Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Recursive Identification Algorithms Based on Minimizing Estimation Error ⋆

Kueiming Lo * **Hidenori Kimura** ** **Donghai Li** ***
Wook-Hyun Kwon ****

* *School of Software, Tsinghua University*
ISS Key Laboratory, Ministry of Education China
Tsinghua University, Beijing 100084, China
(E-mail: gluo@tsinghua.edu.cn)
** *Bio-mimetic Control Center*
Institute of Physical and Chemical Research
Moriyama-ku, Nagoya, JAPAN
*** *Department of Thermal Engineering*
Tsinghua University, Beijing 100084, CHINA
**** *School of Electrical Engineering*
Seoul National University, Seoul 151-742, Korea
(E-mail: whkwon@cisl.snu.ac.kr)

Abstract: Parameter selection for the criterion weighting matrix is concerned based on the information of both modifying the past estimation residuals and renewing the present estimation residual error . After minimizing the system estimation error, an optimal recursive algorithm is given. In this method the system data record can be used efficiently. The consistency of the new recursive algorithm is analyzed. Finally, some simulation examples are included to demonstrate the new method's reliability.

Keywords: system identification, recursive algorithm, optimal recursive algorithm

1. INTRODUCTION

Since the measured input-output data is processed sequentially, adaptation schemes are adopted to re-identify the system online or in real-time operation and form the recursive identification algorithm. With rapid development of computer science and automatic control, this method has gained considerable attention and appeared much work, especially the celebrated RLS algorithm (For example, see Mershed and Sayed [2000], Lo and Zhang [2000], Hassibi and Kaliath [2001], Hellgren and Forssell [2001], and Ahn *et al.* [2004]. These recursive methods were based on the conventional prediction error (PE) criterion Ljung [1999].

Because of a real system complicated, it is difficult to get a precise prediction model. Smaller prediction errors does not mean smaller parameter estimation errors. In fact, if some complicated interferences occurred, a considerable identification error would arise from the RLS algorithms Lo and Kwon [2002], Lo and Kwon [2003], Yazdi et al. [2005], and Chan et al. [2006].

A general recursive algorithm for discrete systems is considered in this paper. First, a simplified recursive algorithm is proposed. There are many tuning parameters contained in this recursive algorithm. Some regulating techniques are established and the free tuning parameters are determined based on the system data record. Then, the identification principle is to construct an identification algorithm to estimate the system parameters. An optimal recursive algorithm is constructed by minimizing the parameter estimation error. This recursive algorithm is able to resist system noise, including color noise, biased noise, and noise for some unmodeled disturbance. Furthermore, the consistency of the optimal recursive algorithm are analyzed, and some simulation examples are included to demonstrate the new method's reliability.

⋆ This work is supported by the Funds NSFC60672110, NSFC60474026, and the JSPS Foundation.

2. RECURSIVE IDENTIFICATION

Consider the SISO linear regression system:

$$y_t = \varphi_t^T \theta + w_t \tag{1}$$

where y_t is the system output. $\varphi_t \in R^n$ the regressor vector, θ is the unknown system parameter and w_t the noise. The system identification often includes the following performance:

$$J_t = [Y_t - \Phi_t \theta]^T Q_t [Y_t - \Phi_t \theta] \tag{2}$$

where

$$\Phi_t = (\varphi_1, \varphi_2, \cdots, \varphi_t)^T$$
$$Y_t = (y_1, y_2, \cdots, y_t)^T$$

and $\theta, \varphi_t \in R^n$. The matrix Q_t:

$$Q_t = \begin{pmatrix} Q_{t-1} & \alpha_t \\ \alpha_t^T & q_t \end{pmatrix}, \qquad \alpha_t \in R^{t-1}, \quad t = 1, 2, 3, \cdots$$

is a $t \times t$ symmetrical matrix. If the matrix $\Phi_t^T Q_t \Phi_t$ is nonsingular, minimizing cost function (2), it is not difficult to get the optimal solution of parameter θ:

$$\widehat{\theta}_t = [\Phi_t^T Q_t \Phi_t]^{-1} \Phi_t^T Q_t Y_t, \qquad t = 1, 2, 3, \cdots\cdots \tag{3}$$

Specifically, estimation (3) becomes a weighted LS algorithm if α_t are taken as zeros. When α_t are not zeros, in general, the calculation of parameter $\widehat{\theta}_t$ is complicated for every sample number t. For this online calculation, computing $\widehat{\theta}_t$ is especially difficult when the sample number t increases. Let:

$$P_t = \Phi_t^T Q_t \Phi_t$$

$$a_t = 1 + \varphi_t^T P_{t-1}^{-1} \Phi_{t-1}^T \alpha_t \quad (4)$$

$$\sigma_t = q_t - \alpha_t^T \Phi_{t-1} P_{t-1}^{-1} \Phi_{t-1}^T \alpha_t \quad (5)$$

$$b_t = a_t + a_t^{-1} \sigma_t \varphi_t^T P_{t-1}^{-1} \varphi_t \quad (6)$$

The recursive algorithm of relation (3) was proposed as (Lo and Kimura [2003], Lo et al. [2006]):

$$\begin{cases} \widehat{\theta}_t = \widehat{\theta}_{t-1} + \dfrac{1}{a_t b_t} P_t^{-1}(a_t \Phi_{t-1}^T \alpha_t + \sigma_t \varphi_t)(y_t - \varphi_t^T \widehat{\theta}_{t-1}) \\ \quad + \dfrac{1}{a_t b_t} P_{t-1}^{-1}(a_t \varphi_t \\ \quad - \varphi_t^T P_{t-1}^{-1} \varphi_t \Phi_{t-1}^T \alpha_t) \alpha_t^T (Y_{t-1} - \Phi_{t-1} \widehat{\theta}_{t-1}) \\ P_t^{-1} = P_{t-1}^{-1} - b_t^{-1} P_{t-1}^{-1}(\varphi_t \alpha_t^T \Phi_{t-1} + \Phi_{t-1}^T \alpha_t \varphi_t^T) P_{t-1}^{-1} \\ \quad + \dfrac{1}{a_t b_t} P_{t-1}^{-1}(\varphi_t^T P_{t-1}^{-1} \varphi_t \Phi_{t-1}^T \alpha_t \alpha_t^T \Phi_{t-1} \\ \quad - \sigma_t \varphi_t \varphi_t^T) P_{t-1}^{-1} \end{cases} \quad (7)$$

Algorithm (7) is composed of two decoupled complementary parts. One part renews the information of the current estimation residual; the other part modifies the estimation on past arithmetic errors. In this scheme it shows that at time t there are t tuning parameters: q_t and $\alpha_t \in R^{t-1}$. Therefore, a common question is posed: how are variables chosen to guarantee that the algorithm is more reliable? If α_t is chosen as: $\alpha_t = 0,\ t = 1, 2, \cdots$ and q_t is a positive constant, from expressions (4)-(6), we have $a_t = 1$, $\sigma_t = q_t$, and

$$b_t = 1 + q_t \varphi_t^T P_{t-1}^{-1} \varphi_t$$

It follows that algorithm (7) is the same as the RLS algorithm, which only considers the present estimation residual in the modification part.

Due to a large number of free variables as well as the tuning parameters in algorithm (7), there is space to improve the algorithm performance. By minimizing the frequency-domain estimate, the recursive empirical frequency-domain optimal parameter (REFOP) estimate was proposed (Lo and Kimura [2003] and Lo et al. [2006]). In that algorithm, the tuning parameters could be generated by the input signal and the output signal of a given discrete system. Since it is derived from the frequency domain, the calculation seems somewhat complex. It could not make the choice of tuning parameters efficient, either. In this paper, we propose an optimal recursive algorithm (ORA) instead of the REFOP method. In the proposed algorithm, the tuning parameters are determined by minimizing the estimation error directly.

Theorem 1. With the previous notations, algorithm (7) can be expressed as:

$$\begin{aligned} P_t^{-1} &= P_{t-1}^{-1} - b_t^{-1} P_{t-1}^{-1}(\varphi_t \alpha_t^T \Phi_{t-1} \\ &\quad + \Phi_{t-1}^T \alpha_t \varphi_t^T) P_{t-1}^{-1} \\ &\quad + \frac{1}{a_t b_t} P_{t-1}^{-1}(\varphi_t^T P_{t-1}^{-1} \varphi_t \Phi_{t-1}^T \alpha_t \alpha_t^T \Phi_{t-1} \\ &\quad - \sigma_t \varphi_t \varphi_t^T) P_{t-1}^{-1} \end{aligned} \quad (8)$$

$$\begin{aligned} \widehat{\theta}_t &= \widehat{\theta}_{t-1} + P_t^{-1} \Big[(\Phi_{t-1}^T \alpha_t + q_t \varphi_t)(y_t - \varphi_t^T \widehat{\theta}_{t-1}) \\ &\quad + \varphi_t \alpha_t^T (Y_{t-1} - \Phi_{t-1} \widehat{\theta}_{t-1}) \Big] \end{aligned} \quad (9)$$

The proof of Theorem 1 is omitted.

Theorem 1 is another form of the recursive algorithm of estimation (3). It comes from algorithm (7); however, instead of the calculated priority order of $\widehat{\theta}_t$ and P_t^{-1} at time t in algorithm (7), the first calculation in Theorem 1 is that of P_t^{-1}, and then the parameter $\widehat{\theta}_t$ is calculated by the result of P_t^{-1}. Since of the complicated calculation in algorithm (7), it is difficult to obtain the following algorithm. Therefore, Theorem 1 is necessary for the next analysis.

3. OPTIMAL RECURSIVE ALGORITHM

Lemma 1. Suppose that matrix P_{t-1} is positive definite. Then matrix P_t is also positive definite if

$$q_t > \alpha_t^T \Phi_{t-1} P_{t-1}^{-1} \Phi_{t-1}^T \alpha_t - (\varphi_t^T P_{t-1}^{-1} \varphi_t)^{-1} a_t^2. \quad (10)$$

The proof of Lemma 1 is omitted.

If the initial matrix P_0 is positive definite and $q_t, (t = 1, 2, 3, \cdots)$ satisfies the condition, Lemma 1 implies that the matrix P_t is also positive definite. Therefore, to optimize the recursive algorithm, the variable α_t can be chosen by minimizing

$$\widetilde{\theta}_t^T \widetilde{\theta}_t, \quad \text{or } \widetilde{\theta}_t^T P_t \widetilde{\theta}_t, \quad \text{or } \widetilde{\theta}_t^T P_t^2 \widetilde{\theta}_t$$

etc. Since

$$\widetilde{\theta}_t^T P_t^2 \widetilde{\theta}_t = ||\widetilde{\theta}_t^T P_t||^2$$

it is easy to see that $\widetilde{\theta} = 0$ if

$$\widetilde{\theta}_t^T P_t = 0$$

Furthermore, since the function $\widetilde{\theta}_t^T P_t^2 \widetilde{\theta}_t$ is more convenient to analyze some properties, it is chosen as the performance function in the following discussion.

Theorem 2. If the vector

$$\alpha_t = (\alpha_t(1), \alpha_t(2), \cdots, \alpha_t(t-1))^T \in R^{t-1}$$

is chosen such that:

$$\begin{aligned} &(W_{t-1}\varphi_t^T + \Phi_{t-1} w_t)(\varphi_t W_{t-1}^T + \Phi_{t-1}^T w_t)\alpha_t \\ &= (W_{t-1}\varphi_t^T + \Phi_{t-1} w_t)(P_{t-1}\widetilde{\theta}_{t-1} - q_t \varphi_t w_t) \end{aligned} \quad (11)$$

then performance function

$$J_t = \widetilde{\theta}_t^T P_t^2 \widetilde{\theta}_t$$

is the minimum, where:

$$\widetilde{\theta}_t = \theta - \widehat{\theta}_t, \qquad W_t = (w_1, w_2, \cdots, w_t)^T.$$

The proof of Theorem 2 is omitted.

Remark **1**. The number of the equations in relationship (12) is only n since $\theta \in R^n$. If t is large enough, n vectors

$$\varphi_{k1}, \cdots, \varphi_{k2}, \quad \varphi_{kn} \in \{\varphi_1, \varphi_2, \cdots, \varphi_t\}$$

can be chosen such that matrix

$$\overline{\Phi}^T w_t + \varphi_t \overline{W}^T$$

is nonsingular, where

$$\begin{aligned} \overline{\Phi} &= (\varphi_{k1}, \varphi_{k2}, \cdots, \varphi_{kn})^T \\ \overline{W} &= (w_{k1}, w_{k2}, \cdots, w_{kn})^T \\ \overline{\alpha}_t &= (\overline{\alpha}_t(k1), \overline{\alpha}_t(k2), \cdots, \overline{\alpha}_t(kn))^T \end{aligned}$$

$\overline{\alpha}_t \in R^n$, and for $l = 1, 2, \cdots, t-1$:

$$\alpha_t(l) = \begin{cases} \alpha_t(l) = \overline{\alpha}_t(l) & l = ki, \ i = 1, 2, \cdots, n \\ 0, & \text{others.} \end{cases}$$

In a simple implementation, however, for a fix number n we can revert a default positive value and use a switching technique for the matrix

$$\overline{F}^T w_t + \varphi_t \overline{W}^T$$

in case of singularity (see Example 1). Similar to the proof of Theorem 2, the condition:

$$\begin{aligned} &(\overline{W}\varphi_t^T + \overline{\Phi} w_t)(\varphi_t \overline{W}^T + \overline{\Phi}^T w_t)\overline{\alpha}_t \\ &= (\overline{W}\varphi_t^T + \overline{\Phi} w_t)(P_{t-1}\widetilde{\theta}_{t-1} - q_t \varphi_t w_t) \end{aligned} \quad (12)$$

is equivalent to relationship (11). Since of the nonsingular nature of the matrix

$$\overline{\Phi}^T w_t + \varphi_t \overline{W}^T$$

the minimum of performance function J_t can be achieved if the variable $\overline{\alpha}_t$ is chosen as:

$$\overline{\alpha}_t = (\varphi_t \overline{W}^T + \overline{\Phi}^T w_t)^{-1}(P_{t-1}\widetilde{\theta}_{t-1} - q_t \varphi_t w_t) \quad (13)$$

Remark **2**. From (12) it is easy to see that $\widetilde{\theta}_t = 0$ if matrix P_t is nonsingular and relation (13) is satisfied. This means that the real parameter θ can be obtained exactly at time t. In situations when symbols can not possibly be confused with each other, in the following discussion the symbols α_t and $\overline{\alpha}_t$, Φ_{t-1} and $\overline{\Phi}$, and W_{t-1} and $\overline{W}$ sometimes are not distinguished from each other.

However, this result is based on the following requirements: the signal $\{w_k\}_1^t$ and the difference $\widetilde{\theta}_{t-1}$ must be known and matrix P_t must be nonsingular. In practice it is impossible to get the precise information of signals $\{w_k\}_1^t$ and $\widetilde{\theta}_{t-1}$ in system identification. Thus, $\overline{\alpha}_t$ is only obtained by estimation:

$$\widehat{\alpha}_t = (\varphi_t \widehat{W}^T + \overline{\Phi}^T \widehat{w}_t)^{-1}(P_{t-1}\widehat{\widetilde{\theta}}_{t-1} - q_t \varphi_t \widehat{w}_t) \quad (14)$$

where $\widehat{\widetilde{\theta}}_{t-1}$ and $\widehat{w}_t$ are the estimation of $\widetilde{\theta}_{t-1}$ and w_t, and

$$\widehat{W} = (\widehat{w}_{k1}, \widehat{w}_{k2}, \cdots, \widehat{w}_{kn})^T$$

At this time the selection of $k1, k2, \cdots, kn$ should satisfy the condition that the matrix

$$\overline{\Phi}^T \widehat{w}_t + \varphi_t \widehat{W}^T$$

is nonsingular. Hereafter, it is discussed how to estimate $\widehat{w}_t$ and $\widehat{\widetilde{\theta}}_{t-1}$. From system (1) the noise can be expressed as:

$$w_t = y_t - \varphi_t^T \theta$$

The most natural way to estimate noise is:

$$\widehat{w}_t = y_t - \varphi_t^T \widehat{\theta}_{t-1} \quad (15)$$

The ORA method is made up of relationships (8), (9), (14), and (15). The justifiability of such an estimation is verified by the next theorem.

Theorem 3. Let q_t be chosen:

$$q_t \in \left(\frac{2|a_t| + (2-\varepsilon)|\overline{\alpha}_t^T \overline{\Phi} P_{t-1}^{-1} \overline{\Phi}^T \overline{\alpha}_t \varphi_t^T P_{t-1}^{-1} \varphi_t - a_t^2|}{(1-\varepsilon)\varphi_t^T P_{t-1}^{-1}\varphi_t}, \frac{1 - 2\varepsilon a_t}{\varepsilon \varphi_t^T P_{t-1}^{-1} \varphi_t} \right) \quad (16)$$

and the parameters $\widehat{\theta}_t$ be estimated by algorithms (8), (9), (14), and (15). Then we have

$$||\widetilde{\theta}_t|| \le (1-\varepsilon)||\widetilde{\theta}_{t-1}||,$$

where ε is a small positive number.

The proof of Theorem 3 is omitted.

Theorem 3 demonstrates that for an appropriate q_t the estimation based on algorithms (8), (9), (14), and (15) is consistent. Since the algorithm contains relationship (15), the result of Theorem 3 also conforms to the validity of noise estimate $\widehat{w}_t$. However, at time t it depends on $\widetilde{\theta}_{t-1}$. From the recursive algorithm the value of $\widetilde{\theta}_{t-1}$ can be calculated by the value of $\widetilde{\theta}_{t-2}$. In contrast, if the value of $\widetilde{\theta}_{t-2}$ is determined from algorithms (8), (9), (14), and (15), the values of α_{t-1} and $\widetilde{\theta}_{t-1}$ are obtained. Therefore, the key determining factor for α_t should be how to choose the initial value of $\widetilde{\theta}_0$. In order to add some flexibility to the algorithm's performance and to make the recursive algorithm more reliable and adaptable, in the following, one method for the estimation of $P_{t-1}\widetilde{\theta}_{t-1}$ is proposed. From (C1) it follows that:

$$P_t\widetilde{\theta}_t = P_0\widetilde{\theta}_0 - \sum_{k=1}^{t} \left[(\overline{\Phi}^T \overline{\alpha}_k + q_k \varphi_k) w_k + \varphi_k \overline{\alpha}_k^T \overline{W} \right] \quad (17)$$

If the initial matrix P_0 is taken as δI, where I is the identity matrix and δ is a small positive number, then $P_0\widetilde{\theta}_0$ can be omitted or the initial value is chosen by an any small vector. Therefore, $P_{t-1}\widetilde{\theta}_{t-1}$ can be estimated by the recursive form:

$$\begin{aligned} P_{t-1}\widehat{\widetilde{\theta}}_{t-1} &= -\sum_{k=1}^{t-1} \left[(\overline{\Phi}^T \widehat{\alpha}_k + q_k \varphi_k)\widehat{w}_k + \varphi_k \overline{\alpha}_k^T \widehat{W} \right] \\ &= P_{t-2}\widehat{\widetilde{\theta}}_{t-2} - (\overline{\Phi}^T \widehat{\alpha}_{t-1} \\ &\quad + q_{t-1}\varphi_{t-1})\widehat{w}_{t-1} - \varphi_{t-1}\overline{\alpha}_{t-1}^T \widehat{W} \end{aligned} \quad (18)$$

We can use another method to deal with $\widetilde{\theta}_{t-1}$ for estimating the values of $\overline{\alpha}_t$. Since

$$\widetilde{\theta}_{t-1} = \theta - \theta_{t-1}$$

in relationship (14), $\widetilde{\theta}_{t-1}$ can be estimated by an adaptive method. The average of the last m estimation values

$$\overline{\theta}_{t-1} = \frac{1}{m} \sum_{k=t-m}^{t-1} \widehat{\theta}_k$$

is used as the parameter θ, where m is a proper integer. Then, we have that:

$$\begin{aligned}\widehat{\widehat{\theta}}_{t-1} &= \overline{\theta}_{t-1} - \widehat{\theta}_{t-1} \\ &= \frac{1}{m}\sum_{k=t-m}^{t-1} \widehat{\theta}_k - \widehat{\theta}_{t-1}\end{aligned} \tag{19}$$

Lemma 1 and Theorem 3 give the choice of domain for q_t. Of course, this is a theoretic result for analyzing the performance of the new identification algorithm. In the following simulations, in fact, q_t is assigned a number from an interval $(0, 1]$.

4. SIMULATIONS

Most of the early system identification work was based on the assumption that the interference system noise was either Gaussian, m.d.s. signals, or white noise. In engineering, these restrictions do not reflect reality since system noises are unknown and are very complicated. In this discussion there is no restriction on the system noise; that is, the disturbance may be non-Gaussian, non-m.d.s., or non-white noise. This performance is demonstrated by the following examples.

To illustrate the behavior of the optimal recursive algorithms, a simulation trial was conducted for comparison with ordinary LS recursive algorithms, which was a special case of algorithms (8) and (9) with

$$\alpha_t = 0,\ t = 1, 2, 3, \cdots$$

For a real system, the output $\{y_t\}_1^N$ was generated by the system with a given input sequence $\{u_t\}_1^N$. The experimental sample number N was 2000. Let

$$\eta = (\sum_{t=1}^{N} w_t^2 / \sum_{k=1}^{N} y_k^2)^{\frac{1}{2}}$$

be the noise-to-signal ratio, which expresses the extent of model signal disturbance. θ was denoted as the real model parameter, while $\widehat{\theta}_{ORA}$ and $\widehat{\theta}_{LS}$ were the optimal recursive algorithm and the recursive least-squares (RLS) estimates.

Example 1. The discrete system was given as:

$$y_t = \frac{b_1}{1 + a_1 q^{-1}} u_{t-1} + w_t \tag{20}$$

The real system parameters were $a_1 = 0.8$ and $b_1 = 3$. The input signal $\{u_t\}_1^N$ was generated by a sine generator. $\{w_t\}_1^N$ was a stochastic disturbance with mean 1.55 and variance 0.52. It was a biased noise rather than a white noise. The output of the system was then generated by (20) with the noise-to-signal ratio $\eta = 0.1475$. The parameter was estimated according to the RLS method and the ORA1 method, which consisted of relationships (8), (9), (14), (15), and (18). The initial parameter $\widehat{\theta}_0$ was the zero vector. At time t the matrix $\overline{\Phi}$ was chosen as

$$\overline{\Phi} = (\varphi_{t-2}, \varphi_{t-1})^T$$

To avoid the singularity of the matrix

$$\varphi_t \widehat{W}^T + \overline{\Phi}^T \widehat{w}_t$$

in (14), the $\overline{\alpha}_t$ was chosen as the zero vector if

$$|\det(\varphi_t \widehat{W}^T + \overline{\Phi}^T \widehat{w}_t)| < 0.1$$

In fact, only six switches were encountered in the full simulation.

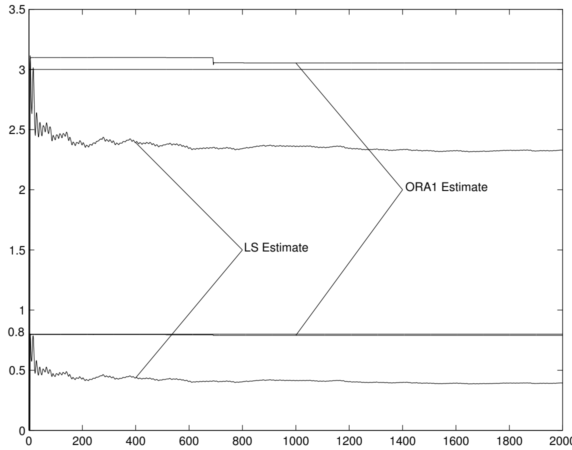

Fig. 1. Simulation of Example 1

The optimal recursive algorithm can be intuitively compared with the RLS method as shown in Figure 1:

which clearly shows that the optimal recursive algorithm method identifies the real system (20) more efficiently than does the RLS method. It is also validated by the following calculated values:

$$\overline{\theta}_{ORA1} = \begin{pmatrix} 0.7916 \pm 0.0025 \\ 3.0684 \pm 0.0206 \end{pmatrix}$$

$$\overline{\theta}_{LS} = \begin{pmatrix} 0.4100 \pm 0.0194 \\ 2.3531 \pm 0.0307 \end{pmatrix}$$

where $\overline{\theta}_{ORA1}$ denotes the average estimate from the 100th ORA value to the 2000th ORA value. $\overline{\theta}_{LS}$ denotes the average estimate from the value of the 100th LS estimate to the 2000th LS estimate. The calculation error is defined by standard deviation.

Example 2. The discrete system is given as:

$$y_t = \frac{u_{t-1}}{1 + a_1 q^{-1} + a_2 q^{-2}} + w_t \tag{21}$$

where

$$w_t = \frac{w1_t}{1 + q^{-1} + 0.2q^{-2}} + w2_t$$

The real system parameters were $a_1 = -1.7, a_2 = 0.8$. The input signal $\{u_t\}_1^N$ was generated by a pulse generator. $\{w1_t\}_1^N$ was an approximate white noise and $\{w2_t\}_1^N$ was a sawtooth signal. The noise $\{w_t\}_1^N$ consisted of these two signals with mean 0.0301 and variance 2.0466. As in example 1, this noise was not a white noise, either. The output of the system was then generated by (21) with a noise-to-signal ratio $\eta = 0.2237$. The parameter was estimated according to the RLS method and the ORA2 method, which consisted of relationships (8), (9), (14), (15), and (19) with $m = 5$. From (19) we can see that the desired estimation $\widehat{\widehat{\theta}}_{t-1}$ would depend on the choice of initial parameter $\widehat{\theta}_0$. In order to get better initial values, the recursive empirical frequency-domain optimal estimate was used in the first 200 steps, where $\overline{\alpha}_t$ was determined by (8) and $v(t) = y_t$ (Lo and Kimura 2003). At time t matrix $\overline{\Phi}$ was chosen as

$$\overline{\Phi} = (\varphi_{t-2}, \varphi_{t-1})^T$$

too. To avoid the singularity of the matrix

$$\varphi_t \widehat{W}^T + \overline{\Phi}^T \widehat{w}_t$$

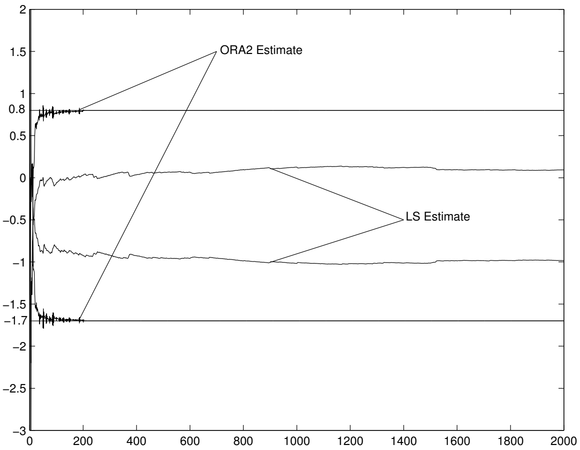

Fig. 2. Simulation of Example 2

in (14), $\overline{\alpha}_t$ was chosen as the zero vector if

$$|\det(\varphi_t \widehat{W}^T + \overline{\Phi}^T \widehat{w}_t)| < 0.1$$

However, no switch occurred in the full simulation. Figure 2 shows the simulation result.

Due to the interference of complicated and biased noise, a large error occurred using the ordinary RLS algorithm. Figure 2 shows that the estimation of the ORA2 method was more precise than that of the RLS method. The calculated values were given by the following:

$$\overline{\theta}_{ORA2} = \begin{pmatrix} -1.6969 \pm 0.0038 \\ 0.8000 \pm 0.0047 \end{pmatrix}$$

$$\overline{\theta}_{LS} = \begin{pmatrix} -0.9769 \pm 0.0383 \\ 0.0860 \pm 0.0387 \end{pmatrix}$$

where $\overline{\theta}_{ORA2}$ denotes the average estimate from the 100th ORA value to the 2000th ORA value. $\overline{\theta}_{LS}$ denotes the average estimate from the value of the 100th LS estimate to the 2000th LS estimate.

5. CONCLUSIONS

This paper presents a recursive algorithm for discrete systems, composed of two decoupled complementary parts. There are many tuning parameters contained in this recursive algorithm. They are included and constructed in a time-varying performance criterion and in the recursive algorithm. Minimizing the parameter estimation error, an ORA method was established. Compared with the previous recursive algorithms, in this method the system data record can be used efficiently, and computational time is less than that of the REFOP estimate (Lo and Kimura [2003]). Furthermore, the viability and consistency of the optimal recursive algorithm were analyzed. Since there was not any restriction on the system disturbance noise, theoretic analysis and simulation results indicated that the new algorithms have the advantage of being anti-interference, which includes protection against color noise, biased noise, and noise for some unmodeled systems.

The purpose of our research is to extend and adapt system identification to be efficiently used in complicated engineering applications. As a special case of algorithms (8) and (9), the LS algorithm should be the simplest choice because

$$\alpha_t = 0, \qquad t = 1, 2, 3, \cdots$$

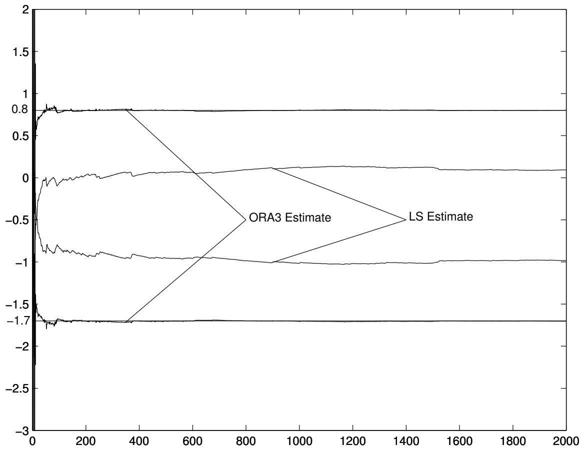

Fig. 3. Simulation of Example 2

But this does not prove that it provides the best method of choosing tuning parameters. In addition to the method discussed in this paper, of course, there should be other methods to choose tuning parameters α_t to make the algorithm more efficient. For instance, in the simulation of Example 2, if the tuning parameters are chosen as:

$$\alpha_t = (\alpha_{1t}, \alpha_{2t}, \cdots, \alpha_{(t-1)(t-1)})^T$$

with

$$\begin{cases} \alpha_{1k} = 0, \quad k = 1, 2, \cdots, t-3 \\ \alpha_{(t-2)t} = \alpha_{(t-1)t} = \sqrt{q_t} = \dfrac{1}{3} \end{cases} \qquad t = 3, 4, 5, \cdots \tag{22}$$

then the simulation results are as shown in Figure 3:

and the calculation values are as follows:

$$\overline{\theta}_{ORA3} = \begin{pmatrix} -1.7030 \pm 0.0051 \\ 0.7999 \pm 0.0053 \end{pmatrix}$$

where $\overline{\theta}_{ORA3}$ denotes the average of the estimation values of algorithms (8), (9), and (22). It is also more exact than the LS method in identifying real system (21).

REFERENCES

K. S. Ahn, J. Cho, and H. K. Baik, Blind adaptive channel equalization. using multichannel linear prediction-based cross-correlation vector estimation. *IEEE Trans. Consumer Electronics*, volume 50(4), papges 1026–1032, 2004.

S. C. Chan, Y. Wen, K. L. Ho. A robust past algorithm for subspace tracking in impulsive noise. *IEEE Trans. Signal Processing*, volume 54(1), pages 105–116, 2006.

B. Hassibi and T. Kaliath. H-infinity bounds for least-squares estimators. *IEEE Trans. Automat. Control*, volume 46(2), pages 309–314, 2001.

J. Hellgren and U. Forssell. Bias of feedback concellation algorithms in hearing aids based on direct closed loop identification. *IEEE Trans. Speech and Audio Processing*, volume 9(8), pages 906–913, 2001.

L. Ljung. System Identification: Theory for the User. Upper Saddle River, NJ: Prentice-Hall, 1999.

K. M. Lo and H. Kimura. Recursive estimation methods for discrete systems. *IEEE Trans. Automat. Control*, volume 48(11), pages 2019–2024, 2003.

K. M. Lo, H. Kimura, W. H. Kwon, and X. J. Yang. Empirical Frequency-domain Optimal Parameter Estimate for Black-box Processes. *IEEE Trans. Circuits and Systems I: Regular Papers*, voluem 53(2), pages 419–430, 2006.

K. M. Lo and W. H. Kwon. A new identification approach for FIR models. *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, volume 49(6), pages 439–446, 2002.

K. M. Lo and W. H. Kwon. New Identification Approaches for Disturbed Models. *Automatica*, volume 39(9), pages 1627–1634, 2003.

K. M. Lo and D. C. Zhang. Stochastic adaptive one-step-ahead optimal controllers based on input matching. *IEEE Trans. Automat. Control*, volume 45(5), pages 980–983, 2000.

R. Mershed and A. J. Sayed. Order-recursive RLS Laguerre adaptive filtering. *IEEE Trans. Signal Processing*, volume 48(11), pages 3000–3010, 2000.

H. S. Yazdi,M. Lotfizad, E. Kabir, and M. Fathy. Clipped input RLS applied to tracking. *EURASIP Journal on Applied Signal Processing*, volume 8, pages 1221–1228, 2005.