Minimax Estimation of Uncertain Systems in the Presence of Bounded Disturbance

Wataru KITAMURA* and Yasumasa FUJISAKI**

Minimax estimation is considered for a single-input single-output discrete-time uncertain system in the presence of bounded disturbance. The given regressors are divided into two sets which have small and large amplitudes respectively, where the amplitude ranges are assumed to be exclusive each other. Then, the nominal parameter of the system is estimated so that the maximal output error is minimized. The bounds of the disturbance and the parameter uncertainty are also estimated by using the output errors for these two sets. For this minimax estimation, the estimation errors are evaluated when the regressors of each set are persistently exciting. Furthermore, probabilistic estimation errors are derived when the regressors of each set are persistently exciting and periodic and have the same amplitude, and the disturbance and the parameter uncertainty are random variables which take their extreme values with a probability. The result implies that the errors converge to zero as the number of samples tends to infinity.

Key Words: Minimax estimation, uncertain systems, bounded disturbance.

1. Introduction

Minimax estimation is frequently used in identification of a system parameter in the presence of unknown bounded disturbance $^{1)\sim 3)}$. This estimation enables us to obtain not only an estimate of the system parameter that describes measured input-output data but also an estimate of the upper bound of the disturbance. Then, the estimates converge to the true values as the number of samples tends to infinity if the regressor is assumed to be persistently exciting and the disturbance is assumed to be random variable which takes its extreme value with a probability⁴⁾. However, the existing literature deals with only a system with additive bounded disturbance, i.e., parameter uncertainty of the system has not been introduced. It should be noticed that the existing estimation algorithm cannot be used for an uncertain system with a parameter uncertainty. This is because the parameter uncertainty causes an output error of the system not as an additive bounded noise but as an unbounded noise whose amplitude depends on the magnitude of input signals.

In this paper, we investigate minimax estimation for uncertain systems in the presence of unknown bounded disturbance. We here propose an identification method which gives not only an estimate of the nominal parameter of the system but also estimates of the upper bounds of bounded disturbance and bounded parameter uncertainty. The given regressors are divided into two sets which have small and large amplitudes respectively, where the amplitude ranges are assumed to be exclusive each other. Then, the maximal output error is minimized so that the nominal parameter of the system is estimated. The bounds of the disturbance and the parameter uncertainty are also estimated by using the output errors for these two sets. In order to evaluate performance of this minimax estimation, we derive deterministic bounds of the estimation errors when the regressors of each set are persistently exciting. Furthermore, we also derive probabilistic bounds of the estimation errors for finite number of samples when the regressors of each set are persistently exciting and periodic and have the same amplitude, and the disturbance and the parameter uncertainty are random variables which take their extreme values with a probability. Then, we prove that the errors converge to zero as the number of samples tends to infinity. It should be noted that, even if there is no parameter uncertainty of the system, the results of this paper give some useful details relative to the existing literature⁴⁾ which explored convergence property of the estimation errors since the results here give an evaluation of the errors with respect to number of samples.

2. Minimax estimation

Let us describe a single-input single-output (SISO) discrete time system as

$$y_i = \phi_i^{\mathrm{T}}(\theta + \eta_i) + v_i \tag{1}$$

where subscript *i* means the time, $y_i \in \mathbb{R}$ is the sys-

^{*} Graduate School of Science and Technology, Kobe University (Currently, Logistics & Automation Division, Engineering Department, Murata Machinery, Ltd.)

^{**} Department of Computer Science and Systems Engineering, Kobe University

53

tem output, $\phi_i \in \mathbb{R}^m$ is the measurable regression vector, $\theta \in \mathbb{R}^m$ is a nominal parameter of the system to be identified, $v_i \in \mathbb{R}$ is an unknown disturbance, and $\eta_i \in \mathbb{R}^m$ is an unknown parameter uncertainty. The disturbance and the parameter uncertainty are bounded, that is,

$$|v_i| \le \epsilon, \qquad \|\eta_i\|_2 \le \delta \tag{2}$$

where the bounds ϵ and δ are unknown, and $\| \bullet \|_2$ is the l_2 norm

$$\|\eta_i\|_2 \doteq \left(\eta_i^{\mathrm{T}} \eta_i\right)^{1/2}.$$

In this paper, we consider an identification problem which is to find estimates of the nominal parameter of the system θ , the bound of disturbance ϵ , and the bound of parameter uncertainty δ based on given input-output data $\{y_i, \phi_i\}$. In the following, we derive the estimates in the framework of minimax estimation.

Note that, if no parameter uncertainty of the system is presented in (1), the minimax estimate of θ is given by $\hat{\theta} \doteq \arg \min_{\hat{\theta}} \max_i |y_i - \phi_i^{\mathrm{T}} \hat{\theta}|$. Then, it turns out that an estimate ϵ is given by the maximum value of output error $\max_i |y_i - \phi_i^{\mathrm{T}} \hat{\theta}|^{-1) \sim 3}$. However, if there exists parameter uncertainty of the system, both upper bounds ϵ and δ have to be estimated. One value such as the maximum value of output error is clearly not sufficient for determining two values ϵ and δ .

In order to estimate both upper bounds ϵ and δ , we first divide given input-output data $\{y_i, \phi_i\}$ into two sets, which are described as $\{y_i^{(1)}, \phi_i^{(1)}\}$ and $\{y_i^{(2)}, \phi_i^{(2)}\}$. Here we classify $\{y_i, \phi_i\}$ into $\{y_i^{(1)}, \phi_i^{(1)}\}$ and $\{y_i^{(2)}, \phi_i^{(2)}\}$ by amplitude of the regressor, that is,

$$a_{1} \leq \|\phi_{i}^{(1)}\|_{2} \leq b_{1},$$

$$a_{2} \leq \|\phi_{i}^{(2)}\|_{2} \leq b_{2},$$

$$0 < a_{1} \leq b_{1} < a_{2} \leq b_{2} < \infty$$
(3)

where $\{y_i^{(1)}, \phi_i^{(1)}\}$ and $\{y_i^{(2)}, \phi_i^{(2)}\}$ contain N_1 and N_2 input-output data respectively.

Then, we obtain the minimax estimate $\hat{\theta}$ of θ as

$$\widehat{\theta}_{j} \doteq \arg\min_{\widehat{\theta}} \max_{i} |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}}\widehat{\theta}|, \qquad j = 1, 2 \qquad (4)$$

by each input-output data set $\{y_i^{(j)},\phi_i^{(j)}\}, j=1,2.$ Defining the maximum output error as

$$e_j \doteq \max_i |y_i^{(j)} - (\phi_i^{(j)})^{\mathrm{T}} \widehat{\theta}_j|, \quad j = 1, 2,$$
 (5)

we can obtain estimates $\hat{\epsilon}, \hat{\delta}$ of ϵ, δ as

$$\widehat{\epsilon} \doteq \frac{b_2 e_1 - b_1 e_2}{b_2 - b_1} \tag{6}$$

$$\widehat{\delta} \doteq \frac{e_2 - e_1}{b_2 - b_1}.\tag{7}$$

The identification proposed in this paper is a method using the equations (4), (6) and (7). In the following sections, we investigate the estimation errors between the estimates $\hat{\theta}_j$, $\hat{\epsilon}$ and $\hat{\delta}$ and the true values θ , ϵ and δ , which shows that the estimates $\hat{\theta}_j$, $\hat{\epsilon}$ and $\hat{\delta}$ are actually relevant.

Notice that the estimates $\hat{\theta}_j$ of θ can be computed by solving a linear programming problem

$$\min_{\hat{\theta}, \nu_j} \quad \nu_j$$
s.t.
$$|y_i^{(j)} - (\phi_i^{(j)})^{\mathrm{T}} \hat{\theta}| \le \nu_j, \qquad \forall i = 1, 2, \dots, N_j \quad (8)$$

with each input-output data set $\{y_i^{(j)}, \phi_i^{(j)}\}, j = 1, 2.$

3. Deterministic analysis of the estimation errors

In this section, we derive upper bounds of the estimation errors in a deterministic setting.

Let us introduce a definition $^{5)}$.

Definition 1. The regressor ϕ_i , i = 1, 2, ..., N is said to be persistently exciting (PE) if there exist some $n \in \mathbb{N}, \alpha \in \mathbb{N}$, and $\beta \in \mathbb{N}$ such that

$$\alpha^2 I \le \sum_{i=i_0}^{i_0+n-1} \phi_i \phi_i^{\mathrm{T}} \le \beta^2 I \tag{9}$$

for any $i_0 \in \mathbb{N}$, $1 \leq i_0 \leq N - n + 1$. Here, \mathbb{N} denotes the set of positive integers and the matrix inequality $A \leq B$ means that B - A is positive semidefinite.

Note that the regressor ϕ_i can be PE regardless of whether it is a deterministic or a stochastic vector. If the regressor is PE, $\hat{\theta}$ of (8) is bounded for all ν_j , and thus the estimation error $\theta - \hat{\theta}_j$ is bounded.

In the following, we assume that each set of the regressors $\{\phi_i^{(1)}\}\$ and $\{\phi_i^{(2)}\}\$ satisfies PE condition (9). Then, the indexes of PE condition are written as (n_1, α_1, β_1) and (n_2, α_2, β_2) for $\{\phi_i^{(1)}\}\$ and $\{\phi_i^{(2)}\}\$ respectively, where $N_j = \ell_j n_j$ and ℓ_j is an integer for simplicity.

Then, we obtain the following theorem.

Theorem 1. Consider the system (1). Assume that the regressor $\phi_i^{(j)}$, j = 1, 2 is deterministic and PE. Then,

$$\|\theta - \widehat{\theta}_j\|_2 \le \frac{2\sqrt{n_j}}{\alpha_j} (\epsilon + b_j \delta), \qquad j = 1, 2.$$
 (10)

Proof. In the following, j is fixed as 1 or 2 since the proof is the same. Let us describe Φ_N^j as

$$\Phi_{N}^{j} \doteq \begin{bmatrix} (\phi_{1}^{(j)})^{\mathrm{T}} \\ (\phi_{2}^{(j)})^{\mathrm{T}} \\ \vdots \\ (\phi_{N_{j}}^{(j)})^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{N_{j} \times m}$$

Recall that the regressor $\phi_i^{(j)}$ is PE and $N_j = \ell_j n_j$. Then,

we have

$$(\Phi_N^j)^{\mathrm{T}} \Phi_N^j = \sum_{i=1}^{\ell_j n_j} \phi_i^{(j)} (\phi_i^{(j)})^{\mathrm{T}} \ge \alpha_j^2 \ell_j I_j$$

from the left inequality of (9). Hence, we obtain

$$\|\{(\Phi_N^j)^{\mathrm{T}}\Phi_N^j\}^{-1}(\Phi_N^j)^{\mathrm{T}}\|_{i2} \le \frac{1}{\alpha_j\sqrt{\ell_j}}$$
(11)

where $\| \bullet \|_{i2}$ is the induced l_2 norm

$$||A||_{i2} \doteq \sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} = \sqrt{\overline{\lambda}(AA^{\mathrm{T}})}$$

and $\overline{\lambda}(AA^{\mathrm{T}})$ is the maximum eigenvalue of AA^{T} for real matrix A. Noting that the estimate $\hat{\theta}$ of θ is obtained from (4), we have

$$\begin{aligned} y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}} \widehat{\theta}_{j} &| \leq \max_{i} |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}} \widehat{\theta}_{j} | \\ &= \min_{\widehat{\theta}} \max_{i} |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}} \widehat{\theta} | \\ &\leq \max_{i} |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}} \theta |. \end{aligned}$$
(12)

On the other hand, from (1), (2), and (3), we have

$$|y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}}\theta| = |v_{i} + (\phi_{i}^{(j)})^{\mathrm{T}}\eta_{i}|$$

$$< \epsilon + \delta b_{j}.$$
(13)

From (12) and (13), we see that

$$|y_i^{(j)} - (\phi_i^{(j)})^{\mathrm{T}}\widehat{\theta}_j| \le \max_i |y_i^{(j)} - (\phi_i^{(j)})^{\mathrm{T}}\theta|$$
$$\le \epsilon + \delta b_j.$$
(14)

Using these inequalities (13) and (14), we see that

$$\begin{aligned} |(\phi_{i}^{(j)})^{\mathrm{T}}(\theta - \widehat{\theta}_{j})| &= |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}}\widehat{\theta}_{j} - \{y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}}\theta\}| \\ &\leq |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}}\widehat{\theta}_{j}| + |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}}\theta| \\ &\leq 2(\epsilon + \delta b_{j}) \end{aligned}$$
(15)

holds for any i = 1, 2, ..., N. This implies

$$\|\Phi_N^j(\theta - \widehat{\theta}_j)\|_2 \le 2\sqrt{\ell_j n_j} (\epsilon + \delta b_j)$$

With (11), we conclude

$$\begin{aligned} \|\theta - \widehat{\theta}_{j}\|_{2} &\leq \|\{(\Phi_{N}^{j})^{\mathrm{T}} \Phi_{N}^{j}\}^{-1} (\Phi_{N}^{j})^{\mathrm{T}}\|_{i2} \|\Phi_{N}^{j}(\theta - \widehat{\theta}_{j})\|_{2} \\ &\leq \frac{2\sqrt{n_{j}}}{\alpha_{j}} (\epsilon + \delta b_{j}). \end{aligned}$$
(16)

This completes the proof.

Theorem 1 shows that the upper bound of the estimation error of θ is linear with respect to ϵ and δ . That is, $\hat{\theta}_1 = \hat{\theta}_2 = \theta$ if $\epsilon = \delta = 0$. We therefore see that the proposed minimax estimation is relevant.

Note that the diameter of the membership set in the presence of disturbance and parameter uncertainty is discussed in Theorem 1⁶⁾, where it is shown that there exist input-output data $\{y_i, \phi_i\}$ such that the diameter does not converge to zero even if the number of samples increases. The upper bound of the estimation error in Theorem 1

also does not depend on the number of samples, which is consist with the previous result $^{6)}$.

We further obtain a similar result on ϵ and $\delta.$

Theorem 2. Consider the system (1). Assume that the regressor $\phi_i^{(j)}$, j = 1, 2 is deterministic and PE. Then,

$$\frac{b_1(\epsilon + b_2\delta)}{b_2 - b_1} \le \epsilon - \hat{\epsilon} \le \frac{b_2(\epsilon + b_1\delta)}{b_2 - b_1}$$
(17)
$$-\frac{\epsilon + b_1\delta}{b_2 - b_1} \le \delta - \hat{\delta} \le \frac{\epsilon + b_2\delta}{b_2 - b_1}.$$
(18)

Proof. Using the inequalities (12) and (13), the maximum output error e_j is evaluated as

$$e_{j} = \max_{i} |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}} \widehat{\theta}_{j}|$$

$$\leq \max_{i} |y_{i}^{(j)} - (\phi_{i}^{(j)})^{\mathrm{T}} \theta|$$

$$\leq \epsilon + \delta b_{j}, \qquad j = 1, 2.$$
(19)

That is,

$$0 \le e_j \le \epsilon + \delta b_j, \qquad j = 1, 2.$$

Then, we obtain

$$\begin{split} \widehat{\epsilon} &\geq -\frac{b_1e_2}{b_2 - b_1} \geq -\frac{b_1(\epsilon + \delta b_2)}{b_2 - b_1} \\ \widehat{\epsilon} &\leq \frac{b_2e_1}{b_2 - b_1} \leq \frac{b_2(\epsilon + \delta b_1)}{b_2 - b_1} \\ \widehat{\delta} &\geq -\frac{e_1}{b_2 - b_1} \geq -\frac{\epsilon + \delta b_1}{b_2 - b_1} \\ \widehat{\delta} &\leq \frac{e_2}{b_2 - b_1} \leq \frac{\epsilon + \delta b_2}{b_2 - b_1}. \end{split}$$

Evaluating $\epsilon - \hat{\epsilon}$, $\delta - \hat{\delta}$ with the above bounds, we obtain the results (17) and (18).

Theorem 2 shows that the lower and upper bounds of the estimation errors of ϵ and δ are linear with respect to ϵ and δ , which is similar to the situation on θ in Theorem 1.

4. Probabilistic analysis of the estimation errors

In this section, we derive upper bounds of the estimation errors in a stochastic setting. Let us first introduce the following definitions on the disturbance $^{4)}$ and the parameter uncertainty $^{6)}$.

Definition 2. Suppose that the disturbance v_i is a random variable satisfying $|v_i| \leq \epsilon$. The bound ϵ is said to be tight if for any $\rho > 0$ and each *i*, there exists some $p_v(\rho) > 0$ such that

$$\operatorname{Prob}\{-\epsilon \le v_i \le -(\epsilon - \rho)\} \ge p_v(\rho)$$
$$\operatorname{Prob}\{\epsilon - \rho \le v_i \le \epsilon\} \ge p_v(\rho)$$

where $Prob\{\bullet\}$ is the probability that the event \bullet occurs.

Definition 3. Suppose that the parameter uncertainty η_i is a random vector satisfying $\|\eta_i\|_2 \leq \delta$. The bound δ is said to be tight if for any $\mu > 0$, each *i*, and any ϕ_i , there exists some $p_\eta(\mu) > 0$ such that

$$\begin{aligned} &\operatorname{Prob}\{-\delta \|\phi_i\|_2 \le \phi_i^{\mathrm{T}} \eta_i \le -(\delta-\mu) \|\phi_i\|_2\} \ge p_\eta(\mu) \\ &\operatorname{Prob}\{(\delta-\mu) \|\phi_i\|_2 \le \phi_i^{\mathrm{T}} \eta_i \le \delta \|\phi_i\|_2\} \ge p_\eta(\mu). \end{aligned}$$

Definitions 2 and 3 introduce stochastic properties into the disturbance v_i and the parameter uncertainty η_i , and the tightness means that v_i and η_i take around their extreme values with nonzero probability.

Then, the next lemma is obtained $^{6)}$.

Lemma 1. Assume that the disturbance v_i and the parameter uncertainty η_i are independent random variables, and their bounds ϵ and δ are tight. Then, for any $\rho > 0$, $\mu > 0$, each i, and any ϕ_i ,

$$\operatorname{Prob}\left\{-(\epsilon+\delta\|\phi_i\|_2) \leq v_i + \phi_i^{\mathrm{T}}\eta_i\right\}$$
$$\leq -(\epsilon-\rho) - (\delta-\mu)\|\phi_i\|_2\right\}$$
$$\geq p_v(\rho)p_\eta(\mu) > 0$$
$$\operatorname{Prob}\left\{\epsilon-\rho + (\delta-\mu)\|\phi_i\|_2 \leq v_i + \phi_i^{\mathrm{T}}\eta_i \leq \epsilon + \delta\|\phi_i\|_2\right\}$$
$$\geq p_v(\rho)p_\eta(\mu) > 0.$$

This lemma means that, if both of the bounds ϵ and δ are tight, the bound $\epsilon + \delta \|\phi_i\|_2$ of $v_i + \phi_i^{\mathrm{T}} \eta_i$ is also tight.

In the following, for simplicity, we assume that the regressor $\phi_i^{(j)}$ is n_j $(n_j \ge m)$ periodic and $N_j = \ell_j n_j$, i.e.,

$$\phi_{i_p+kn_j}^{(j)}=\phi_{i_p}^{(j)}$$

where $i_p = 1, 2, ..., n_j$, $k = 0, 1, 2, ..., \ell_j - 1$, $\ell_j \ge 1$. We also assume that $a_j = b_j$, j = 1, 2. Note that, if the regressors $\phi_{i_p}^{(1)}$ and $\phi_{i_p}^{(2)}$ are PE, we can set $n = n_j$ in (9) without loss of generality, and there exists α_j such that

$$\alpha_j^2 I \le (\Phi_n^j)^{\mathrm{T}} \Phi_n^j, \qquad j = 1,2 \tag{20}$$

where

$$\Phi_n^j \doteq \begin{bmatrix} (\phi_1^{(j)})^{\mathrm{T}} \\ (\phi_2^{(j)})^{\mathrm{T}} \\ \vdots \\ (\phi_{n_j}^{(j)})^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{n_j \times m}.$$

Then, we obtain the following theorem.

Theorem 3. Consider the system (1). Assume that the regressor $\phi_i^{(j)}$ is deterministic, periodic, PE, and $a_1 = b_1, a_2 = b_2$. Furthermore, v_i and η_i are independent random variables, and their bounds are tight. Then, the estimation error of θ is evaluated as

$$\operatorname{Prob}\left\{\|\theta - \widehat{\theta}_j\|_2 \le \frac{\sqrt{n_j}}{\alpha_j}(\rho + b_j\mu)\right\} \ge c_{p_j}, \qquad j = 1, 2$$
(21)

where c_{p_j} is a positive constant satisfying

$$c_{p_j} \le \left\{ 1 - 2 \left(1 - p_v(\rho) p_\eta(\mu) \right)^{\ell_j} \right\}^{n_j}, \qquad j = 1, 2.$$
 (22)

Proof. In the following, j is fixed as 1 or 2 since the proof is the same. We first fix i_p . Then, we have

$$(\phi_{i_p}^{(j)})^{\mathrm{T}} (\theta - \widehat{\theta}_j)$$

= $y_{i_p+kn_j}^{(j)} - (\phi_{i_p}^{(j)})^{\mathrm{T}} \widehat{\theta}_j - \{ v_{i_p+kn_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p+kn_j} \}$

from (1). Since $\hat{\theta}_j$ is obtained by (4), with (14), we have

$$-(\epsilon + \delta b_j) \le y_{i_p + kn_j}^{(j)} - (\phi_{i_p}^{(j)})^{\mathrm{T}} \widehat{\theta}_j \le \epsilon + \delta b_j$$

for any k. If there exists k_u satisfying

$$\epsilon - \rho + (\delta - \mu) \|\phi_{i_p}^{(j)}\|_2 \le v_{i_p + k_u n_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p + k_u n_j} \le \epsilon + \delta \|\phi_{i_p}^{(j)}\|_2,$$
(23)

then, with $\|\phi_{i_p}^{(j)}\|_2 = b_j$, we obtain

$$(\phi_{i_p}^{(j)})^{\mathrm{T}}(\theta - \widehat{\theta}_j) \le \rho + \mu b_j.$$

On the other hand, if there exists k_l satisfying

$$-\epsilon - \delta \|\phi_{i_p}^{(j)}\|_2 \le v_{i_p+k_l n_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p+k_l n_j} \le -(\epsilon - \rho) - (\delta - \mu) \|\phi_{i_p}^{(j)}\|_2, \qquad (24)$$

then, in the same way, we have

$$-(\rho+\mu b_j) \le (\phi_{i_p}^{(j)})^{\mathrm{T}}(\theta-\widehat{\theta}_j).$$

Hence, if there exist both of such k_u and k_l , we have

$$|(\phi_{i_p}^{(j)})^{\mathrm{T}}(\theta - \widehat{\theta}_j)| \le \rho + \mu b_j.$$

$$(25)$$

We therefore see that, for some fixed i_p ,

$$\operatorname{Prob}\left\{ |(\phi_{i_{p}}^{(j)})^{\mathrm{T}}(\theta - \widehat{\theta}_{j})| \leq \rho + \mu b_{j} \right\}$$
$$\geq 1 - 2 \left(1 - p_{v}(\rho)p_{\eta}(\mu)\right)^{\ell_{j}}.$$
(26)

This is because Lemma 1 says that the probability that $v_{i_p+kn_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p+kn_j}$ does not satisfy (23) and (24) for all $k = 0, 1, \ldots, \ell_1 - 1$ is less than $(1 - p_v(\rho)p_\eta(\mu))^{\ell_j}$. Now, if (25) satisfies for each $i_p = 1, 2, \ldots, n_j$, we have

$$\|\Phi_n^j(\theta - \widehat{\theta}_j)\|_2 \le \sqrt{n_j}(\rho + \mu b_j).$$

Then, following an evaluation similar to (16), we see

$$\begin{aligned} \|\theta - \widehat{\theta}_1\|_2 &\leq \|\{(\Phi_n^j)^{\mathrm{T}} \Phi_n^j\}^{-1} (\Phi_n^j)^{\mathrm{T}}\|_{i2} \|\Phi_n^j (\theta - \widehat{\theta}_j)\|_2 \\ &\leq \frac{\sqrt{n_j}}{\alpha_j} (\rho + \mu b_j) \end{aligned}$$

where $\|\{(\Phi_n^j)^{\mathrm{T}}\Phi_n^j\}^{-1}(\Phi_n^j)^{\mathrm{T}}\|_{i2} \leq 1/\alpha_j$. We therefore obtain the theorem since the probability that (26) holds for all $i_p = 1, 2, \ldots, n_j$ is $\{1 - 2(1 - p_v(\rho)p_\eta(\mu))^{\ell_j}\}^{n_j}$.

Theorem 3 shows that the evaluation of the estimation error of θ is linear with respect to the indexes ρ and μ of tightness. That is, the estimation error depends on how close the disturbance and the parameter uncertainty take their extreme value in the data used for identification. The error is independent of the upper bounds of the disturbance and the parameter uncertainty. We then discuss the estimation errors of ϵ and δ . To this end, we need the following lemma.

Lemma 2. Under the same assumptions in Theorem 3,

Prob
$$\{\epsilon - 2\rho + (\delta - 2\mu)b_j \le e_j \le \epsilon + \delta b_j\} \ge c_{p_j},$$

 $j = 1, 2.$ (27)

Proof. In the following, j is fixed as 1 or 2 since the proof is the same. The upper bound of e_j is obvious from (19). Thus, we only discuss the lower bound of e_j . We first fix i_p . For the system (1),

$$\begin{aligned} v_{i_{p}+kn_{j}} &+ (\phi_{i_{p}}^{(j)})^{\mathrm{T}} \eta_{i_{p}+kn_{j}} \\ &= \{ y_{i_{p}+kn_{j}}^{(j)} - (\phi_{i_{p}}^{(j)})^{\mathrm{T}} \widehat{\theta}_{j} \} - (\phi_{i_{p}}^{(j)})^{\mathrm{T}} (\theta - \widehat{\theta}_{j}). \end{aligned}$$

Hence, we see

$$\begin{aligned} |y_{i_p+kn_j}^{(j)} - (\phi_{i_p}^{(j)})^{\mathrm{T}} \widehat{\theta}_j| \\ &\geq |v_{i_p+kn_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p+kn_j}| - |(\phi_{i_p}^{(j)})^{\mathrm{T}} (\theta - \widehat{\theta}_j)| \end{aligned}$$

Then, we have

$$\max_{k} |y_{i_{p}+kn_{j}}^{(j)} - (\phi_{i_{p}}^{(j)})^{\mathrm{T}} \widehat{\theta}_{j}| \\ \geq \max_{k} |v_{i_{p}+kn_{j}} + (\phi_{i_{p}}^{(j)})^{\mathrm{T}} \eta_{i_{p}+kn_{j}}| - |(\phi_{i_{p}}^{(j)})^{\mathrm{T}} (\theta - \widehat{\theta}_{j})|.$$
(28)

If there exist both of k_u and k_l satisfying (23) and (24), we see that the inequality (25) and

$$(\epsilon - \rho) + (\delta - \mu) \|\phi_{i_p}^{(j)}\|_2 \le |v_{i_p + kn_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p + kn_j}|$$

hold. Thus, we have

$$(\epsilon - \rho) + (\delta - \mu) \|\phi_{i_p}^{(j)}\|_2$$

$$\leq \max_k |v_{i_p + kn_j} + (\phi_{i_p}^{(j)})^{\mathrm{T}} \eta_{i_p + kn_j}|.$$
(29)

Substituting these (25) and (29) into (28), we obtain

$$\epsilon - 2\rho + (\delta - 2\mu) \|\phi_{i_p}^{(j)}\|_2 \le \max_k |y_{i_p + kn_j}^{(j)} - (\phi_{i_p}^{(j)})^{\mathrm{T}} \widehat{\theta}_j|.$$

Note that this inequality holds if there exist both k_u and k_l satisfying (23) and (24). Thus, for fixed i_p , we have

$$\operatorname{Prob}\left\{\epsilon - 2\rho + (\delta - 2\mu)b_j \leq \max_k |y_{i_p + kn_j}^{(j)} - (\phi_{i_p}^{(j)})^{\mathrm{T}}\widehat{\theta}_j|\right\}$$
$$\geq 1 - 2(p_v(\rho)p_\eta(\mu))^{\ell_j}.$$

If the above inequality holds for all $i_p = 1, 2, ..., n_j$, from a representation of the maximum output error

$$e_j = \max_{i_p} \max_k |y_{i_p+kn_j}^{(j)} - (\phi_{i_p}^{(j)})^{\mathrm{T}} \widehat{\theta}_j|$$

we see that

$$\epsilon - 2\rho + (\delta - 2\mu)b_j \le e_j \tag{30}$$

which holds with probability $\{1 - 2(p_v(\rho)p_\eta(\mu))^{\ell_j}\}^{n_j}$. This completes the proof.

Then, we obtain the following theorem on the estimation errors of ϵ and δ . **Theorem 4.** Consider the system (1). Assume that the regressor $\phi_i^{(j)}$ is deterministic, periodic, PE, and $a_1 = b_1, a_2 = b_2$. Furthermore, v_i and η_i are independent random variables, and their bounds are tight. Then, the estimation errors of ϵ and δ are evaluated as

$$\operatorname{Prob}\left\{-\frac{2b_{1}(\rho+b_{2}\mu)}{b_{2}-b_{1}} \leq \epsilon - \widehat{\epsilon} \leq \frac{2b_{2}(\rho+b_{1}\mu)}{b_{2}-b_{1}}\right\} \geq c_{p_{1}}c_{p_{2}}$$
(31)
$$\operatorname{Prob}\left\{-\frac{2(\rho+b_{1}\mu)}{b_{2}-b_{1}} \leq \delta - \widehat{\delta} \leq \frac{2(\rho+b_{2}\mu)}{b_{2}-b_{1}}\right\} \geq c_{p_{1}}c_{p_{2}}.$$

Proof. When the inequality of e_j (30) of Lemma 2 holds for j = 1, 2, the estimate $\hat{\epsilon}$ can be evaluated as

$$\begin{split} \widehat{\epsilon} &\geq \frac{b_2 \{ \epsilon - 2\rho + (\delta - 2\mu) b_1 \} - b_1 (\epsilon + \delta b_2)}{b_2 - b_1} \\ &= \epsilon - \frac{2b_2 (\rho + b_1 \mu)}{b_2 - b_1} \\ \widehat{\epsilon} &\leq \frac{b_2 (\epsilon + \delta b_1) - b_1 \{ \epsilon - 2\rho + (\delta - 2\mu) b_2 \}}{b_2 - b_1} \\ &= \epsilon + \frac{2b_1 (\rho + b_2 \mu)}{b_2 - b_1}. \end{split}$$

In the same way, the estimate $\hat{\delta}$ can also be evaluated as

$$\begin{split} \widehat{\delta} &\geq \frac{\epsilon - 2\rho + (\delta - 2\mu)b_2 - (\epsilon + \delta b_1)}{b_2 - b_1} \\ &= \delta - \frac{2(\rho + b_2\mu)}{b_2 - b_1} \\ \widehat{\delta} &\leq \frac{\{\epsilon + \rho + (\delta + \mu)b_2\} - \{\epsilon - 2\rho + (\delta - 2\mu)b_1\}}{b_2 - b_1} \\ &= \delta + \frac{2(\rho + b_1\mu)}{b_2 - b_1}. \end{split}$$

Using these inequalities, we have the estimation errors in (31) and (32). Note that these inequalities require that the evaluation of e_j of Lemma 2 holds for both j = 1, 2. These events are independent each other, and thus the probabilities of the theorem are $c_{p_1}c_{p_2}$.

Theorem 4 shows that the estimation errors of ϵ and δ are linear with respect to the indexes ρ and μ of tightness, which is similar to the case of θ .

If we choose the confidence c_{p_j} in Theorems 3 and 4 as

$$c_{p_j} = \left\{ 1 - 2 \left(1 - p_v(\rho) p_\eta(\mu) \right)^{\ell_j} \right\}^{n_j}, \quad j = 1, 2$$

we see that, for any ρ and $\mu,$

 $c_{p_1} \to 1, \quad c_{p_2} \to 1 \qquad \text{as} \quad \ell_j \to \infty, \quad j = 1, 2.$

This leads to the following corollary.

Corollary 1. Under the same assumptions in Theorems 3 and 4, the estimates $\hat{\theta}_j$, $\hat{\epsilon}$, and $\hat{\delta}$ converge to θ , ϵ , and δ in probability as $\ell_j \to \infty$, j = 1, 2.

The existing literature⁴⁾ shows that the estimates of θ and ϵ converge to their true values for a system without parameter uncertainty (i.e., $\delta = 0$). On the other hand,

this paper presents not only its counterpart for the case $\delta \neq 0$ as Corollary 1 but also quantitative evaluations of the estimation errors as Theorems 3 and 4.

In this section, we derive the probabilistic evaluations of the estimation errors under the assumptions that the regressor is not only PE but also periodic and $a_j = b_j$, j =1, 2. If the assumption $a_j = b_j$, j = 1, 2 does not hold, the estimate $\hat{\theta}_j$ is not uniquely determined from (8) (See more details in Appendix). Thus, we cannot expect that the estimation errors converge to zero. In other words, in order that the estimation errors converge to zero, it should be required that (1) is an FIR system, the regressor ϕ_i consists of identification input, and its amplitude can be tuned. On the other hand, the periodicity assumption can be removed if an involved discussion ^{7), 8)} is employed.

If we know some information about the distributions of the disturbance and the parameter uncertainty so that we can estimate $p_v(\rho)$, $p_\eta(\mu)$, Theorem 3 gives the necessary number of samples explicitly

$$\ell_j \ge \frac{\ln\left(\frac{1 - (c_{p_j})^{1/n_j}}{2}\right)}{\ln(1 - p_v(\rho)p_\eta(\mu))}, \qquad j = 1, 2$$
(33)

where ρ , μ and c_{p_j} are our specified values. A similar remark can be established for Theorem 4.

Note that Theorems 3 and 4 give not only probabilistic upper bounds of the estimation errors but also their deterministic upper bounds. In fact, setting $\rho = 2\epsilon$ and $\mu = 2\delta$ so that $c_{p_j} = 1$, we obtain

$$\begin{aligned} \|\theta - \widehat{\theta}_j\|_2 &\leq \frac{2\sqrt{n_j}}{\alpha_j} (\epsilon + b_j \delta), \quad j = 1, 2\\ -\frac{4b_1(\epsilon + b_2 \delta)}{b_2 - b_1} &\leq \epsilon - \widehat{\epsilon} \leq \frac{4b_2(\epsilon + b_1 \delta)}{b_2 - b_1}\\ -\frac{4(\epsilon + b_1 \delta)}{b_2 - b_1} &\leq \delta - \widehat{\delta} \leq \frac{4(\epsilon + b_2 \delta)}{b_2 - b_1}. \end{aligned}$$

Thus, the deterministic upper bound of θ from Theorem 3 is identical to that of Theorem 1, while the deterministic upper bounds of ϵ and δ from Theorem 4 is four times as conservative as that of Theorem 2.

5. Numerical examples

In this section, we give a numerical example. Let us consider a system

$$\begin{split} y_i^{(j)} &= (\phi_i^{(j)})^{\mathrm{T}}(\theta + \eta_i) + v_i, \\ \theta &= \begin{bmatrix} -1\\ 1 \end{bmatrix}, \phi_i^{(j)} = \begin{bmatrix} u_{1i}^{(j)}\\ u_{2i}^{(j)} \end{bmatrix}, \eta_i = \begin{bmatrix} \eta_{1i}\\ \eta_{2i} \end{bmatrix} \end{split}$$

where v_i and η_i are independent random variables with uniform distributions on $|v_i| \leq 1$ and $||\eta_i||_2 \leq 2$, i.e., $\epsilon = 1$

____ 1 < 0°.0.5 $\theta \parallel$ 0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000 (ψ 500 1000 1500 2000 2500 3000 4000 4500 5000 3500 $\langle \circ \rangle$ 0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000 The number of repetitions

Fig. 1 Estimation errors of θ , ϵ , and δ

and $\delta = 2$. The regressor is periodic with n = 3, that is

$$\begin{split} \phi_1^{(1)} &= \begin{bmatrix} \sqrt{3}/2\\ 1/2 \end{bmatrix}, \phi_2^{(1)} = \begin{bmatrix} 0\\ -1 \end{bmatrix}, \phi_3^{(1)} = \begin{bmatrix} -\sqrt{3}/2\\ 1/2 \end{bmatrix}, \\ \phi_1^{(2)} &= \begin{bmatrix} 0\\ 10 \end{bmatrix}, \phi_2^{(2)} = \begin{bmatrix} 5\sqrt{3}\\ -5 \end{bmatrix}, \phi_3^{(2)} = \begin{bmatrix} -5\sqrt{3}\\ -5 \end{bmatrix} \end{split}$$

where $a_1 = b_1 = 1$, $a_2 = b_2 = 10$. The indexes of PE for $\{\phi_i^{(1)}\}\$ and $\{\phi_i^{(2)}\}\$ are (3, 1.5, 1.5) and (3, 150, 150).

Fig. 1 shows the estimation errors of θ , ϵ , and δ at each 5000 steps from $\ell_j = 1$ to 5000, where the broken line represents the estimation errors of $\hat{\theta}_1$ and the solid line represents that of $\hat{\theta}_2$. After 5000 steps calculation, we obtained the estimates

$$\widehat{\theta}_1 = \begin{bmatrix} -0.9802\\ 1.0105 \end{bmatrix}, \quad \widehat{\theta}_2 = \begin{bmatrix} -1.0241\\ 1.0249 \end{bmatrix}$$
$$\widehat{\epsilon} = 0.9787, \qquad \widehat{\delta} = 1.9637.$$

The result shows that the estimation errors converge to zero as the number of samples increases, which is consistent with Corollary 1.

6. Concluding remarks

In this paper, we have investigated a minimax estimation of uncertain systems in the presence of disturbance and parameter uncertainty. The proposed identification method gives estimates of not only the system parameter but also the upper bound of the disturbance and the parameter uncertainty. We have studied these estimation errors and have shown that, under some conditions, the estimation errors converge to zero as the number of samples tends to infinity.

References

- E. Walter, and H. P.-Lahanier: Estimation of Parameter Bounds from Bounded Error Data- A Survey, *Mathemat*ics and Computers in Simulation, **32**, 449/468 (1990)
- 2) H. P.-Lahanier, and E. Walter: Exact Description of Feasible Parameter Sets and Minimax Estimation, *Int. Journal*

of Adaptive Control and Signal Processing, 8, 5/14 (1994)

- 3) M. Milanese, J. Norton, H. P.-Lahanier, and E. Walter: Bounding Approaches to System Identification, Plenum Press New York and London (1996)
- 4) E.-W. Bai, H. Cho, and R. Tempo: Convergence Properties of the Membership Set, Automatica, 34-10, 1245/1249 (1998)
- 5) S. Dasgupta and Y.-F. Huang: Asymptotically Convergent Modified Recursive Least Squares with Data Dependent Updating and Forgetting Factor for System with Bounded Noise, IEEE Trans. on Information Theory, 33-3, 383/392 (1987)
- 6) W. Kitamura and Y. Fujisaki: Convergence Properties of the Membership Set in the Presence of Disturbance and Parameter Uncertainty, Trans. of the Society of Instrument and Control Engineers, **39**-4,382/387 (2003)
- 7) H. Akçay: The Size of the Membership Set in a probabilistic framework, Automatica, 40-2, 253/260 (2004)
- 8) W. Kitamura and Y. Fujisaki: Stochastic Properties of the Membership Set in the Presence of Parameter Uncertainty - Nonperiodic Regressor Case -, SICE 3rd Annual Conference on Control Systems, 547/550 (2003)

Appendix A. An example that $\hat{\theta}_i$ is not determined uniquely

Through an example, we show that, if the condition $a_j = b_j, \ j = 1, 2$ does not hold, $\hat{\theta}_j$ is not determined uniquely even if the other assumptions in Corollary 1 are satisfied. Here, we fix j as 1 or 2 since the discussions are similar for both cases of $\hat{\theta}_1$ and $\hat{\theta}_2$.

Let us consider a system

$$y_i^{(j)} = (\phi_i^{(j)})^{\mathrm{T}}(\theta + \eta_i) + v_i, \quad \theta = \begin{bmatrix} 1\\1 \end{bmatrix}$$

where $|v_i| \leq 1$ and $||\eta_i||_2 \leq 1$, i.e., $\epsilon = \delta = 1$. The regressor is periodic with $n_i = 2$, that is,

$$\phi_i^{(j)} = \left[\begin{array}{c} 2\\ 0 \end{array} \right], \left[\begin{array}{c} 0\\ 1 \end{array} \right], \left[\begin{array}{c} 2\\ 0 \end{array} \right], \left[\begin{array}{c} 0\\ 1 \end{array} \right], \ldots$$

This is PE with $\alpha_j = 1$ and $\beta_j = 2$, where $a_j \neq b_j$ since $a_j = 1$ and $b_j = 2$. We consider 4 periodic v_i and η_i

$$v_i = -1, -1, 1, 1, \dots$$

$$\eta_i = -\begin{bmatrix} 1\\0 \end{bmatrix}, -\begin{bmatrix} 0\\1 \end{bmatrix}, \begin{bmatrix} 1\\0 \end{bmatrix}, \begin{bmatrix} 0\\1 \end{bmatrix}, \dots$$

Then, the output $y_i^{(j)}$ is 4 periodic as

$$y_i^{(j)} = -1, -1, 5, 3, \dots$$

Note that Corollary 1 is established by using a set of v_i and η_i satisfying both (23) and (24) with $\rho = \mu \simeq 0$, where the assumptions of tightness ensure the existence of such v_i and η_i . Now, for v_i and η_i given above, we have

$$v_i + (\phi_i^{(j)})^{\mathrm{T}} \eta_i = -3, -2, 3, 2, \dots,$$

= $-(\epsilon + \delta \| \phi_1^{(j)} \|_2), -(\epsilon + \delta \| \phi_2^{(j)} \|_2),$
 $\epsilon + \delta \| \phi_1^{(j)} \|_2, \epsilon + \delta \| \phi_2^{(j)} \|_2, \dots,$

which implies that (23) and (24) are satisfied with $\rho = \mu =$ 0. Thus, these v_i and η_i can correspond to the assumption of tightness. In fact, even if v_i and η_i are random variables whose bounds are tight, there exists no better value of $v_i + (\phi_i^{(j)})^{\mathrm{T}} \eta_i$ to improve the estimation error since this sequence of $v_i + (\phi_i^{(j)})^{\mathrm{T}} \eta_i$ takes all extreme value.

From (8), $\hat{\theta}_i$ is obtained by solving an LP

 $\theta.\nu$

$$\begin{split} \min_{\hat{\theta}, \nu_j} & \nu_j \\ \text{s.t.} & |-1 - 2\hat{\theta}_1| \le \nu_j \\ & |-1 - \hat{\theta}_2| \le \nu_j \\ & |5 - 2\hat{\theta}_1| \le \nu_j \\ & |3 - \hat{\theta}_2| \le \nu_j \end{split}$$
 (A. 1)

where $\hat{\theta} = [\hat{\theta}_1 \ \hat{\theta}_2]^{\mathrm{T}}$. Notice here that, from (12) and (13),

$$\nu_j \le \epsilon + \delta b_j$$
$$= 3$$

holds. Since $\hat{\theta}_1$ satisfying (A.1) does not exist if $\nu_j < 3$, we see $\nu_i = 3$. Then, solving (A.1), we have

$$\hat{\theta}_1 = 1, \quad 0 \le \hat{\theta}_2 \le 2.$$

This means that any $\hat{\theta}$ satisfying the inequality above can be an minimax estimate. Obviously, this is not unique.

Wataru KITAMURA (Member)



He received the B.E., M.E., and Ph.D. degrees from Kobe University, Japan in 2000, 2002, 2005, respectively. He is currently a research member of Logistics & Automation Division, Engineering Department, Murata Machinery, Ltd. He is a member of ISCIE. He was studying system identification and robust control theory at Kobe University.

Yasumasa Fujisaki (Member)



He received his B.S., M.S., and Ph.D. degrees from Kobe University, Japan, in 1986, 1988, and 1994, respectively. From 1988 to 1991, he was a Research Member of the Electronics Research Laboratory, Kobe Steel, Ltd., Japan. In 1991, he joined the faculty of Kobe University, where he is currently an Associate Professor of Department of Computer Science and Systems Engineering. His research interests include probabilistic methods for analysis and design of control systems. He is a member of ISCIE, ISIJ, and IEEE.

Reprinted from Trans. of the SICE Vol. 41 No. 2 111/117 2005