_i(k_3|k_2)} v_i(k_2). \] (14)

From the relation (6), the rate of rotation during the interval \( k_2 - k_1 \) is given by
\[ \Delta \theta_i(k_2|k_1) := \frac{\arccos \sigma_i(k_3|k_1)}{k_2 - k_1} \text{ [rad/step]}, \] (15)
so that the extrapolated angle \( \theta_i(k_3|k_2) \) can be evaluated as
\[ \theta_i(k_3|k_2) = \Delta \theta_i(k_2|k_1) \{ k_3 - k_2 + \frac{1}{2} (N + \alpha - 1) \}. \] (16)

Therefore, the estimate of the extended observability matrix at the future time \( k_3 \), \( \tilde{F}_\alpha(k_3|k_2) \), can be computed as
\[ \tilde{F}_\alpha(k_3|k_2) = \begin{bmatrix} \hat{v}_1(k_3|k_2) & \hat{v}_2(k_3|k_2) & \cdots & \hat{v}_n(k_3|k_2) \end{bmatrix}. \] (17)

4 Recursive Implementation of Subspace Prediction

The recursive SP algorithm is derived incorporated with the recursive subspace identification algorithm developed in [7]-[9].

Consider the LQ-factorization of the constructed data matrix:
\[ \begin{bmatrix} U_\alpha(k_3|k_2 - N + 1) \\
U_\beta(k_3|k_2 - N + 1) \end{bmatrix} = \begin{bmatrix} L_1(k_3) & 0 & 0 \\
L_2(k_3) & L_2(k_3) & 0 \end{bmatrix} \begin{bmatrix} Q_{1\beta}(k_2) \\
Q_{2\beta}(k_2) \end{bmatrix}, \] (18)
where \( Y_\alpha(\cdot, \cdot) \) are block Hankel matrices (as appeared in (3)); \( U_\beta(\cdot, \cdot) \) is an instrumental variable matrix constructed similarly from input data. Then, the estimate of the signal subspace is derived by performing the SVD of \( L_2(k_3) L_2^T(k_2) \), and \( L_2(k_3) L_2^T(k_2) \) is renewed by the recursive 4SID algorithm using fixed input and output data size [7].

Now, write the matrices \( L_2(k_3) \) and \( L_2(k_3) L_2^T(k_2) \) as
\[ L_2(k_3) = \begin{bmatrix} s_1(k_3) & s_2(k_3) & \cdots & s_{\alpha}(k_3) \end{bmatrix} \]
\[ L_2(k_3) L_2^T(k_2) := \begin{bmatrix} h_1(k_3) & h_2(k_3) & \cdots & h_{\alpha}(k_3) \end{bmatrix}, \]
where \{s_1(\cdot)\} and \{h_1(\cdot)\} are column vectors. Let \{f_{ij}(\cdot)\} be the (i, j)-element of the matrix \( L_2(k_3) L_2^T(k_2) \) \( (f_{ij}(k_2) \neq 0) \). Then, the column vector of \( L_2(k_3) \): \( L_2(k_3) \) is represented as
\[ h_j(k_2) = f_{1j}(k_2) s_1(k_2) + f_{2j}(k_2) s_2(k_2) + \cdots + f_{\alpha j}(k_2) s_{\alpha}(k_2) \] (j = 1, \ldots, \alpha). (19)

Choosing \( \alpha \) column vectors arbitrarily from among \( \{h_j(k_2)\}_{j=1}^{\alpha} \), and construct a matrix
\[ L_a(k_2) = [h_1(k_2), \ldots, h_\alpha(k_2)] \in R^{\alpha \times n} (i \neq j). \] (20)

Then, the following relation holds:
\[ \text{span}_{\text{col}} \{ \Gamma_\alpha(k_2) \} \cong \text{span}_{\text{col}} \{ L_a(k_2) \} \]
\[ \cong \text{span}_{\text{col}} \{ L_2(k_3) L_2^T(k_2) \}. \] (21)

As a result, the computation of the angle between \( \Gamma_\alpha(k_1) \) and \( \Gamma_\alpha(k_2) \) can be performed by that between \( L_a(k_1) \) and \( L_a(k_2) \)
\[ \hat{F}_\alpha(k_1) \left\{ \hat{F}_\alpha^T(k_1) \right\}^T \hat{F}_\alpha^T(k_1) \]
\[ = L_a(k_2) \left\{ L_a^T(k_2) L_a(k_2) \right\}^T L_a^T(k_2) \]
\[ = U(k_1) S(k_1|k_2) V^T(k_2), \] (22)
and the estimate of \( \text{span}_{\text{col}} \{ \hat{F}_\alpha(k_2) \} \) is given by the principal vectors of \( L_a(k_2) \) as far as the angles between principal vectors at \( k_1 \) and \( k_2 \) hold the relation \( \theta_i(k_3|k_1) < \pi/4 \) (see Appendix).

Consequently, the recursive SP algorithm is summarized as follows:
Subspace Prediction Algorithm

Step 1: Acquire a new data set \{u(k_2), y(k_2)\}, and renew \( L_2(k_2) \) and \( L_1(k_2) \) according to the recursive algorithm proposed in [7].

Step 2: Construct \( L_0(k_2) \) and perform the SVD as (22).

Step 3: Predict the future subspace by the procedure mentioned in Section 4.

Step 4: Derive each unknown system matrices according to the 4SID framework.

5 A Numerical Example

Consider a single-input, single-output two-dimensional time-varying stochastic system with matrices:

\[
A_k = \frac{1}{2} \begin{bmatrix}
\sin(\pi k/1250) & \sin(\pi k/5000) + 0.5 \\
-\sin(\pi k/5000) - 0.5 & \sin(\pi k/1250)
\end{bmatrix},
\]

\[
B_k = \begin{bmatrix}
2.0 \\
-1.0
\end{bmatrix},

C_k = \begin{bmatrix}
1.0 & 2.0
\end{bmatrix},

D_k = 1.5.
\]

The random noises \( w_k \) and \( v_k \) are mutually independent and have common covariance \( \mathbb{E}\{w_k w_j^T\} = 0.1^2 I_2 \delta_{kj} \) and \( \mathbb{E}\{v_k v_j^T\} = 0.1^2 \delta_{kj} \). User-defined parameters are set: \( \alpha = \beta = 5 \). Each figure shows the result of 50-step ahead prediction, i.e., \( L(= k_3 - k_2) = 50 \).

Figure 2 depicts a couple of time evolutions of real and imaginary parts of typical one of predicted conjugate poles. Figures 3 and 4 show the results of 100-sample means and variances of the estimation of the real and imaginary parts of conjugate poles for the case of \( N = 250 \). In the figures predicted and true ones are depicted by chain and broken lines, respectively.

Furthermore, Fig. 5 depicts a couple of time evolutions of real and imaginary parts of typical one of predicted conjugate poles for the case of \( N = 100 \), and Figs. 6 and 7 show the results of 100 Monte Carlo experiments.

In Figs. 2 and 5, sample paths of predicted system poles of both cases track true locus well, and comparison between Figs 3, 4, 6 and 7 shows that sample variances are not extremely affected by the number of observation data \( N \).

6 Conclusion

In this paper, a recursive subspace prediction algorithm for the time-varying system matrices of the linear systems has been proposed. The basic idea of our approach is to apply the recursive 4SID algorithm to the subspace prediction algorithm, and this lightens the computational burden accompanying the SVD. The efficiency was confirmed by simulation experiments.

References


Fig. 4. Sample mean (top) and variance (bottom) of the imaginary part of predicted conjugate poles $(N = 250, L = 50 \text{ (50-step ahead prediction)})$. 

Fig. 5. Time evolutions of real (top) and imaginary (bottom) parts of predicted conjugate poles $(N=100, L = 50 \text{ (50-step ahead prediction)})$. 

Fig. 6. Sample mean (top) and variance (bottom) of the real part of predicted conjugate poles $(N = 100, L = 50 \text{ (50-step ahead prediction)})$. 

Fig. 7. Sample mean (top) and variance (bottom) of the imaginary part of predicted conjugate poles $(N=250, L = 50 \text{ (50-step ahead prediction)})$. 

---

100
Appendix: Condition for Derivation of the Principal Vectors

By the assumption that the noise sequences are mutually uncorrelated with input sequence, the subspace spanned by $L_b(k_2)$ is represented as

$$\text{span}_{\text{col}}\{L_b(k_2)\} = \text{span}_{\text{col}}\{\Gamma_{\ell}(k_2)\}$$

$$+ \text{span}_{\text{col}}\{E(k_2)\},$$

(23)

where $\oplus$ denotes the direct sum; and $\text{span}_{\text{col}}\{E(k_2)\}$ is the noise subspace. So, each column vector of $L_b(k_2)$ is represented as:

$$h_{ij}(k_2) = g_{ij}(k_2)v_{i}(k_2) + \cdots + g_{nj}(k_2)v_{n}(k_2) + g_{n+1,j}(k_2)v_{n+1}(k_2) + \cdots + g_{\alpha \ell,j}(k_2)v_{\alpha \ell}(k_2),$$

(24)

where $v_{i}(k_2) \in \text{span}_{\text{col}}\{\Gamma_{\ell}(k_2)\}$ ($i = 1, \ldots, n$) and $v_{i}(k_2) \in \text{span}_{\text{col}}\{E(k_2)\}$ ($i = n+1, \ldots, \alpha \ell$) are principal vectors of the signal and noise subspaces, respectively (Fig. 8); and $g_{ij}(k_2)$ ($i, j = 1, \ldots, \alpha \ell$) are appropriate coefficients for the basis $v_{i}(k_2)$ ($i = 1, \ldots, \alpha \ell$) in this representation. Then, for the angle between ith principal vector of the signal subspace and jth one of the noise subspace, $\phi_{ij}(k_2)$, holds the relation:

$$\phi_{ij}(k_2) = \frac{\pi}{2} - \theta_{ij}(k_2)$$

(25)

where $\theta_{ij}(k_2)$ is represented by (23) as

$$\theta_{ij}(k_2) = \arccos\left(\frac{\text{span}_{\text{col}}\{\Gamma_{\ell}(k_2)\} \cdot \text{span}_{\text{col}}\{E(k_2)\}}{\text{span}_{\text{col}}\{\Gamma_{\ell}(k_2)\} \cdot \text{span}_{\text{col}}\{E(k_2)\}}\right).$$

On the other hand, the SVD in (22) yields principal vectors of signal subspace at time $k_2$ from $\text{span}_{\text{col}}\{L_b(k_2)\}$ so that the angle between principal vectors of $\text{span}_{\text{col}}\{L_b(k_2)\}$ and $\text{span}_{\text{col}}\{\Gamma_{\ell}(k_1)\}$ becomes minimum.

So, all $\theta_{ij}(k_2)$ have to be smaller than $\phi_{ij}(k_2)$ to derive principal vectors of $\Gamma_{\ell}(k_2)$ as the first $n$ column vectors of $\Gamma_{\ell}(k_2)$, i.e.,

$$\phi_{ij}(k_2) = \frac{\pi}{2} - \theta_{ij}(k_2) \iff \theta_{ij}(k_2) < \frac{\pi}{4}.$$