J Syst Sci Complex (2022) 35: 748–765

LS-Based Parameter Estimation of DARMA Systems with Uniformly Quantized Observations*

JING Lida · ZHANG Ji-Feng

DOI: 10.1007/s11424-021-0314-y
Received: 15 December 2020 / Revised: 21 January 2021

Abstract This paper is concerned with the parameter estimation of deterministic autoregressive moving average (DARMA) systems with quantization data. The estimation algorithms adopted here are the least squares (LS) and the forgetting factor LS, and the signal quantizer is of uniform, that is, with uniform quantization error. The authors analyse the properties of the LS and the forgetting factor LS, and establish the boundedness of the estimation errors and a relationship of the estimation errors with the size of quantization error, which implies that the smaller the quantization error is, the smaller the estimation error is. A numerical example is given to demonstrate theorems.

Keywords Discrete-time linear time-invariant systems, parameter estimation, quantized output.

1 Introduction

Owing to quantized data existing in numerous fields, system identification with quantized data is of importance in understanding the modeling capacity of the systems with limited sensor information[1]. For instance, in networked systems, transmitting quantized data instead of exact data can improve the communication efficiency, and storing quantized data instead of exact data can reduce the storage space. In fact, quantized data is inherent in many digital procedure, since data is usually obtained from a communication channel. So, it is natural to study how to get a desirable estimation of the system parameters by using quantized observations.

In the past two decades, identification with quantized data has become a hot topic and a large number of literature (e.g. [1–10]) emerged. Among others, [1] provided two different parameter estimation frameworks, respectively, for deterministic systems and stochastic systems with binary data. [2] gave the identification of regime-switching systems with binary data.

JING Lida · Zhang Ji-Feng (Corresponding author)
Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190*, China; School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100049*, China.* Email: jld@amss.ac.cn; jif@iss.ac.cn.
*The work was supported by National Key R&D Program of China under Grant No. 2018YFA0703800; the National Natural Science Foundation of China under Grant No. 61877057.
◇ *This paper was recommended for publication by Editor WU Zhengguang.*

[3] introduced a recursive projection algorithm for finite impulse response (FIR) systems and showed the parameter estimation convergence rate. [4] gave an iterative parameter estimate algorithm for systems with batched binary-valued observations based on the maximum likelihood method. [5] developed a parameter estimation approach for multi-input and multi-output FIR linear systems with quantized outputs. [6] proposed a parameter estimation method of quantized deterministic autoregressive moving average (DARMA) systems, and analyzed the boundedness of parameter estimation error. [7] used the method of [6] to distributed parameter identification of quantized DARMA multi-agent systems. Compared with the above-mentioned works, the purpose of this paper is to consider a general case with auto-regression part, weaken the conditions on the regression vectors used in [6] and [11] by the LS method, analyse properties of the parameter estimation errors, and especially, establish a relationship of the estimation errors with the size of quantization error.

Due to its simplicity and easy-to-use property, the LS method has played an important role in parameter estimation and has widely used in practice[12]. A lot of research works (see [12–26]) are on the LS and its variations. Even for the forgetting factor LS, one can find many different variations. For example, [21] designed the forgetting factor according with the current value of the output prediction error. [22] introduced a directional forgetting variant for getting a better exploitation of the incoming information. [23] applied the forgetting factor LS to tracking time-varying linear regression models and analysed the relationship of the parameter tracking error with the forgetting factor and the parameter time-varying speed.

Considering the wide use of quantized data and the LS method, it is of significance to study the LS and the forgetting factor LS with quantized data. In this paper, the LS and the forgetting factor LS of DARMA systems with uniform quantized data are researched. Through the proceedings of proving, we can find that the properties of matrix $P_{n+1}^{-1} = \sum_{i=0}^{n} \psi_i \psi_i^{\mathrm{T}}$ and $Q_{n+1}^{-1} = (1-\mu)Q_n^{-1} + \mu\psi_n\psi_n^{\mathrm{T}}$ are the key to getting the boundedness of parameter estimation errors for both the LS and the forgetting factor LS. In fact, we adopt a classic technique to design the system inputs so as to control the sizes of P_n^{-1} and Q_n^{-1}. An interesting phenomenon is that the forgetting factor LS does not need any condition on the size of the quantization error, while the LS does. One the contrary, in this paper we do not need the stability condition of the auto-regression part to ensure the boundedness of the LS, but need for the forgetting factor LS.

Quantization error is often state- and input-dependent, and not of white noises. Thus, the properties of the parameter estimation with quantization error are more complex. Thanks to the simple forms of the LS, we fortunately established the bounedness property of the LS parameter estimation and the relationship of the bounedness with the quantization error size.

In this paper, for a given vector or matrix x, x^{T} denotes the transpose of x; $||x||$ denotes the Euclidean norm for vector case and the corresponding induced norm for matrix case. The rest of this paper is organized as follows. Section 2 describes the system model and the form of uniform quantizer. Section 3 gives the LS for the quantized DARMA model, and researches the influence of quantization error on the parameter estimation error. Section 4 gives forgetting factor LS for the quantized DARMA model, and analyses the influence of quantization error on the parameter estimation error. Section 5 uses a numerical example to demonstrate the main results. Section 6 presents some conclusion remarks.

2 Model

Consider the following DARMA system:

$$A(z)y_{n+1} = B(z)u_n, \quad n \geq 0, \tag{1}$$

where y_n and u_n are the output and input. For simplicity, we suppose $y_n = u_n = 0$, $\forall n < 0$.

$$\begin{aligned} A(z) &= 1 + a_1 z + a_2 z^2 + \cdots + a_p z^p, \\ B(z) &= b_1 + b_2 z + \cdots + b_q z^{q-1}, \end{aligned}$$

where a_i and b_j are to be estimated, z is the shift-back operator and the orders p, q are assumed known.

The purpose of this paper is to estimate the following parameter vector by using system inputs and quantized outputs.

$$\theta = [-a_1, -a_2, \cdots, -a_p, b_1, b_2, \cdots, b_q]^{\mathrm{T}}. \tag{2}$$

For the convenience of proof, the model (1) can be rewritten as follows:

$$y_{n+1} = \theta^{\mathrm{T}} \varphi_n, \tag{3}$$

where

$$\varphi_n = [y_n, \cdots, y_{n-p+1}, u_n, \cdots, u_{n-q+1}]^{\mathrm{T}}. \tag{4}$$

This paper considers the condition that the system output y_n cannot be directly measured and only its quantized value is known. We want to design parameter estimation algorithm and analyze the influence of the quantization error on parameter estimation error.

For a given constant $\varepsilon > 0$, the quantizer used here is of the following uniform form:

$$s_n = \begin{cases} \vdots \\ -\varepsilon, & y_n \in \left[-\dfrac{3\varepsilon}{2}, -\dfrac{\varepsilon}{2}\right), \\ 0, & y_n \in \left[-\dfrac{\varepsilon}{2}, \dfrac{\varepsilon}{2}\right), \\ \varepsilon, & y_n \in \left[\dfrac{\varepsilon}{2}, \dfrac{3\varepsilon}{2}\right), \\ \vdots \end{cases} \tag{5}$$

From (5) we know that

$$s_{n+1} = \theta^{\mathrm{T}} \psi_n + \varepsilon_{n+1}, \tag{6}$$

where

$$\psi_n = [s_n, \cdots, s_{n-p+1}, u_n, \cdots, u_{n-q+1}]^{\mathrm{T}}. \tag{7}$$

It can be proved that

$$|\varepsilon_{n+1}| \le \frac{\varepsilon}{2}(|a_1| + |a_2| + \cdots + |a_p| + 1). \tag{8}$$

3 Parameter Estimation with the LS

Before going further, we first list the main conditions to be used in their paper.

Assumption 1 $A(z)$ and $B(z)$ are coprime, $a_p \neq 0$.

For any $x \in R^{p+q}$, $||x|| = 1$. Define $x = [x_1, x_2, \cdots, x_{p+q}]^{\mathrm{T}}$ and

$$H_x(z) = x_1 B(z) z + \cdots + x_p B(z) z^p + x_{p+1} A(z) + \cdots + x_{p+q} z^{q-1} A(z) = \sum_{i=0}^{p+q-1} g_i(x) z^i$$

and

$$L_x(z) = \sum_{i=1}^{p} x_i z^{i-1}$$

and

$$g(x) = [g_0(x), g_1(x), \cdots, g_{p+q-1}(x)]^{\mathrm{T}}.$$

Then, by Assumption 1 and Lemma 1 of [11] we know that

$$\min_{||x||=1} ||g(x)||^2 > 0.$$

Assumption 2 There exists a constant $\gamma > 0$ such that $|u_i| \le \gamma$ and

$$\lambda_{\min}\left(\sum_{i=k+1}^{N} U_i U_i^{\mathrm{T}}\right) \ge \delta, \quad \forall k \ge 0,\ N = k + p + q,$$

where

$$U_i = [u_i, u_{i-1}, \cdots, u_{i-p-q+1}]^{\mathrm{T}}, \tag{9}$$

δ is a positive constant.

Lemma 1 *If Assumption 2 is satisfied, then there exists a constant $c_0 > 0$ such that $|(H_x(z)u_i)(L_x(z)\varepsilon_i)| \le \frac{(p+q)c_0\varepsilon}{3(p+q+1)}$, for any $x \in R^{p+q}$, $||x|| = 1$.*

Proof Since $||x|| = 1$, the coefficients of $H_x(z)$ and $L_x(z)$ are bounded. From (8) we know that $|\varepsilon_i|$ is bounded. By Assumption 2, $|u_i|$ is bounded. So, there exists a constant $c_0 > 0$ such that $|(H_x(z)u_i)(L_x(z)\varepsilon_i)| \le \frac{(p+q)c_0\varepsilon}{3(p+q+1)}$. ∎

Lemma 2 *If Assumptions* 1–2 *are satisfied for* $\delta = \frac{(p+q)c_0\varepsilon}{\min_{||x||=1}||g(x)||^2}$, *then there is a constant* $\zeta > 0$ *such that when* $n \geq (p+q)(p+q+1)$,

$$\lambda_{\min}\left(\sum_{i=0}^{n}\psi_i\psi_i^{\mathrm{T}}\right) \geq \zeta n. \tag{10}$$

Proof From Assumption 2 we know that

$$\lambda_{\min}\left(\sum_{i=1}^{n}U_iU_i^{\mathrm{T}}\right) \geq \frac{n-(p+q)}{p+q}\delta \geq \frac{\delta}{p+q+1}n, \quad \forall n \geq (p+q)(p+q+1). \tag{11}$$

Let

$$\phi_n = A(z)\psi_n. \tag{12}$$

Then

$$\phi_n = \left[(zB(z)u_n + \varepsilon_n), \cdots, (z^pB(z)u_n + \varepsilon_{n-p+1}), A(z)u_n, \cdots, z^{q-1}A(z)u_n\right]^{\mathrm{T}}. \tag{13}$$

From (12), for any $x \in R^{p+q}$, $||x|| = 1$, we have

$$\begin{aligned} x^{\mathrm{T}}\left(\sum_{i=0}^{n}\phi_i\phi_i^{\mathrm{T}}\right)x = \sum_{i=0}^{n}\left(x^{\mathrm{T}}\phi_i\right)^2 &= \sum_{i=0}^{n}\left(\sum_{j=0}^{p}a_jx^{\mathrm{T}}\psi_{i-j}\right)^2 \\ &\leq \sum_{j=0}^{p}a_j^2\sum_{i=0}^{n}\sum_{j=0}^{p}\left(x^{\mathrm{T}}\psi_{i-j}\right)^2 \\ &\leq (p+1)\sum_{j=0}^{p}a_j^2\left(x^{\mathrm{T}}\sum_{i=0}^{n}\psi_i\psi_i^{\mathrm{T}}x\right), \end{aligned} \tag{14}$$

where $a_0 = 1$. So,

$$\lambda_{\min}\left(\sum_{i=0}^{n}\psi_i\psi_i^{\mathrm{T}}\right) \geq \frac{1}{(p+1)\sum_{j=0}^{p}a_j^2}\lambda_{\min}\left(\sum_{i=0}^{n}\phi_i\phi_i^{\mathrm{T}}\right).$$

Therefore, from Lemma 1 and (11) it follows that

$$\begin{aligned} x^{\mathrm{T}}\sum_{i=1}^{n}\phi_i\phi_i^{\mathrm{T}}x &= \sum_{i=1}^{n}\left(H_x(z)u_i + L_x(z)\varepsilon_i\right)^2 \\ &= g^{\mathrm{T}}(x)\sum_{i=1}^{n}U_iU_i^{\mathrm{T}}g(x) + 2\sum_{i=1}^{n}\left(H_x(z)u_i\right)\left(L_x(z)\varepsilon_i\right) + \sum_{i=1}^{n}\left(L_x(z)\varepsilon_i\right)^2 \\ &\geq \min_{||x||=1}||g(x)||^2\lambda_{\min}\left(\sum_{i=1}^{n}U_iU_i^{\mathrm{T}}\right) + 2\sum_{i=1}^{n}\left(H_x(z)u_i\right)\left(L_x(z)\varepsilon_i\right) \\ &\geq \min_{||x||=1}||g(x)||^2\lambda_{\min}\left(\sum_{i=1}^{n}U_iU_i^{\mathrm{T}}\right) - \frac{2(p+q)c_0\varepsilon}{3(p+q+1)}n \end{aligned}$$

$$
\begin{aligned}
&= \min_{||x||=1} ||g(x)||^2 \left(\lambda_{\min} \left(\sum_{i=1}^{n} U_i U_i^{\mathrm{T}} \right) - \frac{2\delta}{3\,(p+q+1)} n \right) \\
&\geq \frac{1}{3} \min_{||x||=1} ||g(x)||^2 \, \lambda_{\min} \left(\sum_{i=1}^{n} U_i U_i^{\mathrm{T}} \right),
\end{aligned}
$$

which implies

$$
\lambda_{\min} \left(\sum_{i=0}^{n} \phi_i \phi_i^{\mathrm{T}} \right) \geq \frac{1}{3} \min_{||x||=1} ||g(x)||^2 \, \frac{\delta}{p+q+1} n = \frac{(p+q)\, c_0 \varepsilon}{3\,(p+q+1)} n.
$$

Let $\zeta = \frac{(p+q)c_0\varepsilon}{3(p+q+1)(p+1)\sum_{j=0}^{p} a_j^2}$. This completes the proof. ∎

For θ, we use the following estimation algorithm:

$$
\theta_{n+1} = \left(\sum_{i=0}^{n} \psi_i \psi_i^{\mathrm{T}} \right)^{-1} \sum_{i=0}^{n} \psi_i s_{i+1} = P_{n+1} \sum_{i=0}^{n} \psi_i s_{i+1}, \tag{15}
$$

where

$$
P_{n+1} = \left(P_0^{-1} + \sum_{i=0}^{n} \psi_i \psi_i^{\mathrm{T}} \right)^{-1} = \left(P_n^{-1} + \psi_n \psi_n^{\mathrm{T}} \right)^{-1} = P_n - a_n P_n \psi_n \psi_n^{\mathrm{T}} P_n, \tag{16}
$$

$$
a_n = \left(1 + \psi_n^{\mathrm{T}} P_n \psi_n \right)^{-1}. \tag{17}
$$

From (15)–(17) it follows that

$$
\begin{aligned}
\theta_{n+1} &= \left(P_n - a_n P_n \psi_n \psi_n^{\mathrm{T}} P_n \right) \left(\sum_{i=0}^{n-1} \psi_i s_{i+1} + \psi_n s_{n+1} \right) \\
&= \theta_n - a_n P_n \psi_n \psi_n^{\mathrm{T}} \theta_n + P_n \psi_n s_{n+1} - a_n P_n \psi_n \psi_n^{\mathrm{T}} P_n \psi_n s_{n+1} \\
&= \theta_n - a_n P_n \psi_n \psi_n^{\mathrm{T}} \theta_n + P_n \psi_n \left(1 - a_n \psi_n^{\mathrm{T}} P_n \psi_n \right) s_{n+1} \\
&= \theta_n - a_n P_n \psi_n \psi_n^{\mathrm{T}} \theta_n + a_n P_n \psi_n s_{n+1} \\
&= \theta_n + a_n P_n \psi_n \left(s_{n+1} - \psi_n^{\mathrm{T}} \theta_n \right).
\end{aligned} \tag{18}
$$

Thus, we have obtained the recursive algorithm for the LS estimation.

We set

$$
P_0 = I, \tag{19}
$$

and take θ_0 arbitrarily.

Denote by $\lambda_{\min}(n)$ the smallest eigenvalue of P_{n+1}^{-1}.

Theorem 3 *For* (6), *suppose Assumptions* 1–2 *hold for* $\delta = \frac{(p+q)c_0\varepsilon}{\min_{||x||=1}||g(x)||^2}$. *Then, for all* $0 < \varepsilon < \frac{1}{2\left(1+\sum_{i=1}^{p}|a_i|\right)}$, *when* $n \geq (p+q)\,(p+q+1)$, *we have*

$$
\left|\left| \widetilde{\theta}_{n+1} \right|\right| \leq c_1 \left(\sqrt{\frac{1+\varepsilon}{n}} + \varepsilon \right), \tag{20}
$$

where

$$\widetilde{\theta}_n = \theta - \theta_n, \tag{21}$$

c_1 *is a constant and independent of* n *and* ε.

Proof Noticing $P_{n+1}^{-1} \geq \lambda_{\min}(n)I$, we see that

$$\left\|\widetilde{\theta}_{n+1}\right\|^2 \leq \frac{1}{\lambda_{\min}(n)}\widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1}. \tag{22}$$

Firstly, we need to prove there exist constants c_2, c_3 independent of n and ε such that

$$\widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1} \leq c_2 + c_3\varepsilon\,(n+1)\,. \tag{23}$$

From (17)–(18) it can be seen that

$$\begin{aligned} s_{n+1} - \psi_n^{\mathrm{T}}\theta_{n+1} =& s_{n+1} - \psi_n^{\mathrm{T}}\left(\theta_n + a_n P_n \psi_n \left(s_{n+1} - \psi_n^{\mathrm{T}}\theta_n\right)\right) \\ =& \left(1 - a_n \psi_n^{\mathrm{T}} P_n \psi_n\right)\left(s_{n+1} - \psi_n^{\mathrm{T}}\theta_n\right) \\ =& a_n\left(s_{n+1} - \psi_n^{\mathrm{T}}\theta_n\right). \end{aligned} \tag{24}$$

Hence, by (6), (21) and (24), we can rewrite (18) as

$$\begin{aligned} \widetilde{\theta}_{n+1} =& \widetilde{\theta}_n - a_n P_n \psi_n \left(s_{n+1} - \psi_n^{\mathrm{T}}\theta_n\right) \\ =& \widetilde{\theta}_n - P_n \psi_n \left(s_{n+1} - \psi_n^{\mathrm{T}}\theta_{n+1}\right) \\ =& \widetilde{\theta}_n - P_n \psi_n \left(\widetilde{\theta}_{n+1}^{\mathrm{T}}\psi_n + \varepsilon_{n+1}\right). \end{aligned} \tag{25}$$

We expand $\widetilde{\theta}_{k+1}^{\mathrm{T}} P_{k+1}^{-1} \widetilde{\theta}_{k+1}$ by using (16) and (25),

$$\begin{aligned} &\widetilde{\theta}_{k+1}^{\mathrm{T}} P_{k+1}^{-1} \widetilde{\theta}_{k+1} \\ =& \widetilde{\theta}_{k+1}^{\mathrm{T}}\left(P_k^{-1} + \psi_k\psi_k^{\mathrm{T}}\right)\widetilde{\theta}_{k+1} \\ =& \left[\widetilde{\theta}_k - P_k\psi_k\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)\right]^{\mathrm{T}} P_k^{-1}\left[\widetilde{\theta}_k - P_k\psi_k\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)\right] + \left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k\right)^2 \\ =& \left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k\right)^2 - 2\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)\widetilde{\theta}_k^{\mathrm{T}}\psi_k + \psi_k^{\mathrm{T}} P_k\psi_k\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)^2 + \widetilde{\theta}_k^{\mathrm{T}} P_k^{-1}\widetilde{\theta}_k \\ =& \widetilde{\theta}_k^{\mathrm{T}} P_k^{-1}\widetilde{\theta}_k - 2\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)\left[\widetilde{\theta}_{k+1} + P_k\psi_k\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)\right]^{\mathrm{T}}\psi_k \\ &+ \psi_k^{\mathrm{T}} P_k\psi_k\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)^2 + \left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k\right)^2 \\ =& \widetilde{\theta}_k^{\mathrm{T}} P_k^{-1}\widetilde{\theta}_k + \left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k\right)^2 - \psi_k^{\mathrm{T}} P_k\psi_k\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)^2 - 2\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k + \varepsilon_{k+1}\right)\left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k\right) \\ \leq& \widetilde{\theta}_k^{\mathrm{T}} P_k^{-1}\widetilde{\theta}_k - \left(\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k\right)^2 - 2\varepsilon_{k+1}\widetilde{\theta}_{k+1}^{\mathrm{T}}\psi_k. \end{aligned} \tag{26}$$

Summing up both sides of (26) from 0 to n and letting $c_2 = \widetilde{\theta}_0^{\mathrm{T}} P_0^{-1}\widetilde{\theta}_0$, we get

$$\begin{aligned} \widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1}\widetilde{\theta}_{n+1} \leq& \widetilde{\theta}_0^{\mathrm{T}} P_0^{-1}\widetilde{\theta}_0 - \sum_{i=0}^{n}\left(\widetilde{\theta}_{i+1}^{\mathrm{T}}\psi_i\right)^2 - 2\sum_{i=0}^{n}\varepsilon_{i+1}\widetilde{\theta}_{i+1}^{\mathrm{T}}\psi_i \\ =& c_2 - \sum_{i=0}^{n}\left(\widetilde{\theta}_{i+1}^{\mathrm{T}}\psi_i\right)^2 - 2\sum_{i=0}^{n}\varepsilon_{i+1}\widetilde{\theta}_{i+1}^{\mathrm{T}}\psi_i, \end{aligned} \tag{27}$$

or equivalently,

$$\widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1} + \sum_{i=0}^{n} \left(\widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right)^2 \le c_2 + \left| 2 \sum_{i=0}^{n} \varepsilon_{i+1} \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right|. \tag{28}$$

From (8) and $0 < \varepsilon < \frac{1}{2\left(1+\sum_{i=1}^{p}|a_i|\right)}$, we have

$$\begin{aligned}
\left| 2 \sum_{i=0}^{n} \varepsilon_{i+1} \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right| \le & 2 \sum_{i=0}^{n} |\varepsilon_{i+1}| \left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right| \\
\le & \varepsilon \left(1 + \sum_{i=1}^{p} |a_i| \right) \sum_{i=0}^{n} \left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right| \\
\le & \varepsilon \left(1 + \sum_{i=1}^{p} |a_i| \right) \sum_{i=0}^{n} \left(\left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right|^2 + 1 \right) \\
= & \varepsilon \left(1 + \sum_{i=1}^{p} |a_i| \right) \sum_{i=0}^{n} \left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right|^2 + \left(1 + \sum_{i=1}^{p} |a_i| \right) \varepsilon (n+1) \\
< & \frac{1}{2} \sum_{i=0}^{n} \left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right|^2 + \left(1 + \sum_{i=1}^{p} |a_i| \right) \varepsilon (n+1).
\end{aligned} \tag{29}$$

From (28)–(29) we know that there is a constant c_3 independent of n and ε such that

$$\begin{aligned}
& \widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1} + \sum_{i=0}^{n} \left(\widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right)^2 \\
\le & c_2 + \frac{1}{2} \sum_{i=0}^{n} \left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right|^2 + \left(1 + \sum_{i=1}^{p} |a_i| \right) \varepsilon (n+1) \\
\le & c_2 + \frac{1}{2} \sum_{i=0}^{n} \left| \widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right|^2 + c_3 \varepsilon (n+1).
\end{aligned} \tag{30}$$

Thus, we have

$$\widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1} \le \widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1} + \frac{1}{2} \sum_{i=0}^{n} \left(\widetilde{\theta}_{i+1}^{\mathrm{T}} \psi_i \right)^2 \le c_2 + c_3 \varepsilon (n+1). \tag{31}$$

So, (23) is proved.

From Lemma 2, (22) and (31), it can be seen that

$$\begin{aligned}
\left\| \widetilde{\theta}_{n+1} \right\|^2 \le \frac{1}{\lambda_{\min}(n)} \widetilde{\theta}_{n+1}^{\mathrm{T}} P_{n+1}^{-1} \widetilde{\theta}_{n+1} \le & \frac{c_2 + c_3 \varepsilon (n+1)}{\zeta n} \\
= & c_4 \frac{1}{n} + c_5 \varepsilon + c_5 \frac{\varepsilon}{n} \le c_6 \left(\frac{1+\varepsilon}{n} + \varepsilon \right),
\end{aligned} \tag{32}$$

where $c_4 = \frac{c_2}{\zeta}$, $c_5 = \frac{c_3}{\zeta}$, $c_6 = \max\{c_4, c_5\}$. So,

$$\left\| \widetilde{\theta}_{n+1} \right\| \le \sqrt{c_6} \left(\sqrt{\frac{1+\varepsilon}{n} + \varepsilon} \right). \tag{33}$$

This completes the proof. ∎

Remark 1 When the boundedness of $|a_i|$ is known, a suitable ε can easily be chosen. It is worth pointing out that $0 < \varepsilon < \frac{1}{2\left(1+\sum_{i=1}^{p}|a_i|\right)}$ is only a sufficient condition, which means that for some $\varepsilon \geq \frac{1}{2\left(1+\sum_{i=1}^{p}|a_i|\right)}$, $\left|\left|\widetilde{\theta}_n\right|\right|$ may also be bounded. This can be seen from the numerical example in Section 5.

Remark 2 Here we would like to emphasize that the proofs of Lemma 2 and Theorem 3 are very similar to that of the classic paper [11] where a series of novel and subtle analysis skills were firstly introduced.

4 Parameter Estimation with Forgetting Factor LS

In addition to Assumptions 1–2, in this section we will need the following assumption.

Assumption 3 $A(z)$ is stable, i.e., $A(z) \neq 0$, $\forall\ |z| \leq 1$.

Then, we have the following result.

Lemma 4 (see [6]) *If Assumptions* 1–2 *are satisfied for* $\delta = \frac{(p+q)c_0\varepsilon}{\min_{||x||=1}||g(x)||^2}$, *then there exist an integer* $h > 0$ *and a constant* $c_7 > 0$ *such that*

$$\lambda_{\min}\left(\Phi_h(t)\right) > c_7, \tag{34}$$

where

$$\Phi_h(t) = \sum_{i=t+1}^{t+h} \psi_{i-1}\psi_{i-1}^{\mathrm{T}}, \quad t \geq 0. \tag{35}$$

For θ, we use the following estimation algorithm:

$$\theta_n = \theta_{n-1} + \mu Q_n \psi_{n-1}\left(s_n - \psi_{n-1}^{\mathrm{T}}\theta_{n-1}\right), \quad \mu \in (0,1), \tag{36}$$

where

$$Q_n = \frac{1}{1-\mu}\left(Q_{n-1} - \mu\frac{Q_{n-1}\psi_{n-1}\psi_{n-1}^{\mathrm{T}}Q_{n-1}}{1-\mu+\mu\psi_{n-1}^{\mathrm{T}}Q_{n-1}\psi_{n-1}}\right), \tag{37}$$

with deterministic initial conditions θ_0 and $Q_0 > 0$.

Let

$$R_n \triangleq Q_n^{-1}. \tag{38}$$

Then, from (37) and the matrix inversion formula, we have

$$R_n = (1-\mu)R_{n-1} + \mu\psi_{n-1}\psi_{n-1}^{\mathrm{T}}. \tag{39}$$

Lemma 5 *If Assumptions* 1–2 *are satisfied for* $\delta = \frac{(p+q)c_0\varepsilon}{\min_{||x||=1}||g(x)||^2}$, *then there is a constant* $c_8 > 0$ *such that for any* $\mu_0 \in (0,1)$,

$$\sup_{\mu\in(0,\mu_0]} ||Q_n|| \leq c_8. \tag{40}$$

Proof By (39) we need only to prove the boundedness of the subsequence $||Q_{hk}||$, $k \geq 1$. By (39) we have

$$R_{kh+h} = (1-\mu)^h R_{kh} + \mu \sum_{i=kh+1}^{kh+h} (1-\mu)^{kh+h-i} \psi_{i-1}\psi_{i-1}^{\mathrm{T}}. \tag{41}$$

Hence, by (35) we have

$$\lambda_{\min}(R_{kh+h}) \geq (1-\mu)^h \lambda_{\min}(R_{kh}) + \mu (1-\mu)^h \lambda_{\min}(\Phi_h(kh)), \tag{42}$$

or

$$\alpha_{k+1} \geq \lambda^h (\alpha_k + \mu\beta_{k+1}), \quad \lambda = 1-\mu, \quad k \geq 0, \tag{43}$$

where $\alpha_k \triangleq \lambda_{\min}(R_{kh})$ and $\beta_{k+1} \triangleq \lambda_{\min}(\Phi_h(kh))$.

By (43) and the Schwarz inequality, we have

$$\alpha_k \geq \mu\lambda^h \left(\sum_{i=1}^{k} (\lambda^h)^{k-i}\beta_i\right) \geq \mu\lambda^h \left(\sum_{i=1}^{k} (\lambda^h)^{k-i}\right)^2 \left(\sum_{i=1}^{k} (\lambda^h)^{k-i}\beta_i^{-1}\right)^{-1}. \tag{44}$$

By Lemma 4, there is a constant $K > 0$ such that

$$\sum_{i=1}^{k} \left(\lambda^h\right)^{k-i} \beta_i^{-1} \leq K. \tag{45}$$

So, we have

$$\begin{aligned} \alpha_k^{-1} &\leq K\mu^{-1}\lambda^{-h} \left(\sum_{i=1}^{k} \left(\lambda^h\right)^{k-i}\right)^{-2} \\ &\leq K\mu^{-1}\lambda^{-h} \left(\sum_{i=1}^{k} \left(\lambda^h\right)^{k-i}\right)^{-1} \\ &= K\mu^{-1}\lambda^{-h} \frac{1-\lambda^h}{1-\lambda^{kh}} \\ &\leq Kh(1-\mu_0)^{-h} \frac{1}{1-\lambda^{kh}}, \quad \mu \in (0, \mu_0], \end{aligned} \tag{46}$$

where we have used the inequality used in [23],

$$\mu^{-1}\left(1-\lambda^h\right) \leq h, \quad \mu \in (0,1).$$

Notice that $(1-\mu)^{\frac{1}{\mu}}, \mu \in (0,1)$ is a decreasing function of μ and that

$$\lim_{\mu \to 0} (1-\mu)^{\frac{1}{\mu}} = e^{-1}. \tag{47}$$

Then, we have

$$\sup_{k\geq\mu^{-1}}\alpha_k^{-1}\leq Kh\left(1-\mu_0\right)^{-h}\frac{1}{1-\left(1-\mu\right)^{h/\mu}}\leq Kh\left(1-\mu_0\right)^{-h}\frac{1}{1-e^{-h}}. \tag{48}$$

On the other hand, by (43) we have

$$\alpha_k^{-1}\leq\lambda^{-hk}\alpha_0^{-1}, \tag{49}$$

and hence, for $\mu\in(0,\mu_0]$,

$$\sup_{k\leq\mu^{-1}}\alpha_k^{-1}\leq\left(1-\mu\right)^{-h/\mu}\alpha_0^{-1}\leq\left(1-\mu_0\right)^{-h/\mu_0}\alpha_0^{-1}, \tag{50}$$

where we have used the inequality

$$\left(1-\mu\right)^{-1/\mu}\leq\left(1-\mu_0\right)^{-1/\mu_0}, \quad \mu\leq\mu_0.$$

Let

$$c_8=\min\left\{\frac{Kh\left(1-\mu_0\right)^{-h}}{1-e^{-h}},\left(1-\mu_0\right)^{-h/\mu_0}\alpha_0^{-1}\right\}. \tag{51}$$

This completes the proof. ▮

Theorem 6 *For* (6)*, if Assumptions* 1–3 *are satisfied for* $\delta=\frac{(p+q)c_0\varepsilon}{\min_{||x||=1}||g(x)||^2}$*, then under the estimation algorithm* (36)–(37)*, we have*

$$\left|\left|\widetilde{\theta}_n\right|\right|\leq c_9\left(\left(1-\mu\right)^n+\varepsilon\right), \tag{52}$$

where

$$\widetilde{\theta}_n=\theta-\theta_n, \tag{53}$$

c_9 *is a constant and independent of* n *and* ε*.*

Proof From (6) and (36) it follows that

$$\begin{aligned}\widetilde{\theta}_n=&\widetilde{\theta}_{n-1}-\mu Q_n\psi_{n-1}\left(s_n-\psi_{n-1}^{\mathrm{T}}\theta_{n-1}\right)\\=&\widetilde{\theta}_{n-1}-\mu Q_n\psi_{n-1}\left(\psi_{n-1}^{\mathrm{T}}\theta+\varepsilon_n-\psi_{n-1}^{\mathrm{T}}\theta_{n-1}\right)\\=&\widetilde{\theta}_{n-1}-\mu Q_n\psi_{n-1}\left(\psi_{n-1}^{\mathrm{T}}\widetilde{\theta}_{n-1}+\varepsilon_n\right)\\=&\left[I-\mu Q_n\psi_{n-1}\psi_{n-1}^{\mathrm{T}}\right]\widetilde{\theta}_{n-1}-\mu Q_n\psi_{n-1}\varepsilon_n,\end{aligned} \tag{54}$$

which together with (39) implies

$$\begin{aligned}R_n\widetilde{\theta}_n=&R_n\left[I-\mu Q_n\psi_{n-1}\psi_{n-1}^{\mathrm{T}}\right]\widetilde{\theta}_{n-1}-\mu\psi_{n-1}\varepsilon_n\\=&\left[R_n-\mu\psi_{n-1}\psi_{n-1}^{\mathrm{T}}\right]\widetilde{\theta}_{n-1}-\mu\psi_{n-1}\varepsilon_n\\=&\left(1-\mu\right)R_{n-1}\widetilde{\theta}_{n-1}-\mu\psi_{n-1}\varepsilon_n,\end{aligned} \tag{55}$$

or

$$R_n\widetilde{\theta}_n = (1-\mu)^n R_0\widetilde{\theta}_0 - \sum_{i=1}^{n}(1-\mu)^{n-i}\mu\psi_{i-1}\varepsilon_i. \tag{56}$$

By Assumptions 2–3, one can see that $||\psi_i||$ is bounded. Furthermore, by (56), there exist constants c_{10}, c_{11} independent of n and ε such that

$$\left|\left|R_n\widetilde{\theta}_n\right|\right| \le \left|\left|(1-\mu)^n R_0\widetilde{\theta}_0\right|\right| + \left|\left|\sum_{i=1}^{n}(1-\mu)^{n-i}\mu\psi_{i-1}\varepsilon_i\right|\right| = c_{10}(1-\mu)^n + c_{11}\varepsilon. \tag{57}$$

Therefore, it follows from Lemma 5 that

$$\begin{aligned}\left|\left|\widetilde{\theta}_n\right|\right| = \left|\left|Q_nR_n\widetilde{\theta}_n\right|\right| &\le ||Q_n||\left|\left|R_n\widetilde{\theta}_n\right|\right| \\ &\le c_8\left|\left|R_n\widetilde{\theta}_n\right|\right| \\ &\le c_8c_{10}(1-\mu)^n + c_8c_{11}\varepsilon \\ &\le c_9\left((1-\mu)^n+\varepsilon\right),\end{aligned} \tag{58}$$

where $c_9 = \max\{c_8c_{10}, c_8c_{11}\}$. This completes the proof. ▮

Remark 3 Compared with [6], we can find that the parameter estimation errors of projection algorithm, the LS and the forgetting factor LS are dependent on quantization error. This is natural, because the size of quantization error reflects the accuracy of quantization. The smaller the quantization error is, or the higher the accuracy is, the smaller the estimation error is.

Remark 4 There are some limits of the applicability of the LS and the forgetting factor LS, which are mainly in how to choose suitable quantizers and system inputs. From the proof of Theorem 3 we know that the quantization error should be pretty small so as to get a satisfactory estimation error. For the choice of system inputs, u_i can be designed as periodic so as to make Assumption 2 satisfied.

5 Numerical Example

In this section, we will illustrate our theoretical result with a simulation example.

Consider the following system: $y_n = ay_{n-1} + bu_{n-1}$, where $\theta = [a, b]^{\mathrm{T}} = [0.5, 1]^{\mathrm{T}}$ is the parameter to be estimated, $\theta_0 = [0, 0]^{\mathrm{T}}$. By the condition of Theorem 3, ε should satisfy $0 < \varepsilon < \frac{1}{2\left(1+\sum_{i=1}^{p}|a_i|\right)} = \frac{1}{3}$. In order to clarify Remark 1, let y_n be quantized by (5) under $\varepsilon = 0.1$, $\varepsilon = 0.3$ and $\varepsilon = 1$, respectively. $u_n = 1, -1, -3, 1, -1, -3, \cdots$, which satisfies Assumption 2. We estimate θ by (15) and (36). The simulation results are given by Figures 1–12.

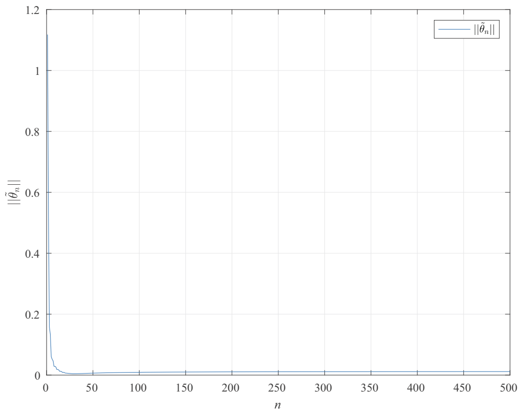

Figure 1 The trajectory of $\left|\left|\widetilde{\theta}_n\right|\right|$ with $\varepsilon = 0.1$ by the LS

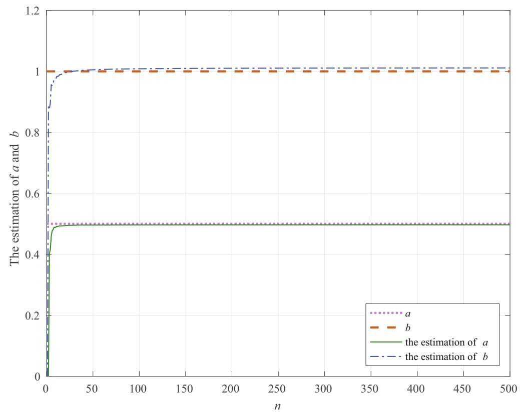

Figure 2 The trajectories of a and b with $\varepsilon = 0.1$ by the LS

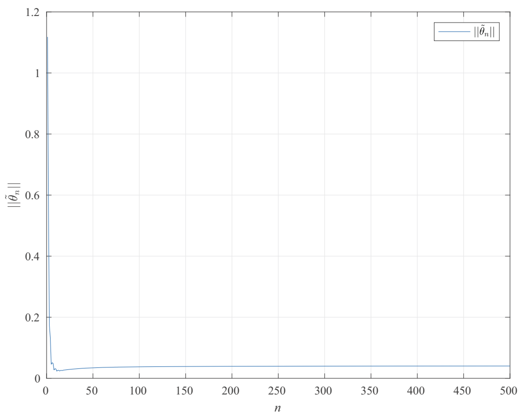

Figure 3 The trajectory of $\left|\left|\widetilde{\theta}_n\right|\right|$ with $\varepsilon = 0.3$ by the LS

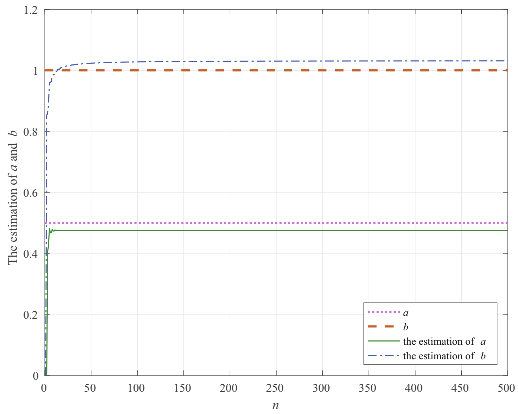

Figure 4 The trajectories of a and b with $\varepsilon = 0.3$ by the LS

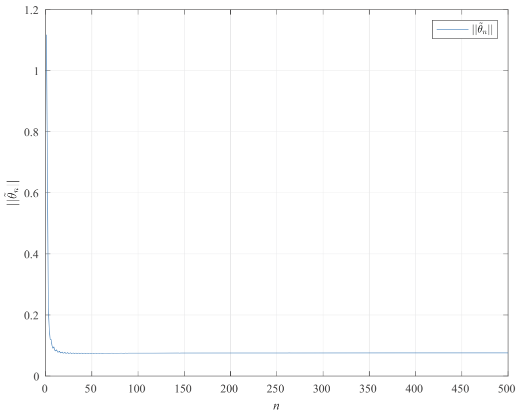

Figure 5 The trajectory of $\left\|\widetilde{\theta}_n\right\|$ with $\varepsilon = 1$ by the LS

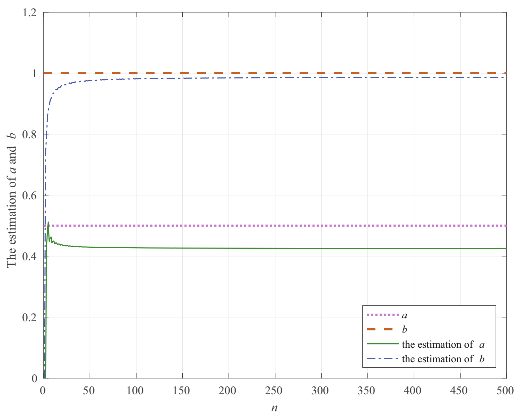

Figure 6 The trajectories of a and b with $\varepsilon = 1$ by the LS

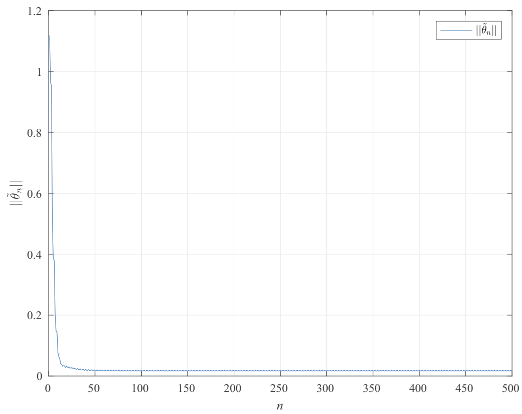

Figure 7 The trajectory of $\left\|\widetilde{\theta}_n\right\|$ with $\varepsilon = 0.1$ by forgetting factor LS

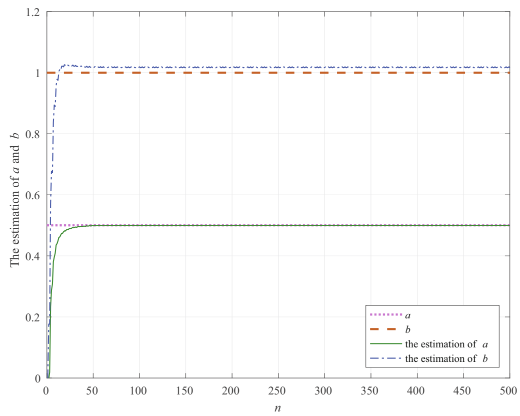

Figure 8 The trajectories of a and b with $\varepsilon = 0.1$ by forgetting factor LS

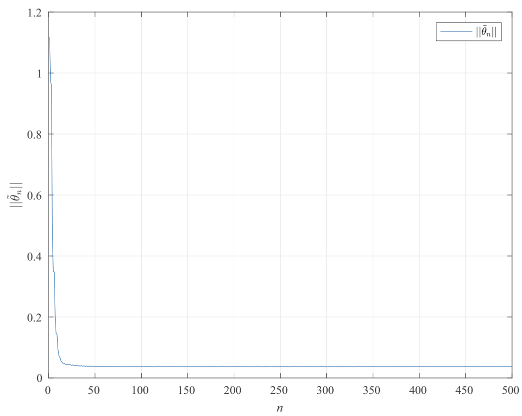

Figure 9 The trajectory of $\left\|\widetilde{\theta}_n\right\|$ with $\varepsilon = 0.3$ by forgetting factor LS

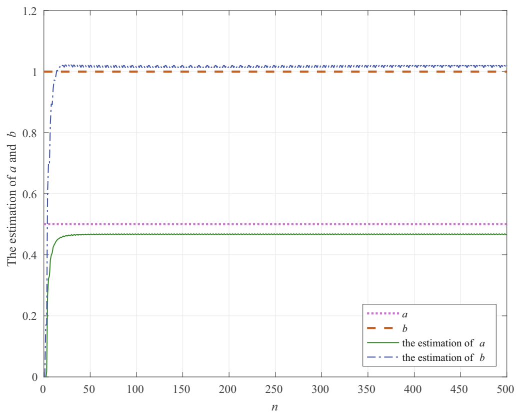

Figure 10 The trajectories of a and b with $\varepsilon = 0.3$ by forgetting factor LS

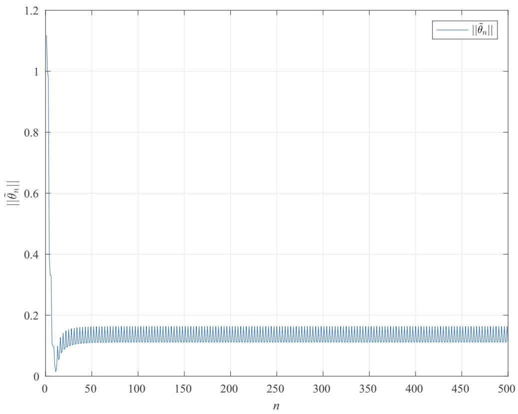

Figure 11 The trajectory of $\left|\left|\widetilde{\theta}_n\right|\right|$ with $\varepsilon = 1$ by forgetting factor LS

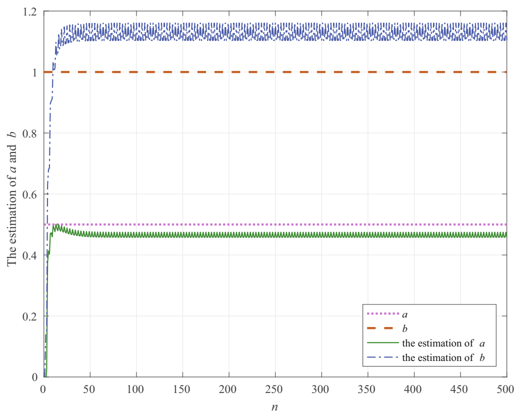

Figure 12 The trajectories of a and b with $\varepsilon = 1$ by forgetting factor LS

From Figures 1–12 it can be seen that the parameter estimation error is bounded and dependent on the size of quantization error for both the LS and forgetting factor LS. The smaller the quantization error is, the smaller the estimation error is.

6 Conclusion

This paper considers the parameter estimation problem of DARMA systems by using uniform quantized data. The LS and the forgetting factor LS are introduced to estimate system parameters. Under some conditions, we show that the estimation errors are bounded and dependent on the size of the quantization error. Here we only consider the case without system noises. For the systems with stochastic noise case, we guess the LS algorithm can still be applicable, but the analysis would be much more complex and difficult. As for further research, the adaptive control of quantized linear systems is worth investigation, which may need to relax the conditions of this paper or take a new way to study.

References

[1] Wang L Y, Zhang J F, and Yin G, System identification using binary sensors, *IEEE Transactions on Automatic Control*, 2003, **48**(11): 1892–1907.

[2] Yin G, Wang L Y, and Kan S, Tracking and identification of regime-switching systems using binary sensors, *Automatica*, 2009, **45**(4): 944–955.

[3] Guo J and Zhao Y L, Recursive projection algorithm on FIR system identification with binary-valued observations, *Automatica*, 2013, **49**(11): 3396–3401.

[4] Zhao Y L, Bi W J, and Wang T, Iterative parameter estimate with batched binary-valued observations, *SCIENCE CHINA Information Sciences*, 2016, **59**(5): 052201.

[5] Godoya B I, Goodwin G C, Agüero J C, et al., On identification of FIR systems having quantized output data, *Automatica*, 2011, **47**(9): 1905–1915.

[6] Jing L D and Zhang J F, Tracking control and parameter identification with quantized ARMAX systems, *SCIENCE CHINA Information Sciences*, 2019, **62**(9): 199203.

[7] Wang Y and Zhang J F, Distributed parameter identification of quantized ARMAX systems, *Proceedings of the 38th Chinese Control Conference*, Guangzhou, China, 2019, 1701–1706.

[8] Guo J and Diao J D, Prediction-based event-triggered identification of quantized input FIR systems with quantized output observations, *SCIENCE CHINA Information Sciences*, 2020, **63**(1): 112201.

[9] Diao J D, Guo J, and Sun C Y, A compensation method for the packet loss deviation in system identification with event-triggered binary-valued observations, *SCIENCE CHINA Information Sciences*, 2020, **63**(12): 229204.

[10] Wang X M, Hu M, Zhao Y L, et al., Credit scoring based on the set-valued identification method, *Journal of Systems Science and Complexity*, 2020, **33**(5): 1297–1309.

[11] Chen H F and Guo L, Adaptive control via consistent estimation for deterministic systems, *International Journal of Control*, 1987, **45**(6): 2183–2202.

[12] Young P C, Recursive approaches to time series analysis, *Bull. Inst. Math. Applic*, 1974, **10**: 209–224.

[13] Åström K J and Wittenmark B, On self tuning regulars, *Automatica*, 1973, **9**: 185–199.

[14] Chen H F and Zhang J F, Identification of linear systems without assuming stability and minimum-phase, *Science in China* (*Series A*), 1990, **33**(6): 641–653.

[15] Chen H F and Zhang J F, On identifiability for multidimensional ARMAX model, *Acta Mathematicae Applicatae Sinica*, 1993, **9**(1): 1–8.

[16] Chen H F, Quasi-least squares identification and its strong consistency, *International Journal of Control*, 1981, **34**(5): 921–936.

[17] Lai T L and Wei C Z, Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems, *The Annals of Statistics*, 1982, **10**(1): 154–166.

[18] Lai T L and Wei C Z, Extended least squares and their applications to adaptive control and prediction in linear systems, *IEEE Transactions on Automatic Control*, 1986, **AC-31**(10): 898–906.

[19] Chen H F and Guo L, Convergence rate of least-squares identification and adaptive control for stochastic systems, *International Journal of Control*, 1986, **44**(5): 1459–1476.

[20] Chen H F, Convergence rate of continuous time ELS parameter estimation, *IEEE Transactions on Automatic Control*, 1987, **AC-32**(3): 267–269.

[21] Fortescue T R, Kershenbaum L S, and Ydstie B E, Implementation of self-tuning regulators with variable forgetting factors, *Automatica*, 1981, **17**(6): 831–835.

[22] Kulhavy R and Karny M, Tracking of slowly varying parameters by directional forgetting, *Proceedings of the* 9*th IFAC Word Conference*, Budapest, Hungary, 1984, 79–83.

[23] Guo L, Ljung L, and Priouret P, Performance analysis of the forgetting factor RLS algorithms, *International Journal of Adaptive Control and Signal Processing*, 1993, **7**(6): 525–537.

[24] Zhang Y Q and Guo L, Convergence of self-tuning regulators under conditional heteroscedastic noises with unknown high-frequency gain, *Journal of Systems Science and Complexity*, 2021, **34**(1): 236–250.

[25] Li C L, Parameter estimation algorithm of generalized linear model for big data, *Journal of Systems Science and Mathematical Sciences*, 2020, **40**(5): 927–940 (in Chinese).

[26] Wang Z H, Tian M Z, and Hou Z M, Measuring quantile effects based on partial least squares path method, *Journal of Systems Science and Mathematical Sciences*, 2020, 40(4): 728–750 (in Chinese).