A Generalized Class of Pseudomeasurements for Identifying Unknown Parameters of Linear Stochastic Systems

Akio Tanikawa

Faculty of Information Science and Technology,
Osaka Institute of Technology
Kitayama, Hirakata-shi, 573-0196, Japan
E-mail: tanikawa@is.oit.ac.jp

Abstract

A new class of pseudomeasurements for discrete-time linear stochastic systems are derived from continuous-time linear stochastic systems with unknown parameters by applying time-discretization and Taylor expansion. Utilizing these pseudomeasurements, we propose new iterative methods which estimate the states of the discrete-time stochastic systems and identify the unknown parameters simultaneously. A numerical example is given to show the effectiveness of the proposed approach.

Keywords: system identification, pseudomeasurement, extended Kalman filter.

1 Introduction

Identification of stochastic dynamical systems with unknown parameters from input and output data is one of the most fundamental and important problems in systems engineering and has been discussed by many researchers. Since late seventies the typical methods which attracted much attention of the researchers and had good reputations have been the maximum likelihood estimation method (see [2], [3] and [10]) and the subspace-based identification method (see [12] and [24]). In the former approach, estimates of unknown parameters are performed by maximizing such criteria as likelihood or a posteriori function. But the difficulty lies on how to preassign each region of the unknown parameters over which the maximization is taken. As for the latter approach, estimates of the system matrices can be obtained from the input and output data, but it should be noted that the identification by this method is performed within a similarity transformation.

The main purpose of this paper is to introduce a new class of pseudomeasurements which are additional observation processes on the unknown parameters of the stochastic dynamical systems. The pseudomeasurements were utilized originally for various tracking problems by several researchers (see e.g., [1], [19], [20] and [25]). Later, Ohsumi and his colleagues applied the idea of pseudomeasurement to various practical problems ([8], [9], [11], [13], [14], [15], [16], [17] and [18]). Recently, Kameyama and Ohsumi applied it to the identification problem of unknown parameters included in the system matrix $A$ for continuous-time linear stochastic systems ([5] and [6]) (see Section 2 for the definition of the system matrix $A$).

In this paper, the identification problem of the same kind is considered and the pseudomeasurement approach for discrete-time linear stochastic systems are presented. Let the known part (resp. the unknown part) of $A$ be denoted by $A_{(k)}$ (resp. $A_{(u)}$) (see Section 3 for the details). We will propose the iterative method which identifies both the states of the system and the unknown part $A_{(u)}$ of $A$ by using $A_{(k)}$ explicitly. This is one of the advantages of the the proposed method in this paper over the subspace-based method since it is impossible to utilize the information about the known part $A_{(k)}$ of $A$ explicitly by the subspace-based method. The first identification method by pseudomeasurements for discrete-time stochastic systems was proposed in [22] for the case where the dimension of the observation process is the same as that of the dynamical system. Later, two methods were introduced respectively in [21] and [23] for the case where the dimension of the observation process is smaller than that of the dynamical system (i.e., partially observed case).

In this paper, we start from the continuous-time linear stochastic systems with unknown parameters given in Section 2. In Section 3, we derive the new discrete-time stochastic system with unknown parameters from the continuous-time system by applying time-discretization and Taylor expansion. After introducing an augmented system, we introduce a new class of pseudomeasurements for the discrete-time stochastic systems in Section 4. Applying the extended Kalman filter to the augmented system, we derive the new identification method in Section 5 which estimates the states of the discrete-time stochastic system and identify the unknown parameters simultaneously. After a remark to the identification method in Section 6, we give a numer-
ical example to show the effectiveness of the proposed approach in Section 7.

2 Continuous-time System

Consider the continuous-time linear stochastic system for \( t \geq 0 \):

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Cu(t) + G\gamma(t), \quad x(0) = \bar{x}, \quad (1) \\
y(t) &= Hx(t) + Sv(t), \quad (2)
\end{align*}
\]

where \( x(t) \in \mathbb{R}^n \), \( y(t) \in \mathbb{R}^n \) and \( u(t) \in \mathbb{R}^l \) denote the state vector of the system, the observation vector and the input vector, respectively. Here, \( A, C, G, H \) and \( S \) are real constant matrices with appropriate sizes, i.e., \( A \in \mathbb{R}^{n \times n} \), \( C \in \mathbb{R}^{n \times l} \), \( G \in \mathbb{R}^{n \times d_1} \), \( H \in \mathbb{R}^{n \times n} \) and \( S \in \mathbb{R}^{n \times d_2} \). We assume that \( H \) is an invertible matrix. The functions \( \gamma(t) \) and \( v(t) \) are independent random processes (white noise) with identically zero means and covariance matrices

\[
\begin{align*}
\text{cov} [\gamma(t), \gamma(\tau)] &= Q \cdot \delta(t-\tau), \\
\text{cov} [v(t), v(\tau)] &= R \cdot \delta(t-\tau), \\
\text{cov} [\gamma(t), v(\tau)] &= O,
\end{align*}
\]

for all \( t \) and \( \tau \), where \( \delta \) is the Dirac delta function, \( T \) denotes transposition of a matrix, \( Q \) and \( R \) are symmetric, nonnegative definite matrices. Moreover, we suppose that \( R \) is positive definite. These are standard assumptions (see e.g. [4]).

We assume that \( C, G, H \) and \( S \) are known but \( A \) is partially unknown. We decompose \( A \) into the two parts, the known part \( A_{(u)} \in \mathbb{R}^{n \times n} \) and the unknown part \( A_{(u)} \in \mathbb{R}^{n \times n} \), where all the known (resp. unknown) elements are included in \( A_{(u)} \) (resp. \( A_{(u)} \)) and the equality \( A = A_{(k)} + A_{(u)} \) holds. Let \( p \) be the number of unknown parameters of \( A \) and \( \theta \) be the vector consisting of these parameters : \( \theta = [\theta_1, \theta_2, \ldots, \theta_p]^T \). We also write the matrix \( A_{(u)} \) as \( A_{(u)}(\theta) \) to indicate the unknown parameter vector \( \theta \) explicitly.

3 Discrete-time System

From (1)-(2), we can obtain the canonical discrete-time system \( \{x_i\} \) with \( x_i = x(i \times \Delta t) \) given by

\[
\begin{align*}
x_{i+1} &= A^{(d)} x_i + C^{(d)} u_i + G^{(d)} w_i, \quad x_0 = \bar{x}, \quad (3) \\
y_i &= H x_i + S v_i, \quad (4)
\end{align*}
\]

where \( \Delta t \) is a small time increment. Here, the matrices \( A^{(d)} \) and \( C^{(d)} \) have the following explicit forms respectively

\[
\begin{align*}
A^{(d)} &= e^{(\Delta t)A}, \\
C^{(d)} &= \int_0^{\Delta t} e^{\tau A} d\tau \cdot C,
\end{align*}
\]

\( A^{(d)} \) is an \( n \times d_1 \) matrix, and \( \{w_i\} \) and \( \{w_i\} \) are the mutually independent i. i. d. random noise sequences with zero mean and covariance matrices

\[
\begin{align*}
\text{cov} [w_i, w_j] &= Q \cdot \delta(i - j), \\
\text{cov} [v_i, v_j] &= R \cdot \delta(i - j).
\end{align*}
\]

Applying Taylor expansion

\[
\begin{align*}
A^{(d)} &= I + (\Delta t) A + \frac{(\Delta t)^2}{2} A^2 + o((\Delta t)^2), \\
C^{(d)} &= \int_0^{\Delta t} \left(I + \tau A + \frac{\tau^2}{2} A^2\right) d\tau \cdot C + o((\Delta t)^2)
\end{align*}
\]

\[
\begin{align*}
&= \left\{(\Delta t) I + \frac{(\Delta t)^2}{2} A\right\} C + o((\Delta t)^2)
\end{align*}
\]

to (5) and (6) respectively (noting that the term with the notation \( o((\Delta t)^2) \) converges to 0 at the rate faster than \( (\Delta t)^2 \) as \( \Delta t \to 0 \)), we define

\[
\begin{align*}
\tilde{A} &\triangleq I + (\Delta t) A + \frac{(\Delta t)^2}{2} A^2, \quad (7) \\
\tilde{C} &\triangleq \left\{(\Delta t) I + \frac{(\Delta t)^2}{2} A\right\} C. \quad (8)
\end{align*}
\]

Thus, we have the discrete-time linear stochastic system

\[
x_{i+1} = \tilde{A} x_i + \tilde{C} u_i + \tilde{G} w_i, \quad x_0 = \bar{x} \quad (9)
\]

which is an approximation to (3), where we put \( \tilde{G} = G^{(d)} \).

From now on, we use the notation \( x(i) \) instead of \( x_i \) for convenience of notations. In other words, we drop \( \Delta t \) from \( x(i(\Delta t)) \). This notational change will not cause confusion since we always use the notation \( i \) in \( x(i) \) in this sense and do not deal with the continuous-time system \( \{x(t)\} \) directly any more except for arguments on the unknown parameters of the matrix \( A \). We also use the similar notations \( y(i), u(i), v(i) \) and \( w(i) \) instead of \( y_i, u_i, v_i \) and \( w_i \) respectively. Thus we consider the following discrete-time stochastic process:

\[
\begin{align*}
x(i + 1) &= \tilde{A} x(i) + \tilde{C} u(i) + \tilde{G} w(i), \quad x(0) = \bar{x}, \quad (10) \\
y(i) &= H x(i) + S v(i), \quad (11)
\end{align*}
\]

where the pair \( (A, H) \) is assumed tacitly to be observable.

Similarly to the matrix \( A \), we have a decomposition of \( \tilde{A} \)

\[
\tilde{A} = \tilde{A}_{(k)} + \tilde{A}_{(u)}(\theta), \quad (12)
\]

where \( \tilde{A}_{(k)} \) (resp. \( \tilde{A}_{(u)}(\theta) \)) denotes the known (resp. unknown) part of \( \tilde{A} \). Actually, we have

\[
\begin{align*}
\tilde{A} &= I + (\Delta t) A_{(k)} + A_{(u)} + \frac{(\Delta t)^2}{2} \left(A_{(k)}^2 + A_{(u)}^2\right) \\
&= \left\{I + (\Delta t) A_{(k)} + \frac{(\Delta t)^2}{2} A_{(k)}^2\right\} \\
&\quad + \left\{(\Delta t) A_{(u)} + \frac{(\Delta t)^2}{2} A_{(u)}^2\right\} \\
&= \tilde{A}_{(k)} + \tilde{A}_{(u)}(\theta).
\end{align*}
\]
We similarly have a decomposition of $\tilde{C}$
\[
\tilde{C} = \tilde{C}_{(k)} + \tilde{C}_{(u)}(\theta),
\]
where $\tilde{C}_{(k)}$ (resp. $\tilde{C}_{(u)}(\theta)$) denotes the known (resp. unknown) part of $\tilde{C}$. Here, $\tilde{C}_{(k)}$ and $\tilde{C}_{(u)}(\theta)$ have the following explicit forms:
\[
\begin{align*}
\tilde{C} &= \begin{cases} (\Delta t)I + \frac{(\Delta t)^2}{2}A(k) \end{cases} C \\
&= \begin{cases} (\Delta t)I + \frac{(\Delta t)^2}{2}A(k) \end{cases} C + \frac{(\Delta t)^2}{2}A(u) C \\
&= \tilde{C}_{(k)} + \tilde{C}_{(u)}(\theta).
\end{align*}
\]
(15)

For a vector $r \in \mathbb{R}^n$, we denote its $j$-th coordinate by $r_j$ ($j = 1, 2, \cdots, n$) and write $r = [r_1, r_2, \cdots, r_n]^T$. We now investigate the unknown matrices $\tilde{A}_{(u)}(\theta)$ and $\tilde{C}_{(u)}(\theta)$ in detail. First of all, we can easily find a matrix $X^{(0)}(x(i)) \in \mathbb{R}^{n \times p}$ such that
\[
\tilde{A}_{(u)}(\theta) x(i) = X^{(0)}(x(i)) \theta
\]
holds. Actually, $X^{(0)}(x(i))$ has the form
\[
X^{(0)}(x(i)) = M_0 (x(i) \otimes I_p),
\]
(17)
where the symbol $\otimes$ denotes the Kronecker’s product of matrices, $I_p$ denotes the $p \times p$ identity matrix and $M_0$ is an $n \times np$ matrix whose elements are 0 or 1 (see [6] for the details).

Since $\tilde{A}$ contains not only $A$ but $A^2$, we note that $\tilde{A}_{(u)}(\theta)$ has the form
\[
\tilde{A}_{(u)}(\theta) = \tilde{A}_{(u,1)}(\theta) + \tilde{A}_{(u,2)}(\theta),
\]
(18)
where each element of $\tilde{A}_{(u,1)}(\theta)$ is a linear function of $\theta$ and each element of $\tilde{A}_{(u,2)}(\theta)$ is a quadratic function of $\theta$. Then, similarly to (16) and (17), we can find matrices $M_{(u,1)} \in \mathbb{R}^{n \times np}$ and $X^{(1)}(x(i)) \in \mathbb{R}^{n \times p}$ such that
\[
\tilde{A}_{(u,1)}(\theta) x(i) = M_{(u,1)} (x(i) \otimes I_p) \theta = X^{(1)}(x(i)) \theta
\]
holds. But, on the matrix $\tilde{A}_{(u,2)}(\theta)$, we have the expression
\[
\tilde{A}_{(u,2)}(\theta) x(i) = [r_1(x(i)), r_2(x(i)), \cdots, r_n(x(i))]^T = r(x(i)),
\]
(20)
where $r_j(x(i))$ can be written as
\[
r_j(x(i)) = \frac{1}{2} \theta^T D^{(j)}(x(i)) \theta
\]
(21)
with a symmetric matrix $D^{(j)}(x(i)) \in \mathbb{R}^{p \times p}$ depending on $x(i)$ for $j = 1, 2, \cdots, n$.

We now try to find a relation between the unknown vector $\theta$ and the observation process $\{y(i)\}$. Since $H$ is invertible, we have
\[
x(i) = H^{-1} y(i) - H^{-1} S v(i)
\]
(22)
from (11). Substituting this equality to (10), we have
\[
\begin{align*}
H^{-1} y(i + 1) &= x(i + 1) + H^{-1} S v(i + 1) \\
&= \tilde{A}_k x(i) + \tilde{A}_u x(i) + \tilde{C}_k u(i) + \tilde{C}_u u(i) + G w(i) + H^{-1} S v(i + 1) \\
&= \tilde{A}_k H^{-1} y(i) + \tilde{A}_u H^{-1} y(i) + \tilde{C}_k u(i) + \tilde{C}_u u(i) + G w(i) - \tilde{A}_k H^{-1} S v(i) - \tilde{A}_u H^{-1} S v(i) + H^{-1} S v(i + 1).
\end{align*}
\]
(23)

Defining $y^{(b)}(i)$ by
\[
y^{(b)}(i) \triangleq H^{-1} y(i + 1) - \tilde{A}_k H^{-1} y(i) - \tilde{C}_k u(i),
\]
(24)
we obtain
\[
\begin{align*}
y^{(b)}(i) &= \tilde{A}_{(u,1)}(\theta) H^{-1} y(i) + \tilde{C}_{(u)}(\theta) u(i) \\
&= \tilde{A}_{(u,2)}(\theta) H^{-1} y(i) + \text{(noise terms)}.
\end{align*}
\]
(25)
from (23). The first term of the right hand side of (25) can be written by
\[
\tilde{A}_{(u,1)}(\theta) H^{-1} y(i) = M_{(u,1)}(\theta) \left[(H^{-1} y(i)) \otimes I_p\right] \theta
\]
(26)
due to (19), and the second term can be written by
\[
\tilde{C}_{(u)}(\theta) u(i) = \frac{(\Delta t)^2}{2} M_0 [(C u(i)) \otimes I_p] \theta
\]
(27)
due to (15), (16) and (17). Hence, using the matrix $\tilde{H}^{(b)}(i)$ defined by
\[
\tilde{H}^{(b)}(i) \triangleq \tilde{A}_{(u,1)}(\theta) [(H^{-1} y(i)) \otimes I_p] \\
+ \frac{(\Delta t)^2}{2} M_0 [(C u(i)) \otimes I_p],
\]
(28)
we have
\[
y^{(b)}(i) = \tilde{H}^{(b)}(i) \theta + \tilde{A}_{(u,2)}(\theta) H^{-1} y(i) \\
+ \text{(noise terms)},
\]
(29)
where
\[
\tilde{A}_{(u,2)}(\theta) H^{-1} y(i) = [r_1(H^{-1} y(i)), r_2(H^{-1} y(i)), \cdots, r_n(H^{-1} y(i))]^T = r(H^{-1} y(i)).
\]
(30)

Note that $r_j(\cdot)$ was defined by (21).

## 4 Augmented System

For identifying the unknown (constant) vector $\theta$, we usually consider it as a function $\theta(i)$ of $i$ and moreover treat it as a random vector-valued process
\[
\theta(i + 1) = \theta(i) + G^{(b)} w^{(b)}(i)
\]
(31)
by allowing some ambiguity $w^{(b)}(i)$. Here, $\{w^{(b)}(i) \in \mathbb{R}^d\}$ is a newly introduced zero-mean...
Gaussian white noise sequence with covariance matrix $Q^{(b)}$ which is independent with the previously given random sequences. We note that the matrices $G^{(b)}$ and $Q^{(b)}$ can be preassigned by the users. Actually, we may choose $G^{(b)} = O$ if we need no ambiguity in (31).

Introducing the new state vector defined by $z(i) := [x(i)^T, \theta(i)^T]^T$, we obtain the following augmented system from (10) and (31):

$$z(i + 1) = f(z(i)) + C^{(a)}u(i) + G^{(a)}w^{(a)}(i), \quad (32)$$

where

$$f(z(i)) = \{f_1(z(i)), f_2(z(i)), \ldots, f_{n+p}(z(i))\}^T = \begin{bmatrix} \tilde{A} & O \\ O & I \end{bmatrix} z(i) \in \mathbb{R}^{n+p}$$

$$= \begin{bmatrix} \tilde{A}(i) & X^{(1)}(x(i)) \\ O & I \end{bmatrix} z(i) + \begin{bmatrix} \tilde{A}(u, 2) & O \end{bmatrix} z(i)$$

$$= \begin{bmatrix} \tilde{A}(i) & X^{(1)}(x(i)) \\ O & I \end{bmatrix} z(i) + \begin{bmatrix} r(x(i)) \end{bmatrix} ,$$

$$C^{(a)} = \begin{bmatrix} C \\ O \end{bmatrix} \in \mathbb{R}^{(n+p)\times \ell},$$

$$G^{(a)} = \begin{bmatrix} G & O \\ O & G^{(b)} \end{bmatrix} \in \mathbb{R}^{(n+p)\times (d_1 + d_2)} ,$$

$$w^{(a)}(i) = \begin{bmatrix} w(i) \\ w^{(b)}(i) \end{bmatrix} \in \mathbb{R}^{d_1 + d_2} .$$

Since the noise terms in (29) are a linear combination of the mutually independent Gaussian white noises $w(i), v(i)$ and $v(i + 1)$ and $\theta$ is a constant vector, they can be written as a single Gaussian noise. Thus, the equality (29) can be rewritten as

$$y^{(b)}(i) = \tilde{H}^{(b)}(i) \theta(i) + r \{H^{-1}y(i)\} + S^{(b)}v^{(b)}(i) \quad (33)$$

with $S^{(b)} \in \mathbb{R}^{n \times d_4}$, where $\{y^{(b)}(i) \in \mathbb{R}^{d_4}\}$ is a zero-mean Gaussian (but non-white) noise sequence with a covariance matrix $R^{(b)}$ which has a correlation with the previously given Gaussian white noise sequences $\{v(i)\}$ and $\{w(i)\}$. In view of (33), we can regard $\{y^{(b)}(i)\}$ as the observation process of the unknown vector $\theta$. We call $y^{(b)}(i)$ pseudomeasurement of $\theta$.

Introducing the new $(m + n)$-dimensional vectors $y^{(a)}(i) := [y(i)^T, y^{(b)}(i)^T]^T$ and the new $(d_2 + d_4)$-dimensional vectors $v^{(a)}(i) := [v(i)^T, v^{(b)}(i)^T]^T$, we obtain the following augmented observation process:

$$y^{(a)}(i) = \begin{bmatrix} H^{(a)}(i) z(i) + r^{(a)}(i) + S^{(a)}v^{(a)}(i) \\ h(i)(z(i)) + S^{(a)}v^{(a)}(i) \end{bmatrix}$$

$$\triangleq h^{(a)}(i)z(i) + S^{(a)}v^{(a)}(i), \quad (34)$$

from (11) and (33), where the vector $r^{(a)}(i) \in \mathbb{R}^{2n}$ is defined by

$$r^{(a)}(i) = \begin{bmatrix} 0_n^T \\ r(H^{-1}y(i))^T \end{bmatrix} ,$$

where $0_n$ is the $n$-dimensional zero vector.

Moreover, the matrices $H^{(a)}(i)$ and $S^{(a)}$ are defined respectively by

$$H^{(a)}(i) = \begin{bmatrix} H & O \\ O & \tilde{H}^{(b)}(i) \end{bmatrix} \in \mathbb{R}^{(m+n)\times (n+p)} ,$$

$$S^{(a)} = \begin{bmatrix} S & O \\ O & S^{(b)} \end{bmatrix} \in \mathbb{R}^{(m+n)\times (d_2 + d_4)} .$$

5 Extended Kalman Filter

Let us use the notation $Y_i$ for the $\sigma$-field $\{y(j), 0 \leq j \leq i\}$ generated by the observation data $y(j)$ up to the time $i$. Let $\tilde{z}(ii|i)\tilde{z}(ii|i) - 1$ be the least mean square estimates of the state $z(i)$ respectively with respect to $Y_i$ and $Y_{i-1}$. We define the matrices $P(i|i)$ and $P(i + 1|i)$ as follows:

$$P(i|i) = \mathcal{E} \{ (z(i) - \tilde{z}(i|i)) (z(i) - \tilde{z}(i|i))^T \} \quad (35)$$

and

$$P(i + 1|i) = \mathcal{E} \{ (z(i + 1) - \tilde{z}(i + 1|i)) (z(i + 1) - \tilde{z}(i + 1|i))^T \} , \quad (36)$$

where $\mathcal{E}\{ \cdot \}$ denotes the mathematical expectation operator.

Supposing that the state estimate $\tilde{z}(i|i)$ of $z(i)$ has been obtained, we apply the linear approximation to the nonlinear function $f$ around $\tilde{z}(i|i)$ and have

$$f(z(i)) = f(\tilde{z}(i|i)) + \tilde{F}(i)(z(i) - \tilde{z}(i|i)) + \text{(higher order terms)},$$

where

$$\tilde{F}(i) = \begin{bmatrix} \frac{\partial f(z)}{\partial z} \end{bmatrix}_{z = \tilde{z}(i|i)} = \begin{bmatrix} \frac{\partial f^T(z)}{\partial z} \end{bmatrix}_{z = \tilde{z}(i|i)} .$$

Then, we obtain the linearized system of (32)

$$z(i + 1) = \tilde{F}(i)z(i) + C^{(a)}u(i) + G^{(a)}w^{(a)}(i)$$

$$+ f(\tilde{z}(i|i)) - \tilde{F}(i)\tilde{z}(i|i). \quad (37)$$

Although the system equation (37) does not satisfy all the standard conditions for applying the Kalman filter, we formally apply the Kalman filter to it and have the following results.

**Proposition 5.1** The optimal estimates $\tilde{z}(i + 1|i)$ and $\tilde{z}(i + 1|i + 1)$ obtained by the Kalman filter applied to the linearized system (37) with the observation process (34) have the following form:

$$\tilde{z}(i + 1|i) = f(\tilde{z}(i|i)) + C^{(a)}u(i) , \quad (38)$$

$$\tilde{z}(i + 1|i + 1) = \tilde{z}(i + 1|i)$$

$$+ K(i + 1) \left[ y^{(a)}(i + 1) - H^{(a)}(i + 1)\tilde{z}(i + 1|i) \right] . \quad (39)$$
where
\[
K(i + 1) = P(i + 1|i) H^T(i) P(i + 1|i) H^T(i) + S^T(i) R(i) S(i)\]
\[
\times \left[ H^T(i) P(i + 1|i) H^T(i) + S^T(i) R(i) S(i) \right]^{-1}.
\]
(40)

Here, \( \hat{z}(0|1) = \bar{z}_o \) and \( P(0|1) \) are the initial estimates. Moreover, the matrices \( P(i + 1|i) \) and \( P(i|i) \) satisfy the recursive relations
\[
P(i + 1|i) = \hat{F}(i) P(i|i) \hat{F}(i) + C(i) Q(i) G(i)^T\]
(41)
and
\[
P(i|i) = P(i|i) - K(i) H^T(i) P(i|i)\]
(42)
respectively, where
\[
Q(i) = \begin{bmatrix} Q & 0 \\ 0 & Q^{(b)} \end{bmatrix} \in \mathbb{R}^{(d_1+d_3) \times (d_1+d_3)}.
\]

(Proof) Applying the Kalman filter to the linear system
(37) with (34), we have
\[
\hat{z}(i + 1|i) = \hat{F}(i) \hat{z}(i|i) + C(i) u(i) + f(\hat{z}(i|i)) - \hat{F}(i) \hat{z}(i|i)
\]
and
\[
\hat{z}(i + 1|i + 1) = \hat{z}(i + 1|i) + K(i + 1)[y(i) - H(i) \hat{z}(i + 1|i)].
\]
The rest of the results can be obtained by the standard argument on the Kalman filter. \( \quad \square \)

Proposition 5.2 The equalities (38) and (39) can be written explicitly as follows:
\[
[ \hat{x}(i + 1|i) \quad \hat{\theta}(i + 1|i) ] = \begin{bmatrix} \tilde{A}(k) + \tilde{A}(u)(\hat{\theta}(i|i)) & O \\ O & I \end{bmatrix} \begin{bmatrix} \hat{z}(i|i) \\ \hat{\theta}(i|i) \end{bmatrix} + \begin{bmatrix} \tilde{C}(k) + \tilde{C}(u)(\hat{\theta}(i|i)) \\ O \end{bmatrix} u(i)
\]
(43)
and
\[
[ \hat{z}(i + 1|i + 1) \quad \hat{\theta}(i + 1|i + 1) ] = \begin{bmatrix} \hat{z}(i + 1|i) \\ \hat{\theta}(i + 1|i) \end{bmatrix} + K(i + 1)
\times \begin{bmatrix} y(i + 1) - H \hat{z}(i + 1|i) \\ y(i + 1) - H(\hat{\theta}(i + 1|i)) \hat{z}(i + 1|i) \end{bmatrix} - \begin{bmatrix} 0 \\ \tilde{A}(u,2)(\hat{\theta}(i + 1|i)) \end{bmatrix} H^{-1} y(i + 1) \]
(44)

Remark 5.3 We note that
\[
\frac{\partial f^T(z)}{\partial z} = \frac{\partial}{\partial z} \left[ z^T \{ \tilde{A}(k) + \tilde{A}(u)(\theta) \} \theta^T \right] = \frac{\partial}{\partial \theta} \left[ z^T \{ \tilde{A}(k) + \tilde{A}(u)(\theta) \} \theta^T \right]
\]
holds due to the equality
\[
\frac{\partial}{\partial \theta} x^T \tilde{A}(u)(\theta) = \frac{\partial}{\partial \theta} \{ x^T \tilde{A}(u)(\theta) \} T = \{ x^T \tilde{A}(u)(\theta) \} T + \{ \tilde{A}(u) \theta \} T.
\]

Thus we have
\[
\hat{F}(i) = \left( \frac{\partial f^T(z)}{\partial z} \right)_{z=\hat{z}(i|i)} = \begin{bmatrix} \{ \tilde{A}(k) + \tilde{A}(u)(\theta) \} T & O \\ \{ {\hat{X}}(1)(x) + {\hat{X}}(2)(x) \} T & I \end{bmatrix}_{x=\hat{z}(i|i), \theta=\hat{\theta}(i|i)}
\]
(45)
We similarly have
\[
\frac{\partial h(z)}{\partial z} = \begin{bmatrix} H & O \\ O & \hat{H}(b)(i) + Y(y(i)) \end{bmatrix}_{z=\hat{z}(i|i), \theta=\hat{\theta}(i|i)}
\]
(46)
where
\[
Y(y) = \begin{bmatrix} D(1)(H^{-1}) y(1) \theta, D(2)(H^{-1}) y(2), \cdots, D(n)(H^{-1}) y(n) \end{bmatrix} T.
\]

6. A Remark to the Identification Method in Section 5

In this section, we put \( S^{(b)} = r I \) for \( r > 0 \) and \( R^{(b)} = I \) to simplify the following statement. We show in the next proposition that the optimal estimate \( \hat{z}(i + 1|i + 1) \) obtained in Proposition 5.1 converges to the optimal estimate given by the conventional identification method without pseudomeasurement when \( r \to \infty \). The proof is omitted here.

Proposition 6.1 We assume that \( S > 0 \) and \( R > 0 \) hold and that the observations \{ \( y(j) : j = 0, 1, 2, \cdots \) \} are uniformly bounded, i.e., there exist \( L > 0 \) and \( r_0 > 0 \) such that
\[
|y(j)| \leq L
\]
for any \( j \geq 0 \) and any \( r \geq r_0 \). Then, when \( r \to \infty \), we obtain that the gain matrix \( K(i+1) \) given by (40) in Proposition 5.1 converges to the gain matrix of the conventional identification (and state estimation) method without pseudomeasurement. Hence, the optimal estimate \( \tilde{z}(i+1|i+1) \) obtained in Proposition 5.1 converges to the optimal estimate given by the conventional identification method without pseudomeasurement.

### 7 Numerical Simulation

Consider a mass-spring-dashpot system subject to random noise

\[
m \ddot{x}(t) + d_0 \dot{x}(t) + k \dot{x}(t) = c u(t) + g \gamma(t),
\]

(47)

where \( \gamma(t) \) is the displacement from the equilibrium state; \( u(t) \) is the known input; \( \gamma(t) \) is the random disturbance; and \( m, k, d_0 \) are mass, stiffness and damping coefficients respectively.

The state space model of (47) is given by

\[
\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -k/m & -d_0/m \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ c/m \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ g/m \end{bmatrix} \gamma(t), \quad x(0) = \bar{x},
\]

where \( x(t) = [\xi(t), \dot{\xi}(t)]^T \),

\[
A = \begin{bmatrix} 0 & 1 \\ -k/m & -d_0/m \end{bmatrix}, \quad C = \begin{bmatrix} 0 \\ c/m \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ g/m \end{bmatrix}.
\]

Let \( m \) be 1 for simplicity. We assume that \( k \) is known but \( d_0 \) is unknown and so we put \( \theta = d_0 \). Then we have

\[
A = \begin{bmatrix} 0 & 1 \\ -k & -\theta \end{bmatrix}, \quad C = \begin{bmatrix} 0 \\ c \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ g \end{bmatrix}
\]

and hence

\[
A_{(k)} = \begin{bmatrix} 0 & 1 \\ -k & 0 \end{bmatrix}, \quad A_{(\theta)} = \begin{bmatrix} 0 & 0 \\ 0 & -\theta \end{bmatrix}.
\]

By applying time-discretization and Taylor expansion to (47), we have

\[
x(i+1) = \tilde{A} x(i) + \tilde{C} u(i) + \tilde{G} w(i), \quad x(0) = \bar{x},
\]

(51)

where \( \{w_i\} \) is an i. i. d. random noise sequence with zero mean,

\[
\tilde{A} = I + (\Delta t) A + \left( \frac{(\Delta t)^2}{2} A^2 \right)
= \begin{bmatrix} 1 - k (\Delta t)^2/2 & \Delta t \\ -k (\Delta t) & 1 - k (\Delta t)^2/2 \end{bmatrix}
+ \begin{bmatrix} k \theta / 2 (\Delta t)^2 & - \theta (\Delta t) + \theta^2 / 2 (\Delta t)^2 \\ - \theta (\Delta t) + \theta^2 / 2 (\Delta t)^2 \end{bmatrix}
= \tilde{A}_{(k)} + \tilde{A}_{(\theta)}(\theta)
\]

(52)

\[
\tilde{C} = \left( (\Delta t) I + \left( \frac{(\Delta t)^2}{2} A \right) \right) C
= \begin{bmatrix} \Delta t & \frac{(\Delta t)^2}{2} \\ -k (\Delta t)^2/2 & \Delta t \end{bmatrix}
C + \begin{bmatrix} 0 & 0 \\ 0 & -\theta (\Delta t)^2/2 \end{bmatrix} C
\equiv \tilde{C}_{(k)} + \tilde{C}_{(\theta)}(\theta),
\]

(53)

and \( \tilde{G} \) can be determined from \( G \) similarly to \( \tilde{C} \) but it is omitted here. Here, \( (\Delta t) \) is a small time increment.

For the discrete-time state space model (51) with the matrices given by (52) and (53), we suppose that the observation process

\[
y(i) = H x(i) + S v(i)
\]

(54)

is available, where the noise sequences \( \{w(i)\} \) and \( \{v(i)\} \) are mutually independent white Gaussian noises with zero means. The parameters concerning the system structure are set as \( H = S = I_2 \), \( Q = I_2 \), \( R = 0.01 I_2 \) and \( c = 1/4 \).

In our simulation, we chose \( k = 2 \) and \( \theta = 1 \). The inputs \( \{u(i)\} \) were selected from white Gaussian noises with unit variances as identification input. We performed simulation of the proposed identification method from the initial state \( x(0) = [3, 3]^T \) for \( 25 \) fixed pseudo-random numbers. The identification method was started from the initial estimates; \( \tilde{x}(0) = [0, 0]^T \), \( \tilde{\theta}(0) - 1 = 5 \), \( P(0) - 1 = I_3 \). User-defined parameters were set as \( G^{(b)} = I_1 \), \( S^{(b)} = I_2 \), and \( Q^{(b)} = 0 \), whereas \( R^{(b)} \) was set as the four different forms (see the following table for the detail). For each fixed pseudo-random numbers, the average values of the errors \( \| \tilde{\theta}(i) - \theta \| \) of the unknown parameter \( \theta \) for the iterations between 401 and 1000 by the identification method with several choices of the matrix \( R^{(b)} \) are recorded and the best choice of \( R^{(b)} \) is determined.

And the number of simulations is counted where its choice was accepted to be the best one for each \( R^{(b)} \). The following table contains these numbers.

<table>
<thead>
<tr>
<th>( R^{(b)} )</th>
<th># of best choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10 \times I_2 )</td>
<td>1</td>
</tr>
<tr>
<td>( I_2 )</td>
<td>3</td>
</tr>
<tr>
<td>( 0.1 \times I_2 )</td>
<td>2</td>
</tr>
<tr>
<td>( 0.01 \times I_2 )</td>
<td>0</td>
</tr>
<tr>
<td>( 0.001 \times I_2 )</td>
<td>19</td>
</tr>
<tr>
<td>no pseudo.</td>
<td>6</td>
</tr>
</tbody>
</table>

The table also contains the result of the conventional method (without pseudomeasurement) for comparison. From the table, we can observe the effectiveness of the proposed approach.

### 8 Conclusions

We derived a new class of pseudomeasurements for discrete-time stochastic systems from continuous-time
linear stochastic systems with unknown parameters by applying time-discretization and Taylor expansion. Utilizing these pseudomeasurements, we proposed the new iterative methods which estimate the states of the discrete-time stochastic systems and identify the unknown parameters simultaneously. These identification methods can be expected to be very useful for the discrete-time systems obtained from the originally continuous-time stochastic systems. A numerical example is given to show the effectiveness of the proposed approach.

References


