A least squares identification algorithm for a state space model with multi-state delays

Ya Gu a, Rui Ding b, ∗

a Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, PR China
b School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, PR China

A R T I C L E   I N F O
Article history:
Received 16 January 2013
Received in revised form 22 February 2013
Accepted 22 February 2013

Keywords:
Parameter estimation
Recursive identification
Least squares
State space model
Time delay system

A B S T R A C T
A parameter estimator is presented for a state space model with time delay based on the given input–output data. The basic idea is to expand the state equations and to eliminate some state variables, and to substitute the state equation into the output equation to obtain the identification model which contains the information vector and parameter vector. A least squares algorithm is developed to estimate the system parameter vectors. Finally, an illustrative example is provided to verify the effectiveness of the proposed algorithm.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

System identification deals with the problem of estimating the unknown parameters of systems by using the measured input–output data [1–6]. In other words, identification is to build mathematical models of dynamical systems based on the observed data of the systems [7–10]. System identification has been widely used in many areas [11–15], e.g., mechanical systems, biological systems, and environmental systems. A large amount of work has been published in this research field [16,17]. For example, Hu and Ding proposed an iterative least squares estimation algorithm for controlled moving average systems based on matrix decomposition [18]. Ding et al. presented the gradient-based and least squares-based iterative estimation algorithms for multi-input multi-output systems, the key is to replace the unknown noise terms and residuals contained in the information vector with their corresponding estimates at the previous iteration [19]. Wang et al. proposed the hierarchical gradient-based iterative algorithm to interactively estimate the parameter matrix and the parameter vector by using the hierarchical identification principle and the gradient search, the advantage of the hierarchical gradient-based iterative algorithm is that it has fast convergence rate compared with the hierarchical stochastic gradient algorithm [20]. Zhang studied the unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods [21].

The least squares method has been the dominant algorithm for parameter estimation due to its simplicity in concept and convenience in implementation [22,23]. In this literature, Liu et al. presented the least squares estimation for a class of non-uniformly sampled systems based on the hierarchical identification principle [24]. Ding and Chen proposed the hierarchical least squares identification methods for multivariable systems [25]. Wang studied the least squares-based recursive and iterative estimation for output error moving average systems using data filtering, this study combines the auxiliary model

∗ This work was supported by the National Natural Science Foundation of China (No. 61273194) and the 111 Project (B12018).
* Corresponding author.
E-mail addresses: guya927819@163.com (Y. Gu), rding2003@yahoo.cn (R. Ding).

0893-9659/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aml.2013.02.005
identification idea with the filtering theory, transforms an output error moving average system into two identification models and presents a filtering and auxiliary model-based recursive least squares identification algorithm [26]. Han et al. discussed two recursive least squares parameter estimation algorithms for multirate multiple-input systems by using the auxiliary model [27].

In the recognition of the impact on system function, mathematical models incorporating time delays have been developed for an ever widening range of mode control and stabilization of systems [28–31]. Recently, Shi and Yu proposed the robust mixed \( H-2/H\)-infinity control of networked control systems with random time delays in both forward and backward communication links [32]; Ding and Gu analyzed the performance of the auxiliary model-based least squares identification algorithm for one-step state-delay systems [33]. Gu et al. presented an auxiliary model-based least squares identification method for a state space model with a unit time-delay [34] and the parameter and state estimation algorithms for a multi-variable state space system with \( d\)-step state-delay [35] using the hierarchical identification principle [36], and for a state space model with one unit time-delay [37]. This paper studies the parameter identification problem of general state space models multi-state delays based on the input–output data. This paper is different from the previous work in [34,35,37].

This paper is organized as follows. Section 2 derives an input–output representation of a state space system with time delay. Section 3 presents a parameter identification algorithm for state space systems with time delay. Section 4 provides an illustrative example for the results in this paper. Finally, we offer some concluding remarks in Section 5.

2. The system description and identification model

Let us introduce some notation. "\( A \) := \( X \)" or "\( X \) := \( A \)" stands for "\( A \) is defined as \( X \)"; the symbol \( I \) \( (i) \) stands for an identity matrix of appropriate size \( (n \times n) \); the superscript \( T \) denotes the matrix transpose; \( z \) represents a unit forward shift operator: \( z x(t) = x(t + 1) \) and \( z^{-1} x(t) = x(t - 1) \); \( \hat{\theta}(t) \) denotes the estimate of \( \theta \) at time \( t \).

Different from the plant model in [34,35,37], this paper considers the following state space system with multi-state delays,

\[
x(t + 1) = A x(t) + B_1 x(t - 1) + B_2 x(t - 2) + \cdots + B_r x(t - r) + f u(t),
\]

\[
y(t) = c x(t) + v(t),
\]

where \( x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T \in \mathbb{R}^n \) is the state vector, \( u(t) \in \mathbb{R} \) is the system input, \( y(t) \in \mathbb{R} \) is the system output, and \( v(t) \in \mathbb{R} \) is a random noise with zero mean. \( A \in \mathbb{R}^{n \times n}, B_i \in \mathbb{R}^{n \times n}, f \in \mathbb{R}^n \) and \( c \in \mathbb{R}^1 \) are the system parameter matrices/vectors, where the system matrices \( A, B_i \) and the system vector \( f \) are the unknown parameters to be estimated from input–output data \( \{u(t), y(t)\} \).

The following transforms the time delay state space model in (1) and (2) into an identification model. Referring to the method in [36] and according to the definition of \( A, B_i \) and \( f \), expanding (1) gives

\[
x_1(t + 1) = x_2(t) + b_{11} x(t - 1) + b_{12} x(t - 2) + \cdots + b_{1r} x(t - r) + f_1 u(t),
\]

\[
x_2(t + 1) = x_3(t) + b_{21} x(t - 1) + b_{22} x(t - 2) + \cdots + b_{2r} x(t - r) + f_2 u(t),
\]

\[
x_3(t + 1) = x_4(t) + b_{31} x(t - 1) + b_{32} x(t - 2) + \cdots + b_{3r} x(t - r) + f_3 u(t),
\]

\[
\vdots
\]

\[
x_{n-1}(t + 1) = x_n(t) + b_{1,n-1} x(t - 1) + b_{2,n-1} x(t - 2) + \cdots + b_{r,n-1} x(t - r) + f_{n-1} u(t),
\]

\[
x_n(t + 1) = a_n x_1(t) + a_{n-1} x_2(t) + \cdots + a_1 x_n(t) + b_{10} x(t - 1) + b_{20} x(t - 2) + \cdots + b_{r0} x(t - r) + f_0 u(t).
\]

Multiplying the \( i \)th equation of the above polynomials by \( z^{-i} \) gives

\[
x_1(t) = x_2(t - 1) + b_{11} x(t - 2) + b_{21} x(t - 3) + \cdots + b_{1r} x(t - r - 1) + f_1 u(t - 1),
\]

\[
x_2(t - 1) = x_3(t - 2) + b_{12} x(t - 3) + b_{22} x(t - 4) + \cdots + b_{2r} x(t - r - 2) + f_2 u(t - 2),
\]

\[
\vdots
\]

\[
x_{n-1}(t - 1) = x_n(t - 2) + b_{1,n-1} x(t - 3) + b_{2,n-1} x(t - 4) + \cdots + b_{r,n-1} x(t - r - 1) + f_{n-1} u(t - 1),
\]

\[
x_n(t - 1) = a_n x_1(t - 1) + a_{n-1} x_2(t - 1) + \cdots + a_1 x_n(t - 1) + b_{10} x(t - 2) + \cdots + b_{r0} x(t - r) + f_0 u(t - 2).
\]
Applying the matrix inversion formula in the parameter estimation algorithm [35]. The following discusses the parameter estimation algorithm

\[
x_3(t - 2) = x_4(t - 3) + b_{13}x(t - 4) + b_{23}x(t - 5) + \cdots + b_{r3}x(t - r - 3) + f_3u(t - 3),
\]

\[
:\vdots
\]

\[
x_{n-1}(t - n + 2) = x_n(t - n + 1) + b_{1,n-1}x(t - n) + b_{2,n-1}x(t - n - 1) + \cdots + b_{r,n-1}x(t - r - n + 1) + f_{n-1}u(t - n + 1),
\]

\[
x_n(t - n + 1) = ax(t - n) + b_{1n}x(t - n - 1) + b_{2n}x(t - n - 2) + \cdots + b_{rn}x(t - r - n) + f_nu(t - n),
\]

\[
\theta = [a_n, a_{n-1}, \ldots, a_1].
\]

Adding all expressions gives

\[
x_1(t) = b_{11}x(t - 2) + (b_{12} + b_{21})x(t - 3) + \cdots + (a + b_{1,n-1} + b_{2,n-2} + \cdots + b_{n-1,1})x(t - n)
\]

\[
+ \cdots + (b_{r-1,n} + b_{r,n-1})x(t - r - n + 1) + b_{rn}x(t - r - n)
\]

\[
+ f_1u(t - 1) + f_2u(t - 2) + \cdots + f_nu(t - n).
\]

Define the information vector \( \phi(t) \) and the parameter vector \( \theta \)

\[
\phi(t) := [x^T(t - 2), x^T(t - 3), \ldots, x^T(t - n), x^T(t - r - n + 1), x^T(t - r - n), u(t - 1), u(t - 2), \ldots, u(t - n)]^T \in \mathbb{R}^{n \times (r + n - 1)},
\]

\[
\theta := [b_{11}, b_{12} + b_{21}, \ldots, a + b_{1,n-1} + b_{2,n-2} + \cdots + b_{n-1,1}, \ldots, b_{r-1,n} + b_{r,n-1}, b_{rn}, f]^T \in \mathbb{R}^{n \times (r + n - 1)}.
\]

Substituting (3) into (2) gives the following identification model of the state space system with time delay,

\[
y(t) = b_{11}x(t - 2) + (b_{12} + b_{21})x(t - 3) + \cdots + (a + b_{1,n-1} + b_{2,n-2} + \cdots + b_{n-1,1})x(t - n)
\]

\[
+ \cdots + (b_{r-1,n} + b_{r,n-1})x(t - r - n + 1) + b_{rn}x(t - r - n)
\]

\[
+ f_1u(t - 1) + f_2u(t - 2) + \cdots + f_nu(t - n) + v(t)
\]

\[
= \phi^T(t)\theta + v(t).
\]

3. The parameter estimation algorithm

When the states are unavailable, we can estimate the parameters of the systems by using the hierarchical identification principle in [36] or the combined parameter and state estimation algorithm in [35]. The following discusses the parameter identification algorithm by using the measured state variables and input–output data.

Minimizing the criterion function,

\[
J(\theta) = \sum_{j=1}^t v^2(j) = \sum_{j=1}^t [y(j) - \phi^T(j)\theta]^2.
\]

Based on the identification model in (4), we can obtain the following recursive least squares algorithm for estimating the parameter vector \( \theta \) [23,38]:

\[
\hat{\theta}(t) = \hat{\theta}(t - 1) + P(t)\phi(t)[y(t) - \phi^T(t)\hat{\theta}(t - 1)],
\]

\[
P^{-1}(t) = P^{-1}(t - 1) + \phi(t)\phi^T(t), \quad P(0) = p_0I.
\]

\[
\phi(t) = [x^T(t - 2), x^T(t - 3), \ldots, x^T(t - n), x^T(t - r - n + 1), x^T(t - r - n), u(t - 1), u(t - 2), \ldots, u(t - n)]^T.
\]

Applying the matrix inversion formula

\[
(A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1}
\]

to Eq. (6), the recursive least squares algorithm can be summarized as

\[
\hat{\theta}(t) = \hat{\theta}(t - 1) + P(t)\phi(t)[y(t) - \phi^T(t)\hat{\theta}(t - 1)],
\]

\[
P^{-1}(t) = P^{-1}(t - 1) + \phi(t)\phi^T(t)P^{-1}(t - 1) + \phi^T(t)P^{-1}(t - 1)\phi(t), \quad P(0) = p_0I.
\]

\[
\phi(t) = [x^T(t - 2), x^T(t - 3), \ldots, x^T(t - n), x^T(t - r - n + 1), x^T(t - r - n), u(t - 1), u(t - 2), \ldots, u(t - n)]^T.
\]
Let $\mathbf{L}(t) := \mathbf{P}(t)\varphi(t)$. The covariance matrix can be expressed as

\[
\mathbf{P}(t) = \mathbf{P}(t - 1) - \frac{\mathbf{P}(t - 1)\varphi(t)\varphi^T(t)\mathbf{P}(t - 1)}{1 + \varphi^T(t)\mathbf{P}(t - 1)\varphi(t)} \\
= \mathbf{P}(t - 1) - \mathbf{P}(t - 1)\varphi(t)[1 + \varphi^T(t)\mathbf{P}(t - 1)\varphi(t)]^{-1}\varphi^T(t)\mathbf{P}(t - 1) \\
= \mathbf{P}(t - 1) - \mathbf{L}(t)[1 + \varphi^T(t)\mathbf{P}(t - 1)\varphi(t)]\mathbf{L}^T(t) \\
= [\mathbf{I} - \mathbf{L}(t)\varphi^T(t)]\mathbf{P}(t - 1) \\
= \mathbf{P}(t - 1) - \mathbf{L}(t)[\mathbf{P}(t - 1)\varphi(t)]^T.
\]

Here, we use the symmetric matrix property of $\mathbf{P}(t)$. To avoid a large amount of calculation, the recursive least squares algorithm is equivalently expressed as

\[
\hat{\theta}(t) = \hat{\theta}(t - 1) + \mathbf{L}(t)[y(t) - \varphi^T(t)\hat{\theta}(t - 1)], \\
\mathbf{L}(t) = \frac{\mathbf{P}(t - 1)\varphi(t)}{1 + \varphi^T(t)\mathbf{P}(t - 1)\varphi(t)}, \\
\mathbf{P}(t) = \mathbf{P}(t - 1) - \mathbf{L}(t)\varphi^T(t)\mathbf{P}(t - 1), \quad \mathbf{P}(0) = \mathbf{p}_0\mathbf{I}, \\
\hat{v}(t) = y(t) - \varphi^T(t)\hat{\theta}(t), \\
\varphi(t) = [\mathbf{x}^T(t - 2), \mathbf{x}^T(t - 3), \ldots, \mathbf{x}^T(t - n), \ldots, \mathbf{x}^T(t - r - n), \mathbf{x}^T(t - r - n), \\
u(t - 1), u(t - 2), \ldots, u(t - n)]^T.
\]

Let $t = 1$, set the initial values $\hat{\theta}(0) = \mathbf{1}_n/p_0, p_0 = 10^6, u(i) = 0, y(i) = 0$ and $\hat{v}(i) = 1/p_0$ for $i \leq 0$.

4. Example

Consider the following state space system with 2-step time delay:

\[
\mathbf{x}(t + 1) = \begin{bmatrix} 0 & 1 \\ -0.45 & -0.8 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 0.20 & -0.15 \\ 0.15 & -0.2 \end{bmatrix}\mathbf{x}(t - 1) + \begin{bmatrix} 0.20 \\ 0.18 \end{bmatrix}\mathbf{x}(t - 2) + \begin{bmatrix} 1 \\ -1 \end{bmatrix}u(t), \\
y(t) = [1, 0]\mathbf{x}(t) + v(t).
\]

The parameter vector to be identified is

\[
\theta = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8]^T \\
= [a + b_{11}, b_{12} + b_{21}, b_{22}. f_1]^T \\
= [a_3 + a + b_{11}, a + b_{12}, b_{13} + b_{21}, b_{14} + b_{22}, b_{23}, b_{24}, f_1, f_2]^T \\
= [-0.25, -0.95, 0.35, -0.38, 0.18, -0.05, 1.00, -1.00]^T.
\]

In simulation, the input $[u(t)]$ is taken as an uncorrelated persistent excitation signal sequence with zero mean and unit variance, and $[v(t)]$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$, respectively. Assuming that the system states are known and applying the least squares parameter estimation algorithm in (11)–(15) to estimate the parameters, the parameter estimates and their estimation errors are shown in Table 1 with $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$ and the parameter estimation errors $\delta$ versus $t$ are shown in Fig. 1.

From Table 1 and Fig. 1, we can draw the following conclusions: (1) the parameter estimation errors $\delta$ become smaller (in general) with the increasing of $t$ — see Fig. 1; (2) the parameter estimation accuracy becomes higher as the data length $t$ increases — see Table 1; (3) the parameter estimation errors under the same data lengths becomes higher for low noise levels — see Table 1.

5. Conclusions

In this paper, the identification problems are studied for linear systems based on the state-delay state space models with unknown parameters. A new parameter estimation algorithm has been presented to estimate the parameters of the systems. The proposed method can be extended to other linear or nonlinear systems with colored noises [39–49].
Table 1

The parameter estimates and errors with $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$.

<table>
<thead>
<tr>
<th>$\sigma^2$</th>
<th>$t$</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
<th>$\theta_4$</th>
<th>$\theta_5$</th>
<th>$\theta_6$</th>
<th>$\theta_7$</th>
<th>$\theta_8$</th>
<th>$\delta$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10^2</td>
<td>100</td>
<td>-0.28299</td>
<td>-0.97586</td>
<td>0.32661</td>
<td>-0.38171</td>
<td>0.16798</td>
<td>-0.06587</td>
<td>1.00546</td>
<td>-0.98879</td>
<td>2.95817</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>-0.26318</td>
<td>-0.95017</td>
<td>0.33684</td>
<td>-0.37992</td>
<td>0.16909</td>
<td>-0.04759</td>
<td>1.00712</td>
<td>-0.99403</td>
<td>1.30710</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>-0.25256</td>
<td>-0.94766</td>
<td>0.34848</td>
<td>-0.37166</td>
<td>0.17979</td>
<td>-0.04759</td>
<td>1.00641</td>
<td>-0.98916</td>
<td>0.64431</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-0.24648</td>
<td>-0.94091</td>
<td>0.35348</td>
<td>-0.37390</td>
<td>0.18190</td>
<td>-0.04248</td>
<td>1.00183</td>
<td>-0.99664</td>
<td>0.81912</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.25437</td>
<td>-0.95348</td>
<td>0.34938</td>
<td>-0.37713</td>
<td>0.18073</td>
<td>-0.05303</td>
<td>1.00133</td>
<td>-0.99911</td>
<td>0.39905</td>
</tr>
<tr>
<td></td>
<td>3000</td>
<td>-0.24761</td>
<td>-0.94788</td>
<td>0.34864</td>
<td>-0.38137</td>
<td>0.17933</td>
<td>-0.04617</td>
<td>1.00150</td>
<td>-0.99939</td>
<td>0.31130</td>
</tr>
<tr>
<td>0.50^2</td>
<td>100</td>
<td>-0.39755</td>
<td>-1.05834</td>
<td>0.23756</td>
<td>-0.39829</td>
<td>0.11678</td>
<td>-0.11493</td>
<td>1.02158</td>
<td>-1.02473</td>
<td>13.06648</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>-0.31071</td>
<td>-0.94443</td>
<td>0.28621</td>
<td>-0.38066</td>
<td>0.12506</td>
<td>-0.03269</td>
<td>1.03318</td>
<td>-1.01146</td>
<td>6.14516</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>-0.27513</td>
<td>-0.93693</td>
<td>0.34237</td>
<td>-0.33959</td>
<td>0.17795</td>
<td>-0.03720</td>
<td>1.02235</td>
<td>-1.00702</td>
<td>3.13488</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>-0.23201</td>
<td>-0.90412</td>
<td>0.36723</td>
<td>-0.35002</td>
<td>0.18910</td>
<td>-0.01198</td>
<td>1.00886</td>
<td>-0.99151</td>
<td>4.09297</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>-0.27173</td>
<td>-0.96726</td>
<td>0.34678</td>
<td>-0.36580</td>
<td>0.18346</td>
<td>-0.06491</td>
<td>1.00647</td>
<td>-0.99977</td>
<td>1.96240</td>
</tr>
<tr>
<td></td>
<td>3000</td>
<td>-0.23804</td>
<td>-0.93936</td>
<td>0.34319</td>
<td>-0.38689</td>
<td>0.17660</td>
<td>-0.03080</td>
<td>1.00741</td>
<td>-0.99976</td>
<td>1.55023</td>
</tr>
</tbody>
</table>

True values: $-0.25000$ $-0.95000$ $0.35000$ $-0.38000$ $0.18000$ $-0.05000$ $1.00000$ $-1.00000$

Fig. 1. The parameter estimation errors $\delta$ versus $t$ with $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$.

References


