PARAMETER IDENTIFICATION FOR NONSTATIONARY NONLINEAR SYSTEMS

Sunahara, Yoshifumi
Kyoto Institute of Technology

Fukuda, Tokuo
Osaka University

PARAMETER IDENTIFICATION FOR NONSTATIONARY NONLINEAR SYSTEMS

By

Y. SUNAHARA* and T. FUKUDA**

Abstract

In this paper, a method is presented for modelling a class of nonlinear systems which involves time-varying parameters. The system model proposed here is a type of nonlinear difference equation, where unknown parameters are assumed to be linearly involved in the system model.

The principal line of attack is to assume that the nonlinear time-varying function in the system can be expanded into the M known functions with unknown constant coefficients.

First, the estimation process of unknown parameters is given by using the least squares method. Secondly, consistency properties and asymptotic normality conditions of the estimator are shown. Finally, two numerical examples are shown in order to demonstrate asymptotic properties of the estimator derived here.

1. Introduction

Up to the present time, the stochastic system modelling has been widely investigated by many time series analysists, system engineers and economists. Such phases are observed by literatures [1]~[4]. It may be fair to say that most of the previous works have been carried out within the framework of linear stationary models. However, in modelling real data, we have often encountered such data whose outstanding features are that it may have non-Gaussian distribution, nonstationarity or that it may have both non-Gaussianity and nonstationarity. It is obvious that such data mentioned above are not well modelled by a stationary linear model whose input disturbance is Gaussian.

Restricting to works of the system modelling by using either staitonary nonlinear models or nonstationary linear ones, the investigation has gradually become popular. For instance, in relation to the stationary nonlinear models, Netravari and De Figueirado [5] presented the consistent estimators of stationary nonlinear systems by using the stochastic approximation technique. Robinson [6] and Sunahara, et al. [7] presented the consistent estimators of interesting classes of nonlinear MA models by using moment methods. Furthermore, Poznyak [8] proved the consistency of estimators of

* Kyoto Institute of Technology, Matsugasaki, Kyoto 606, Japan.
** Osaka University, Suita, Osaka 565, Japan.

nonlinear regression models by using the least squares method. On the other hand, Lee [9], and Kozin and Nakajima [10] proposed the consistent estimators of linear time-varying models.

In this paper, introducing a class of nonstationary nonlinear models, a method is presented for identifying unknown parameters and asymptotic properties are investigated.

Throughout this paper, we shall use standard notation and terminology; the mathematical expectation is denoted by $E\{\cdot\}$, and the conditional one conditioned on "$*$" by $E\{\cdot\,|*\}$. The prime denotes the transpose of a vector or a matrix, and $c_i(i=1, 2, \cdots)$ denote constants.

2. Nonstationary Nonlinear Model

Let $\{y_k; k=0, 1, 2, \cdots\}$ be the observed n-dimensional discrete output data and assume that the precess $\{y_k\}$ actually comes from the following nonstationary nonlinear input/output relation,

$$(2.1) \qquad y_{k+1}-y_k=\phi_k(\mathring{\theta}, y_k)+B_{k+1}e_{k+1}, \qquad y_0: \text{initial constant,}$$

where B_{k+1} is the $n\times m$ known matrix and e_k is an m-vector unobservable input noise process which satisfies the following basic assumption;

[A.1] Let $\mathcal{F}_k(k=1, 2, \cdots)$ be the increasing σ-algebra generated by the sequence $\{e_i; i=1, 2, \cdots, k\}$. Then $\{e_k\}$ satisfies

$$(2.2) \qquad \begin{cases} E\{e_k|\mathcal{F}_{k-1}\}=0 & \text{w. p. 1} \\ E\{e_k e_k'|\mathcal{F}_{k-1}\}=\sigma^2 I_m & \text{w. p. 1,} \end{cases}$$

where $\sigma^2 I_m$ is the unknown variance of $\{e_k\}$. Furthermore, the time-varying nonlinear function $\phi_k(\theta, y_k)$ is, in this paper, assumed to be

$$(2.3) \qquad \phi_k(\mathring{\theta}, y_k)\triangleq\phi_{0,k}(y_k)+\sum_{j=1}^{M}\mathring{\theta}_j\phi_{j,k}(y_k),$$

where $\{\phi_{j,k}(y_k); j=1, 2, \cdots, M\}$ are known time-varying functions of y_k, and $\{\mathring{\theta}_j; j=1, 2, \cdots, M\}$ are unknown scalar constant parameters.

From (2.3), it is obvious that the system model (2.1) is a nonstationary nonlinear model. Thus, the problems are (i) to estimate the unknown parameters $\{\mathring{\theta}_j\}$ and (ii) to investigate asymptotic properties of the estimators.

3. Parameter Identification

Using the observation data $Y_N\triangleq\{y_1, y_2, \cdots, y_N\}$, the estimator $\hat{\theta}_N\triangleq[\hat{\theta}_{1,N}, \cdots, \hat{\theta}_{M,N}]'$, of $\mathring{\theta}\triangleq[\mathring{\theta}_1, \cdots, \mathring{\theta}_M]'$ is obtained by minimizing

$$(3.1) \qquad L_N(\theta)\triangleq\sum_{k=0}^{N}[y_{k+1}-y_k-\phi_k(\theta, y_k)]'[y_{k+1}-y_k-\phi_k(\theta, y_k)]$$

with respect to $\theta\triangleq[\theta_1, \cdots, \theta_M]'$. The minimum value of $L_N(\theta)$ is easily obtained by

setting $\partial L_N(\theta)/\partial\theta=0$ because of the quadratic form of (3.1) with respect to θ. Hence, defining s_N and Q_N as the M-vector and $M\times M$ matrix whose the i-th element and (i, j)-th component are respectively given by

$$(3.2) \qquad s_i(N)\triangleq\sum_{k=0}^{N}[y_{k+1}-y_k-\phi_{0,k}(y_k)]'\phi_{i,k}(y_k)$$

$$(3.3) \qquad q_{ij}(N)\triangleq\sum_{k=0}^{N}[\phi_{i,k}(y_k)]'\phi_{j,k}(y_k),$$

we have

$$(3.4) \qquad s_N=Q_N\hat{\theta}_N.$$

In order to avoid a numerical difficulty due to the singularity of Q_N, we introduce a matrix Γ_N defined by

$$(3.5) \qquad \Gamma_N\triangleq[Q_N+\rho I_M]^{-1}$$

where ρ is an arbitrary small positive constant given a priori. Therefore, by using this newly introduced matrix Γ_N instead of Q_N^{-1}, the unknown parameter vector θ is uniquely estimated by

$$(3.6) \qquad \hat{\theta}_N\triangleq\Gamma_N s_N.$$

Invoking the matrix inversion lemma [11], we have the recursive version of (3.6) by using (3.3) to (3.6) as follows:

$$(3.7) \quad \hat{\theta}_N=\hat{\theta}_{N-1}+\Gamma_N F_N[I+F_N'\Gamma_{N-1}F_N]^{-1}[y_{N+1}-y_N-\phi_{0,N}(y_N)-F_N'\hat{\theta}_{N-1}], \qquad \hat{\theta}_0=\Gamma_0 s_0,$$

$$(3.8) \quad \Gamma_N=\Gamma_{N-1}-\Gamma_{N-1}F_N[I+F_N'\Gamma_{N-1}F_N]^{-1}F_N'\Gamma_{N-1}, \qquad \Gamma_0=[Q_0+\rho I_M]^{-1},$$

where F_N is defined by

$$(3.9) \qquad F_N\triangleq[\phi_{1,N}(y_N), \cdots, \phi_{M,N}(y_N)]'.$$

4. Consistent Property of Estimators

First, define the estimation error $\tilde{\theta}_N$ by

$$(4.1) \qquad \tilde{\theta}_N\triangleq\mathring{\theta}-\hat{\theta}_N.$$

Then, from (3.6), it follows that

$$(4.2) \qquad \tilde{\theta}_N\triangleq\mathring{\theta}-\Gamma_N s_N.$$

Furthermore, from (2.1), (2.3), (3.3) and (3.5), we have

$$(4.3) \qquad s_N=(\Gamma_N^{-1}-\rho I)\mathring{\theta}+\sum_{k=0}^{N}F_k B_{k+1}e_{k+1}.$$

Hence, substituting (4.3) into (4.2), the estimation error $\tilde{\theta}_N$ can be decomposed as follows:

$$(4.4) \qquad \tilde{\theta}_N=\rho\Gamma_N\mathring{\theta}-\Gamma_N\sum_{k=1}^{N}F_k B_{k+1}e_{k+1}.$$

In the following, we concentrate our attention to prove that $\tilde{\theta}_N$ converges to zero w. p. 1 in two cases of the single parameter and the multi parameter.

4.1. Single-parameter Case

Let us restrict first our attention on the simplest case, i. e., the case where $\mathring{\theta}$ is scalar. In this case, $\mathring{\phi}_k(\mathring{\theta}, y_k)$ and Γ_k defined in Chapters 2 and 3 yield respectively

$$(4.5) \qquad \mathring{\phi}_k(\mathring{\theta}, y_k) = \phi_{0,k}(y_k) + \mathring{\theta}\phi_{1,k}(y_k) \qquad (n\text{-vector})$$

$$(4.6) \qquad \Gamma_N = \left\{ \sum_{k=0}^{N} [\phi_{1,k}(y_k)]'\phi_{1,k}(y_k) + \rho \right\}^{-1} \qquad (\text{scalar}).$$

Then, we have the following theorem.

THEOREM 1. *With* [A. 1], *assume that*
(C.1) $\|B_k\|$ *is uniformly bounded,*
(C.2) *for any* $\kappa \in R^n$, *there exists* $\varepsilon > 0$ *which satisfies*

$$\limsup_{k\to\infty} P\{\|\phi_{1,k+1}(\kappa + B_{k+1}e_{k+1})\| > \varepsilon\} > 0.$$

Then,

$$\hat{\theta}_N \to \mathring{\theta} \qquad as \ N\to\infty \qquad \text{w. p. 1.}$$

The proof of this theorem is given in Appendix 1.

4.2. Multi-parameter Case

First, define

$$(4.7) \qquad \tilde{\phi}_k(\theta, z) \triangleq z + \phi_k(\theta, z).$$

Then, the following conditions are sufficient to prove the consistency of the estimator $\hat{\theta}_N$ in the multi-parameter case;
(C.3) $\tilde{\phi}_k(\theta, z)$ satisfies the uniform Lipschitz condition and is bounded at the initial time $k=0$, i. e., for all $z_1, z_2 \in R^n$ and for all $\theta \in R^M$

$$\|\tilde{\phi}_k(\theta, z_1) - \tilde{\phi}_k(\theta, z_2)\| \leqq c_1\|z_1 - z_2\|$$

$$\|\tilde{\phi}_k(\theta, 0)\| \leqq c_2$$

and furthermore, we assume $c_1 < 1$, which implies that the system (2.1) is stable;
(C.4) the functions $\{\phi_{i,k}(\theta, \cdot); i=1, \cdots, M\}$ defined in (2.3) satisfy the uniform Lipschitz condition and are uniformly bounded at $k=0$, i. e.,

$$\|\phi_{i,k}(\theta, z_1) - \phi_{i,k}(\theta, z_2)\| \leqq c_3\|z_1 - z_2\|$$

$$\|\phi_{i,k}(\theta, 0)\| \leqq c_4;$$

(C.5) for all $\lambda \in R^n$,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N} E\{F_k(\lambda + B_{k+1}e_{k+1})[F_k(\lambda + B_{k+1}e_{k+1})]' \mid \mathcal{F}_k\} \geqq c_5 I > 0, \qquad \text{w. p. 1.}$$

(C.6)　$E\{(e_k e_k')^2 | \mathcal{F}_{k-1}\} \leqq c_6 I.$

THEOREM 2. *Let* [A.1] *hold. Furthermore, assume that the conditions* (C.3) *to* (C.6) *hold. Then*

$$\hat{\theta}_N \to \hat{\theta} \quad as \ N \to \infty \quad w. p. 1.$$

The proof of Theorem 2 is given in Appendix 2.

5. Asymptotic Normality of Estimators

In this chapter, in order to evaluate the asymptotic accuracy of the estimators, we show the asymptotic normality of the estimator $\hat{\theta}_N$ given in Chapter 3. For this purpose, we first define

(5.1)　　　　　　　　$P_N \triangleq \Gamma_N^{-1} / N$

(5.2)　　　　　　　　$U_N \triangleq \dfrac{1}{N} \sum_{k=0}^{N} E\{F_k B_{k+1} B_{k+1}' F_k'\}$

(5.3)　　　　　　　　$\tilde{\theta}_N \triangleq \hat{\theta} - \hat{\theta}_N - \rho \Gamma_N \theta.$

Then, the following theorem is obtained.

THEOREM 3. *Assume that* [A.1] *and* (C.3) *to* (C.6) *hold. Furthermore assume that the following condition holds;*

(C.8)　　　　　　　　$B_k' B_k > 0 \ for \ all \ k = 1, 2, \cdots.$

Then

(5.9)　　　　　　$\sqrt{N} U_N^{-1/2} P_N \tilde{\theta}_N \xrightarrow{law} N(0, I) \quad as \ N \to \infty,$

where "law" denotes the convergence in law and $N(0, I)$ *denotes the Gaussian probability distribution with zero mean and unit covariance.*

The proof of this theorem is also given in Appendix 3.

6. Digital Simulation Studies

The purpose of this chapter is to examine numerically asymptotic properties of the estimators.

For both cases of the single- and the multi-parameter, the consistency and asymptotic normality of the estimator $\hat{\theta}_N$ are examined. First, a simple single-parameter case is shown.

EXAMPLE 1. Consider the following mathematical model of a damped oscilation system,

(6.1)　　　　$\ddot{x}(t) + a \ \mathrm{sgn}\,(\dot{x}(t)) \dot{x}(t)^2 + b x(t) = n(t), \quad x(0) = x_0 \ and \ \dot{x}(0) = \dot{x}_0$

where $n(t)$ is a white Gaussian noise, and the coefficient a is unknown while b is known. The system model (6.1) can be represented by the following state space form by setting $x_1 \triangleq x$ and $x_2 \triangleq \dot{x}$:

$$(6.2) \qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \text{ sgn}(x_2)x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} n(t).$$

Let $\{t_k; k=0, 1, \cdots\}$ be the partitioned time as $\delta = t_{k+1} - t_k$. Then, by setting $y_k \triangleq [y_{1,k}, y_{2,k}]' = [x_1(t_k), x_2(t_k)]'$, we have the following approximated discrete model for (6.2);

$$(6.3) \qquad \begin{cases} y_{k+1} - y_k = \phi_k(\mathring{\theta}, y_k) + B e_{k+1} \\ \phi_k(\mathring{\theta}, y_k) \triangleq \phi_{0,k}(y_k) + \mathring{\theta}\phi_{1,k}(y_k) \end{cases}$$

where $\mathring{\theta} \triangleq -a$, $e_k \triangleq \sqrt{\delta} n(t_k)$, $B \triangleq [0, 1]'$ and $\phi_{0,k}(y_k)$ and $\phi_{1,k}(y_k)$ are respectively given by

$$(6.4) \qquad \phi_{0,k}(y_k) = \begin{bmatrix} -\delta y_{1,k} \\ b y_{2,k} \end{bmatrix}, \qquad \phi_{1,k}(y_k) = \begin{bmatrix} 0 \\ -\delta \text{ sgn}(y_{2,k}) y_{2,k} \end{bmatrix}.$$

It is easily verified that the system model (6.3) satisfies all conditions in Theorem 1. In digital simulation experiments, the time interval δ was set as $\delta = 0.01$ and the true
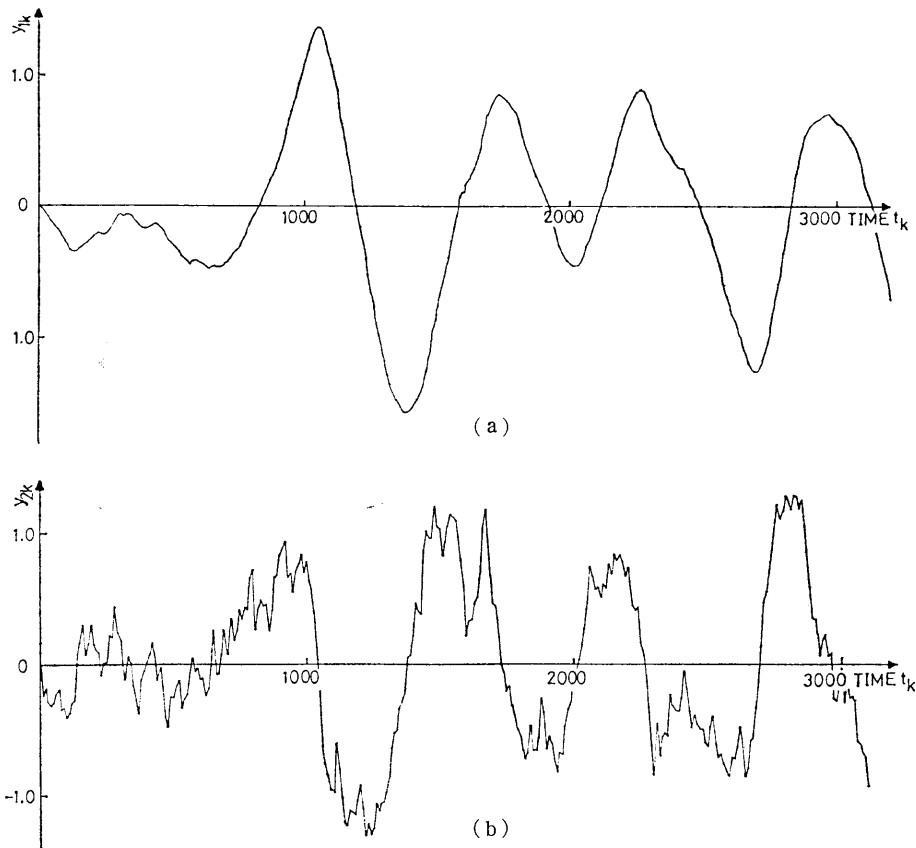

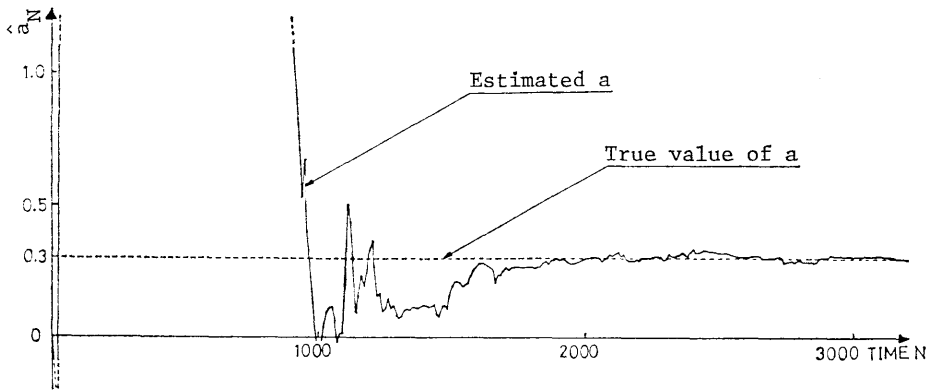Fig. 1. Sample runs of $y_{1,k}$ and $y_{2,k}$ processes of Example 1.

Fig. 2.   A sample run of the estimated $a(=\overset{\circ}{\theta})$.
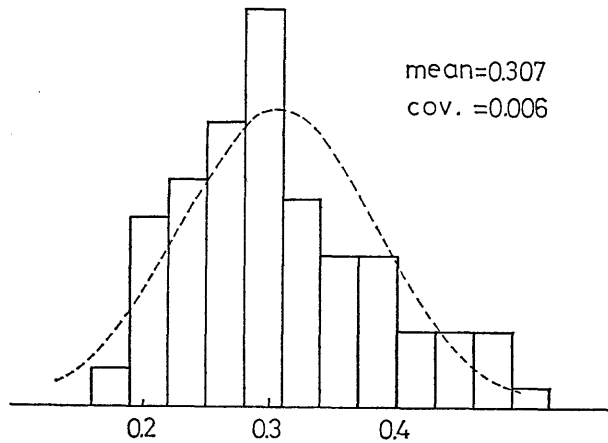

Fig. 3.   The histogram for θ_N at $N=3000$ where the dotted line
denotes the fitted normal curve.

value of $\overset{\circ}{\theta}=-a$ was set -0.3. The values of known constant b and the covariance of e_k were set respectively 0.1 and 0.2. With these values of parameters, the output sequence of $y_{1,k}$ was obtained by simulating (6.2) on the digital computer, and $y_{2,k}$ is also obtained by numerically differentiating $y_{1,k}$. Sample paths of $y_{1,k}$- and $y_{2,k}$-processes are depicted in Fig. 1. Figure 2 shows the convergence property of the estimator $\hat{a}_N(=-\hat{\theta}_N)$. The histogram of 100 sample runs of $\hat{\theta}_N$ at $N=3000$ is depicted in Fig. 3 in order to show the asymptotic normality of $\hat{\theta}_N$.

EXAMPLE 2.   A slightly complexed system is examined which is known as a model of the rolling motion of a ship [12].

(6.5) $\ddot{x}(t)+\alpha\dot{x}(t)+\beta g(x(t))=n(t)$, $x(0)=x_0$ and $\dot{x}(0)=x_0$

where α and β are unknown constant parameters. As a model of the ship rolling motion, the function $g(x)$ is given by $g(x)=1+\gamma x^2$. However, in order to satisfy the conditions (C.3) and (C.4), we assume
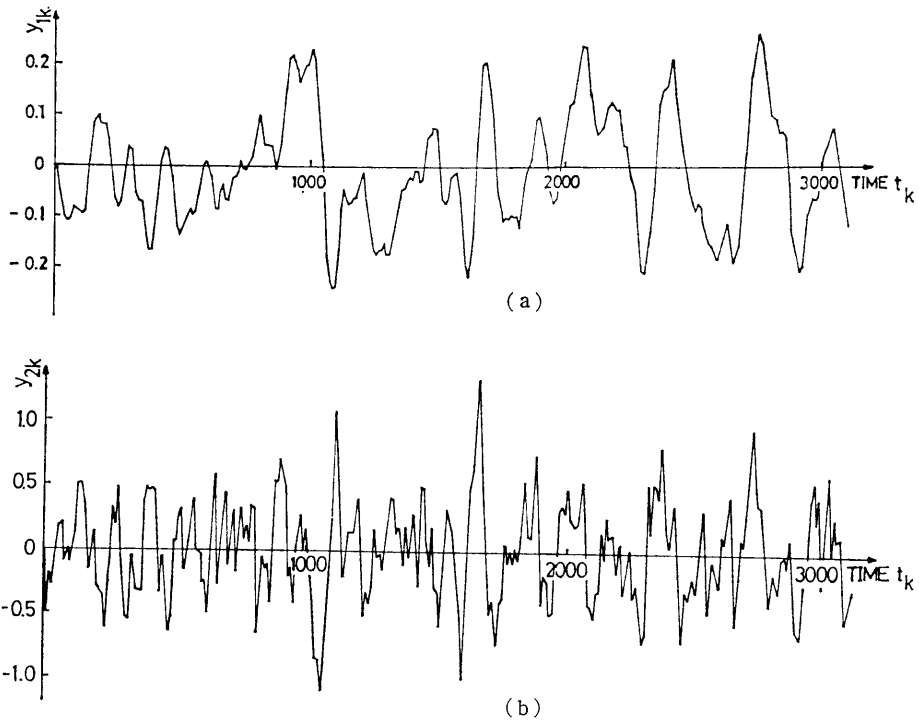
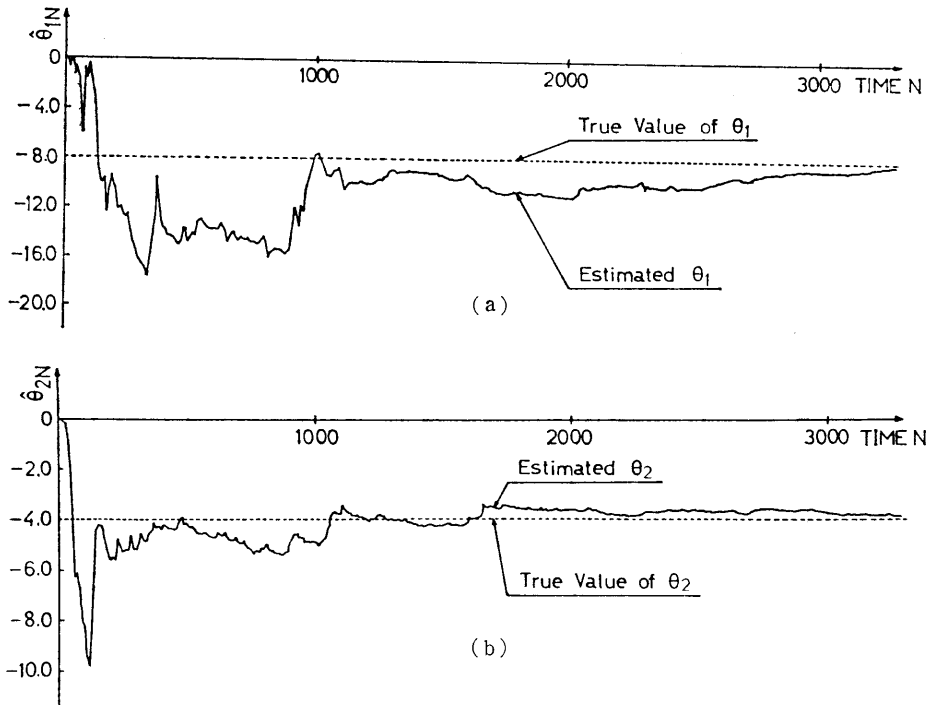Fig. 4.   Sample runs of $y_{1,k}$ and $y_{2,k}$ processes of Example 2.

Fig. 5.   Sample runs of the estimated $\overset{\circ}{\theta}_1$ and $\overset{\circ}{\theta}_2$.
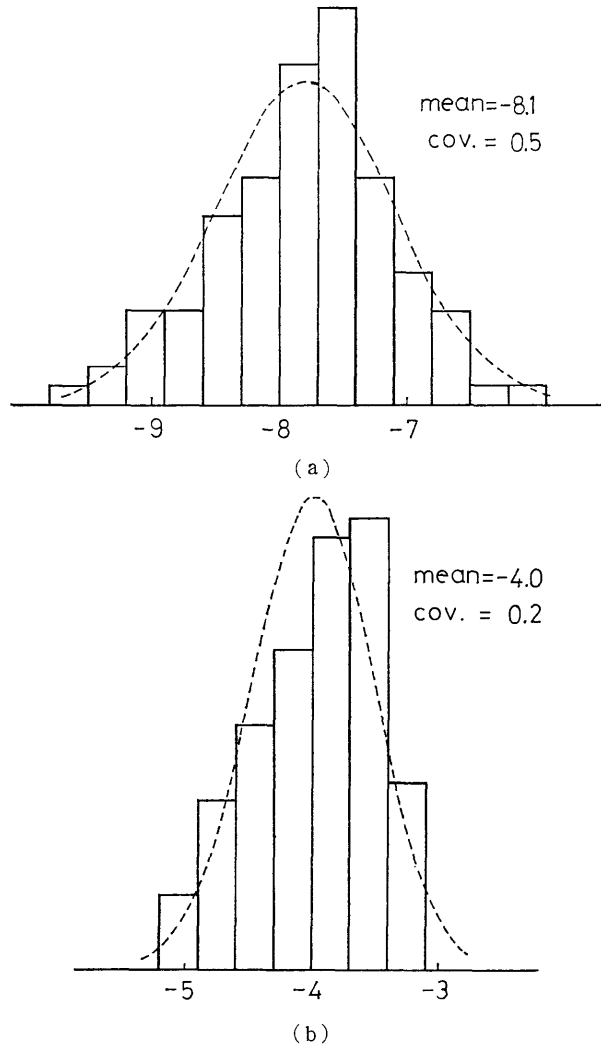
Fig. 6. The histograms for $\theta_{1,N}$ and $\theta_{2,N}$ at $N=3000$ where the dotted lines denote the fitted normal curves.

$$(6.6) \qquad g(x)=\begin{cases}1+\gamma\mu & \text{for } \|x\|^2\geqq\mu \\ 1+\gamma x^2 & \text{for } \|x\|^2\leqq\mu\end{cases}$$

where γ and μ are given constants. By the same procedure as in Example 1, we have the following discrete state space model for (6.5),

$$(6.7) \qquad \begin{cases}y_{k+1}-y_k=\phi_k(\mathring{\theta},\ y_k)+Be_{k+1} \\ \phi_k(\mathring{\theta},\ y_k)\triangleq\phi_{0,k}(y_k)+\mathring{\theta}_1\phi_{1,k}(y_k)+\mathring{\theta}_2\phi_{2,k}(y_k)\end{cases}$$

where $\mathring{\theta}\triangleq[-\beta,\ -\alpha]'$, $B\triangleq[0,\ 1]$. Furthermore, $\phi_{0,k}(y_k)$ to $\phi_{2,k}(y_k)$ are respectivery given by

$$\dot{\varphi}_{0.\,k}(y_k) = \begin{bmatrix} \bar{\delta} y_{2.\,k} \\ 0 \end{bmatrix}, \qquad \dot{\varphi}_{1.\,k}(y_k) = \begin{bmatrix} 0 & -\varepsilon\bar{\delta} \\ \bar{\delta} g(y_{1.\,k}) & 0 \end{bmatrix} y_k,$$

$$\dot{\varphi}_{2.\,k}(y_k) = \begin{bmatrix} 0 \\ \bar{\delta} y_{2.\,k} \end{bmatrix},$$

where we added $-\varepsilon\bar{\delta}(\varepsilon > 0)$ to the (1, 2)-th component of $\dot{\varphi}_{1.\,k}(y_k)$ in order to satisfy the condition (C.5). The selected true values of $\mathring{\theta}_1$, $\mathring{\theta}_2$ and σ^2 were set respectively as -8.0, -4.0 and 1.0. The sample paths of $y_{1.\,k}$ and $y_{2.\,k}$ are depicted in Fig. 4. The convergence features of $\hat{\theta}_{1.\,N}$ and $\hat{\theta}_{2.\,N}$ are shown in Fig. 5 and the histograms of 100 sample runs of $\hat{\theta}_{1.\,N}$ and $\hat{\theta}_{2.\,N}$ at $N = 3000$ are shown in Fig. 6.

7. Conclusions

A method has been presented for modelling a class of nonstationary nonlinear systems which could be expected to perform a good fit to real data by using a small number of parameters.

The key assumption in this paper is that the nonlinear time-varying function can be well approximated by the series of M known functions with unknown constant coefficients.

The estimator of unknown parameters is obtained by using the least squares method. Both consistency and asymptotic normality of the estimator have been proved by using martingale properties. It should be emphasized that in the case of one unknown parameter, the consistency of the estimator holds without stable conditions, while in the case of many unknown parameters, the stability conditions due to the concept of the bounded inputs bounded outputs are required.

The present approach may be expanded to the case where the unknown input covariance is required to be estimated. That is, let $\hat{\sigma}_N{}^2$ be the estimate of input covariance σ^2. The estimator is given by

$$\hat{\sigma}_N \triangleq \frac{1}{mN} \sum_{k=0}^{N} [y_{k+1} - y_k - \phi_k(\hat{\theta}_k, y_k)]' \tilde{B}_{k+1} [y_k - y_k - \phi_k(\hat{\theta}_k, y_k)]$$

if B_{k+1}^{-1} exists, where

$$\tilde{B}_{k+1} \triangleq [B_{k+1} B_{k+1}']^{-1}.$$

APPENDIX 1. *Proof of Theorem 1*

In order to prove Theorem 1, we need the following lemma.

LEMMA 1. *Let [A.1] and (C.2) hold. Then the following scalar quantity h_N defined by*

(A.1) $\qquad h_N = \sum_{k=0}^{N} \Gamma_k F_k B_{k+1} e_{k+1} \qquad (N=0, 1, \cdots),\ (h_N=0 \ for \ N<0)$

converges w.p.1 to some random variable h with the property $E\{h^2\} < \infty$.

PROOF. Since both F_k and Γ_k are $\mathcal{F}_k$-measurable, it is easily varified that $\{h_N, \mathcal{F}_{N+1}\}$ is a martingale. From [A.1], (C.1) and (4.6), we have

$$E\{(\Gamma_0 F_0 B_1 e_1)^2\} \leqq c_7 < +\infty$$

$$E\{(\Gamma_j F_j B_{j+1} e_{j+1})^2 \,|\, \mathcal{F}_j\} \leqq c_8{}^2 \sigma^2 (\Gamma_{j-1} - \Gamma_j)$$

where $\Gamma_j{}^2(\Gamma_j{}^{-1} - \Gamma_{j-1}{}^{-1}) \leqq \Gamma_{j-1} - \Gamma_j$ has been used. Hence

$$E\{h_N{}^2\} = E\{E\{(\Gamma_j F_j B_{j+1} e_{j+1})^2 \,|\, \mathcal{F}_j\}\} + E\{(\Gamma_0 F_0 B_1 e_1)^2\}$$

$$\leqq c_8{}^2 \sigma^2 (\Gamma_0 - E\{\Gamma_N\}) + c_7 \leqq c_8{}^2 \sigma^2 / (Q_0 + \rho) + c_8$$

which means that $\{h_N, \mathcal{F}_{N+1}\}$ is the martingale bounded in L^2. Therefore from the L^2 martingale convergence theorem (see Loeve [13] p. 396), we have $h_N \to h$ w.p.1 as $N \to \infty$.

PROOF OF THEOREM 1: By using Lemma 1, from (4.4) it follows that

(A.2) $\qquad \tilde{\theta}_N = \rho \Gamma_N \mathring{\theta} - \sum_{k=1}^{N} \Gamma_N F_k F_k{}'(h - h_{k-1}) - (h_N - h) - \Gamma_N .$

From (3.5), we have

(A.3) $\qquad \sum_{k=1}^{N} |\Gamma_N F_k F_k{}'| = \sum_{k=1}^{N} \dfrac{F_k F_k{}'}{\sum_{l=0}^{N} F_l F_l{}' + \rho} < 1$

and $\Gamma_N F_k F_k{}' \to 0$ w.p. 1 as $N \to \infty$, provided that $\Gamma_N \to 0$ as $N \to \infty$. Then, by using Toeplitz lemma (see Loeve [13] p. 238), it is easily proved that the 2nd term converges to zero w.p. 1 as $N \to \infty$. Furthermore, it is obvious that other terms of (A.2) converges to zero w.p. 1, provided that $\Gamma_N \to 0$ as $N \to \infty$. Hence the remaining problem is to prove that $\Gamma_N \to 0$ as $N \to \infty$. Recalling that $\Gamma_N{}^{-1} = \sum_{k=1}^{N} [\phi_{1,k}(y_k)]' \phi_{1,k}(y_k) + \rho$, for any $l=1, 2, \cdots$, it follows that

(A.4) $\qquad P\{\sup_l |\Gamma_{N+l}{}^{-1} - \Gamma_N{}^{-1}| > \delta\} > P\{\|\phi_{1,N+1}(y_{N+1})\|^2 > \delta\} .$

Then, by invoking the elementary inequality, $\inf_{n>N}(a_n - b_n) \leqq \sup_{n>N} a_n - \sup_{n>N} b_n$ for any a_n and b_n, from (A.4) we have

(A.5) $\qquad \liminf_{N \to \infty} P\{\sup_l |\Gamma_{N+l}{}^{-1} - \Gamma_N{}^{-1}| \leqq \delta\}$

$$\leqq 1 - \limsup_{N \to \infty} P\{\|\phi_{1,k+1}(\kappa + B_{k+1} e_{k+1})\|^2 > \delta\} < 1$$

where $\kappa \triangleq y_n + \phi_k(\mathring{\theta}, y_k)$ and the condition (C.2) has been used. Hence from the Cauchy criterion for the convergence w.p. 1 (see Lukacs [14] p. 45), $\Gamma_N{}^{-1}$ diverges to infinity as $N \to \infty$ w.p. 1, which means $\Gamma_N \to 0$ w.p. 1 as $N \to \infty$.

APPENDIX 2. *Proof of Theorem 2*

Define the scalar random variables γ_N and β_N by

(A.6) $\gamma_N \triangleq \sum_{k=0}^{N} \alpha' F_k B_{k+1} e_{k+1}$

(A.7) $\beta_N \triangleq \sum_{k=0}^{N} \alpha'(F_k F_k' - E\{F_k F_k' | \mathcal{F}_{k-1}\})\alpha$

respectively, where α is an arbitrary vector such that $\alpha \neq 0$. Then, it is easily verified that $\{\gamma_N, \mathcal{F}_N\}$ and $\{\beta_N, \mathcal{F}_{N-1}\}$ are martingales. Since it is easily verified that, with the assumption [A.1], (C.1), (C.3), (C.4) and (C.7)

$$E\{\|y_k\|^2\} \leqq c_9, \qquad E\{\|y_k\|^4\} \leqq c_{10},$$

then from (C.4) we have following evaluation:

(A.8) $E\{\gamma_N^2\} \leqq \sigma^2 c_{11} \|\alpha\|^2 (N+1)$

(A.9) $E\{\beta_N^2\} \leqq c_{12} \|\alpha\|^2 (N+1).$

Recall here the following lemma by Khasminskii [15].

LEMMA 2. *Suppose that* $\{u_k, \mathcal{F}_k\}$ *is a martingale which satisfies* $E\{u_k^2\} \leqq c_{13} k$, *then* $u_k/k \to 0$ *w.p.* 1 *as* $k \to \infty$.

By regarding γ_N or β_N and N as u_k and k respectively, we may have

(A.10) $\frac{1}{N}\gamma_N = \frac{1}{N}\sum_{k=0}^{N} \alpha' F_k B_{k+1} e_{k+1} \longrightarrow 0 \qquad$ as $N \to \infty$ w.p. 1

(A.11) $\frac{1}{N}\beta_N = \frac{1}{N}\left[\sum_{k=0}^{N}\{\alpha' F_k F_k'\alpha - \alpha' E\{F_k F_k' | E_{k-1}\}\alpha\}\right] \longrightarrow 0 \qquad$ as $N \to \infty$ w.p. 1

From (A.11), it is easily verified that, with (C.5) and with the basic inequality, $\inf_{n>N}(a_n - b_n) \leqq \inf_{n>N} a_n - \inf_{n>N} b_n$,

(A.12) $N\dot{\Gamma}_N = \left[\frac{1}{N}\left(\sum_{k=0}^{N} F_k F_k' + \rho I_M\right)\right]^{-1} \leqq c_{14} I_M \qquad$ w.p. 1

for a sufficiently large N. Hence, from (A.10) and (A.12), we can conclude that the 1st and 2nd terms in the R.H.S. of (4.4) converge to zero w.p. 1 as $N \to \infty$.

APPENDIX 3. *Proof of Theorem 3*

Define

(A.13)
$$\begin{cases} x_{N,i} \triangleq -\frac{1}{\sqrt{N}}\alpha' U_N^{-1/2} F_i B_{i+1} e_{i+i} \\[2mm] S_{N,k} \triangleq \sum_{i=0}^{k} x_{N,i} \\[2mm] \tilde{\sigma}_{N,k}^2 \triangleq E\{x_{N,k}^2 | \mathcal{F}_{N,k-1}\} \\[2mm] v_{N,k}^2 \triangleq \sum_{i=0}^{k} \tilde{\sigma}_{N,i}^2 \end{cases}$$

where $\mathcal{F}_{N,k-1}$ is the σ-algebra generated by $\{x_{N,1}, \cdots, x_{N,k-1}\}$. We introduce the following lemma by Brown and Eagleson [16].

LEMMA 3. *If* $\{S_{N,k}, \mathcal{F}_{N,k}; k=1, 2, \cdots, N\}$ *forms a matingale for each* N *and*

(i) $v_{N,N}^2 \to \tilde{\sigma}^2$ (const.) as $N \to \infty$ in prob.

(ii) for each $\varepsilon > 0$

$$\sum_{k=0}^{N} E\{x_{N,k}^2 I(|x_{N,k}| > \varepsilon)|\mathscr{F}_{N,k-1}\} \longrightarrow 0 \qquad as \ N \to 0 \ in \ prob.$$

Then

$$S_{N,N} \xrightarrow{\ law\ } N(0, \sigma^2) \qquad as \ N \to \infty.$$

Hence in order to prove Theorem 3, it is sufficient to show that all conditions in Lemma 3 are satisfied. Since, from (A.13),

$$\tilde{\sigma}_{N,k}^2 = (\sigma^2 \alpha' U_N^{-1/2} F_k B_{k+1} B_{k+1}' F_k' U_N^{-1/2} \alpha)/N,$$

we have

(A.14) $$v_{N,N}^2 = \|\alpha\|^2,$$

which implies that $v_{N,N}$ satisfies the condition (i) of Lemma 3. Since it is easily verified with the conditions of Theorem 3 that the expectation of $x_{N,k}^2$ is bounded, then it follows that

(A.15) $$E\{x_{N,k}^2 I(|x_{N,k}| > \varepsilon)\} \leqq E\{x_{N,k}^2 I(|\sqrt{N} x_{N,k}| > \sqrt{k} \varepsilon)\} \longrightarrow 0 \qquad as \ k \to \infty.$$

Hence by using (A.15) and the Markov inequality, we have

$$P\{|\sum_{k=0}^{N} E\{x_{N,k}^2 I(|x_{N,k}| > \varepsilon)|\mathscr{F}_{N,k-1}\}| > \delta\}$$

$$\leqq \frac{1}{\delta} \sum_{k=1}^{N} E\{x_{N,k}^2 I(|x_{N,k}| > \varepsilon)\} \longrightarrow 0 \qquad as \ N \to \infty$$

which implies the condition (ii) of Lemma 3.

References

[1] G. Box and G. Jenkins, *Time Series Analysis; Forecasting and Control*, Holden-Day, New York, (1970).

[2] K. J. Åstrom and P. Eykhoff, *System Identification-A Survey*, Automatica, Vol. 7, (1970), pp. 1-38.

[3] P. L. Kashyap and A. R. Rao, *Dynamic Stochastic Models from Empirical Data*, Academic Press, New York, (1976).

[4] R. K. Mehra and D. G. Lainiotis, *System Identification: Advances and Case Studies*, Academic Press, New York, (1976).

[5] A. Netravari and R. J. P. De Figueirado, *On the Identification of Nonlinear Dynamical Systems*, IEEE Trans. Automat. Contr., Vol. AC-16, (1971), pp. 28-36.

[6] P. M. Robinson, *The Estimation of a Nonlinear Moving Average Model*, Stochastic Processes and Their Applications, Vol. 5, (1977), pp. 81-90.

[7] Y. Sunahara, A. Ohsumi, T. Fukuda and T. Kinoshita, *On Modelling of Random Data via Nonlinear MA-Models*, Proc. 23rd Joint Automatic Control Conference (JACC), Kyoto, 1980.

[8] A. S. Poznyak, *Estimating the Parameters of Autoregression Processes by Method of Least Squares*, Int. J. Systems Sci., Vol. 11, (1980), pp. 577-588.

[9] T. Lee, *Consistent Estimators for a Class of Nonstationary Models*, Ph. D. Thesis, Polytechnic Institute of New York (1976).

[10] F. KOZIN and F. NAKAJIMA, *The Order Determination Problem for Time-Varying AR Models*, IEEE Trans. Automat. Contr., Vol. AC-25, (1980), pp. 250-257.

[11] A. P. SAGE and J. L. MELSA, *Estimation Theory with Applications to Communications and Control*, McGraw-hill, New York, (1971).

[12] T. OZAKI and H. ODA, *Nonlinear Time Series Model Identification by Akaike's Information Criterion*, Proc. of IFAC Workshop on "Information and Systems", Compiegue, Oct. 1977, (ed. B. Dubuisson, Pergamon Press).

[13] M. LOEVE, *Probability Theory*, 3rd. ed., D. Van Nostrand. Princeton. N. J. (1963).

[14] E. LUKACS, *Stochastic Convergence*, Academic Press, New York, (1975).

[15] R. Z. KHASMINSKII, *Stability of Systems of Differential Equation under Random Disturbances of Their Parameters*, Nauka, Moscow (1969).

[16] B. M. BROWN and G. K. EAGLESON, *Martingale Convergence to Infinity Divisible Laws with Finite Variance*, Trans. Am. Soc., Vol. 162, (1971).