ICROS-SICE International Joint Conference 2009
August 18-21, 2009, Fukuoka International Congress Center, Japan

Asymptotic System Identification Method Based on Particle Swarm Optimization

Hideo Muroi[1] and Shuichi Adachi[1]

[1]Department of Applied Physics and Physico-Informatics, Keio University, Yokohama, Japan
(E-mail:adachi@appi.keio.ac.jp)

Abstract: In general, structure of a system to be identified is unknown for users a *priori*. This makes the model complex and high order structure. In this paper, we introduce the asymptotic method (ASYM) to deal with the problem. ASYM calculates a high-order model using the well-known least squares method, then reduces the identified model to a simple one. For this model reduction, various model reduction techniques such as balanced realization and output error reduction were developed. In this paper, a new method to reduce the high-order model using the particle swarm optimization in the frequency domain is proposed. Effectiveness of the proposed method is examined through numerical examples.

Keywords: System identification, asymptotic method, particle swarm optimization, high-order estimation, model reduction, curve fitting.

1. INTRODUCTION

System identification is sophisticated methods to obtain mathematical models based on input and output data of the system. However, models obtained by the conventional system identification methods often have complex structure with high order. In general it is difficult to use them for control systems design.

In this paper, we make use of the asymptotic method (ASYM) to deal with the problem. The ASYM consists of two steps. In the first step, a high-order model is identified by using the least squares method, and in the second step, the high-order model is reduced by minimizing the negative log-likelihood function based on the asymptotic theory [1]. For this minimization, Wahlberg proposed the frequency weighted balanced realization [2], and Zhu used the output error method [3]. In this paper, we propose a new ASYM model reduction idea of curve fitting in the frequency domain by using modal analysis, broken line approximation and Particle Swarm Optimization (PSO) [4].

2. ASYMPTOTIC METHOD

Consider a single input, single output (SISO) linear discrete-time system described by

$$y(k) = G^o(q)u(k) + H^o(q)e(k) \tag{1}$$

where q denotes the time-shift operator. $G^o(q)$ and $H^o(q)$ are the system to be identified and the noise transfer functions, respectively. $u(k)$ and $y(k)$ are the input and output, respectively, and $e(k)$ is white noise with zero mean and variance R. The purpose of this paper is to estimate $G^o(q)$ in Eq.(1) accurately based on input-output data.

2.1 The asymptotic theory

Ljung proposed the asymptotic theory [1]. Suppose that both the model order n and the number of the data samples for the estimation N tend to infinity, the main results of this theory are summarized as follows.

$$\lim_{N\to\infty} \begin{bmatrix} \hat{G}_N^n(e^{j\omega}) \\ \hat{H}_N^n(e^{j\omega}) \end{bmatrix} = \begin{bmatrix} G^o(e^{j\omega}) \\ H^o(e^{j\omega}) \end{bmatrix}, \text{ w.p.1} \tag{2}$$

$$\lim_{N\to\infty} \sqrt{N} \begin{bmatrix} \hat{G}_N^n(e^{j\omega}) - \mathrm{E}[\hat{G}_N^n(e^{j\omega})] \\ \hat{H}_N^n(e^{j\omega}) - \mathrm{E}[\hat{H}_N^n(e^{j\omega})] \end{bmatrix} = \mathcal{N}(\mathbf{0}, \boldsymbol{P}_n(\omega)) \tag{3}$$

$$\lim_{n\to\infty} \frac{1}{n}\boldsymbol{P}_n(\omega) = \Phi_v(\omega)\Phi^{-1}(\omega) \tag{4}$$

where

$$\Phi(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{eu}(\omega) \\ \Phi_{ue}(\omega) & R \end{bmatrix}$$

$$\Phi_v(\omega) = |H^o(e^{j\omega})|^2 R,$$

Φ_u and Φ_v denote the power spectrum of u and v, respectively, Φ_{ue} denotes the cross spectrum between u and e. $\mathcal{N}(\mathbf{0}, \boldsymbol{P}_n(\omega))$ means Gaussian distribution with zero mean and variance $\boldsymbol{P}_n(\omega)$.

Eq.(2) means that the model estimates are consistent asymptotically. Eq.(3) means that the error of the transfer functions at each frequency converges to a Gaussian distribution. These results are proved in [1].

2.2 High-order estimation

Dividing Eq.(1) by $H^o(q)$,

$$\frac{1}{H^o(q)}y(k) = \frac{G^o(q)}{H^o(q)}u(k) + e(k). \tag{5}$$

is obtained. This form is considered to be a high-order ARX model, because it can be expressed by

$$A^o(q)y(k) = B^o(q)u(k) + e(k) \tag{6}$$

where $A^o(q)$ and $B^o(q)$ are polynomials in q, respectively. So, the coefficients of these polynomials can be easily estimated by the ordinary least-squares method.

PR0002/09/0000-4499 ¥400 © 2009 SICE

Then, the estimated system model $\hat{G}^n(q)$ and estimated noise model $\hat{H}^n(q)$ are derived as

$$\hat{G}^n(q) = \frac{\hat{B}(q)}{\hat{A}(q)}, \quad \hat{H}^n(q) = \frac{1}{\hat{A}(q)} \tag{7}$$

respectively, where hat denotes the estimate.

2.3 Asymptotic variance of the system model

From Eq.(4), an expression for the asymptotic variance of the system model is approximated by

$$\mathrm{var}[\hat{G}^n_N(e^{j\omega})] \approx \frac{n}{N}\frac{\Phi_v(\omega)R}{\Phi_u(\omega)R - |\Phi_{ue}(\omega)|^2}. \tag{8}$$

In particular, $\Phi_{ue} = 0$ for the open-loop case. Thus Eq.(8) is simplified to

$$\mathrm{var}[\hat{G}^n_N(e^{j\omega})] \approx \frac{n}{N}\frac{\Phi_v(\omega)}{\Phi_u(\omega)}. \tag{9}$$

Eq.(9) means that the asymptotic variance of the system model at some frequency ω is proportional to the model order n and the noise power spectrum Φ_v, and inverse proportional to the number of the samples and signal power spectrum Φ_u. Thus, the intensity of input signal should be strengthened for the frequency band of interest.

Then, negative log-likelihood function [2] in the frequency domain is defined as

$$V = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\hat{G}^n(e^{j\omega}) - \hat{G}^l(e^{j\omega})\right|^2 \frac{\Phi_u(\omega)}{\Phi_v(\omega)} \mathrm{d}\omega \tag{10}$$

where $\hat{G}^l$ is a reduced system model. Note that Φ_v can be replaced by estimated noise transfer function $\hat{H}^n(e^{j\omega})$. By minimizing Eq.(10) with respect to $\hat{G}^l(e^{j\omega})$, the asymptotic maximum likelihood estimator is obtained if the estimate converges to the global minimum.

3. MODEL REDUCTION USING SYSTEM IDENTIFICATION

Zhu proposed a minimization method based on output error system identification method [3]. Let $N \to \infty$ and apply Parseval's theorem to Eq.(10), the negative log-likelihood function (10) becomes

$$V = \frac{1}{N}\sum_{k=1}^{N}\left\{\left(\hat{G}^n(q) - \hat{G}^l(q)\right)\frac{u(k)}{\hat{H}^n(q)}\right\}^2 \tag{11}$$

in the time domain. This form corresponds to the output error criterion of system identification in which input and output are $(1/\hat{H}^n)u$ and $(\hat{G}^n/\hat{H}^n)u$, respectively. This can be expressed by

$$\hat{y}^n(k) = \hat{G}^n(q)u_f(k), \tag{12}$$

where

$$u_f(k) = \frac{1}{\hat{H}^n(q)}u(k).$$

Then, we can use an OE model for identification. In a similar way, the reduced disturbance model can be obtained from the high order estimate $\hat{H}^n$. Finally, the model structure becomes OE model (or BJ model including the disturbance model). The model order is selected by minimizing the asymptotic criterion (ASYC)

$$V_{\mathrm{ASYC}} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\left|\hat{G}^n(q) - \hat{G}^l(q)\right|^2 - \frac{n}{N}\frac{\Phi_u}{\Phi_v}\right|\mathrm{d}\omega.$$

4. MODEL REDUCTION USING CURVE FITTING WITH PSO IN THE FREQUENCY DOMAIN

4.1 Particle Swarm Optimization

Particle Swarm Optimization (PSO [4]) is one of the most powerful metaheuristic optimization methods. PSO optimizes an objective function by updating positions and velocities of many particles which search the optimal solution.

4.2 System identification using frequency response based on PSO

Wada and Sugie proposed a system identification method using frequency response based on PSO [5]. This method consists of two steps. First step is to obtain a frequency response of the system. Second step is to minimize the criterion

$$J = \sum_{\omega_n}\gamma_{\omega_n}[(M_{\omega_n} - \hat{M}_{\omega_n})^2 + P_{\omega_n} - \hat{P}_{\omega_n})^2] \tag{13}$$

using PSO, where M_{ω_n} and P_{ω_n} denote gain [dB] and phase [deg] of the system at frequency ω_n, hat denotes the model obtained by PSO and γ_{ω_n} is frequency weighting parameter. PSO optimizes the parameters a_i and b_j $(i = 0, 1, \cdots, n_a,\ j = 0, 1, \cdots, n_b)$ of the transfer function

$$\hat{G}(s) = \frac{b_{n_b}s^{n_b} + \cdots + b_1 s + b_0}{s^{n_a} + \cdots + a_1 s + a_0}. \tag{14}$$

This method is easy to understand. However, the criterion (13) includes two quantities in different units. Thus, this criterion might need scaling if there is large difference between these values.

4.3 Model reduction using curve fitting in the frequency domain

The minimization of the likelihood function (10) can be interpreted as a curve fitting problem with a frequency weighting $\Phi_u(\omega)/\hat{H}^n(e^{j\omega})$. We propose a new model reduction method using PSO based on the idea in [5]. In the new method, the criterion (10) is replaced by

$$V = \sum_{\omega=\omega_L}^{\omega_H}\left|\hat{G}^n(e^{j\omega}) - \hat{G}^l(j\omega)\right|^2 \frac{\Phi_u(\omega)}{\hat{H}(e^{j\omega})} \tag{15}$$

where ω_L and ω_H denote the frequencies of interest. The reduced-order model G^l which has structure (14) is optimized.

4.4 Initial value

From the frequency response of the high-order model, a transfer function of the system can be approximated by modal analysis or broken line approximation. These approximated transfer function is available for the initial value of a_i and b_i in Eq.(14). Initial values of the PSO are arranged around the approximated transfer function.

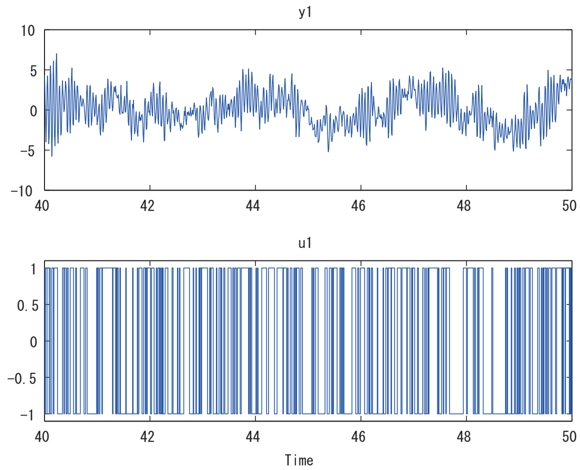

Fig.1 Input and output data of the simulation

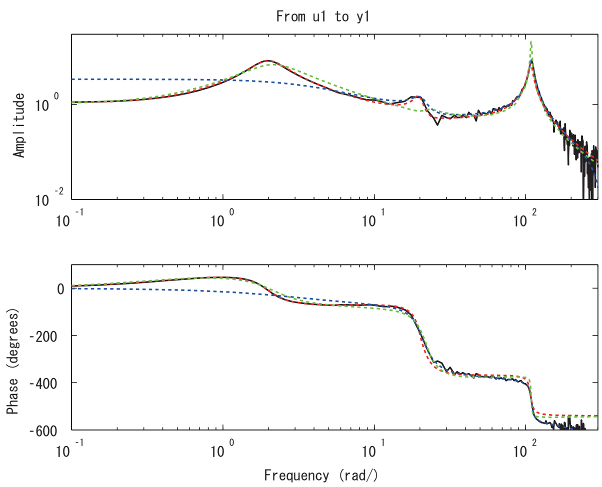

Fig.2 The frequency responses by the proposed method (red), the conventional method (1) (green), the conventional method (2) (blue), and high-order model

5. NUMERICAL EXAMPLES

5.1 Numerical simulation

5.1.1 Identification object

Suppose a continuous-time system described by

$$G(s) = \frac{B(s)}{A(s)}$$

where

$$\begin{aligned} A(s) &= s^6 + 11s^5 + 12438s^4 + 62864s^3 \\ &\quad +4900096s^2 + 5001600s + 19200000 \\ B(s) &= 6000s^4 + 33200s^3 \\ &\quad +2468000s^2 + 40900800s + 19200000. \end{aligned}$$

This system is 6th-order system, and its natural frequencies are $\omega = 2, 20, 110$ rad/s. Because the bandwidth of this system is high and the system has three resonance modes, this system is difficult to identify.

The true system is discretized using Zeroth-Order-Hold with sampling interval $T = 0.1$s. Then, the discrete system

$$G(q) = \frac{B(q)}{A(q)}$$

is obtained. In this section, the output error form

$$y(k) = G(q)u(k) + e(k),$$

is assumed.

5.1.2 System identification condition

The input signal $u(k)$ was a GBN (Generalized Binary Noise) with average switch time of 4 samples. The identification time was 40000 s. The disturbance power was 30 % of that of noise-free output.

The input-output data is shown in Fig.1. First, a 200th-order ARX model was estimated based on the input-output data. The estimated high-order ARX model was reduced by following the three methods:

(i) *The proposed method* : PSO with criterion (15)

Table 1 Comparison of errors in the frequency domain

Method	Mean value of criterion
Proposed method	35.8
Conventional method (1)	317
Conventional method (2)	1100

(ii) *The conventional method* (1) : Output error system identification method [3]

(iii) *The conventional method* (2) : PSO with criterion (13) [5].

These methods were compared by Bode diagram and the square sum of difference between reduced model and high order model at each frequency :

$$V_c = \sum_{\omega=\omega_L}^{\omega_H} |G^n(e^{j\omega}) - G^l(e^{j\omega})|^2, \tag{16}$$

where ω_L=1 rad/s and ω_H=118 rad/s. 245 points of evaluation between ω_L and ω_H were utilized.

It is noted that PSO's parameters were the same values as [5] and the number of particles was 200, the number of iterations was 300. This simulation ran for ten times, and the mean value of (16) was compared.

5.1.3 Identification results

Frequency responses of the estimated models are shown in Fig.2. The proposed method could identify the frequency response of high-order model over wide frequency range. However, the conventional method (1) couldn't identify in the low frequency and the conventional method (2) couldn't identify the second resonance peak of the system.

The mean values of the evaluation criterion (16) are summarized in Table 1. For the narrow frequency band, the effectiveness of the conventional method (1) is validated [3]. However, for the broad frequency band like this simulation case, the accuracy of the method deteriorates because this method is based on discrete system identification which is influenced by sampling interval.

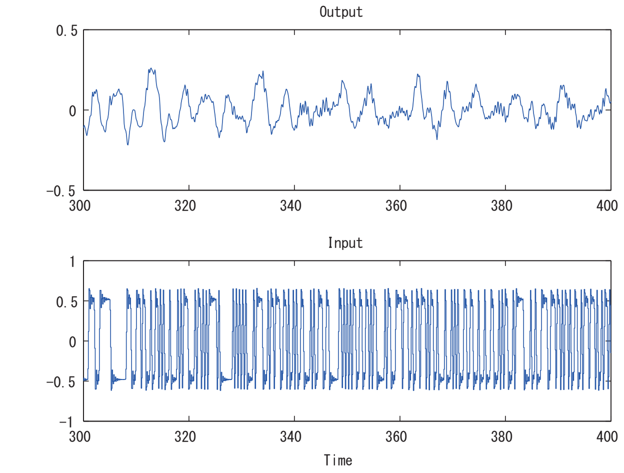

Fig.3 Input and Output data of a real system

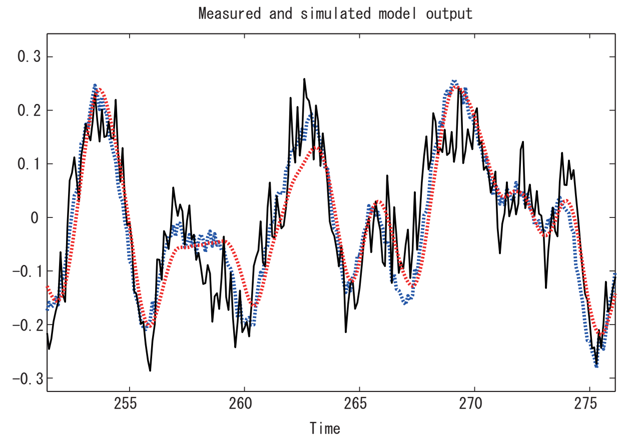

Fig.5 Comparison of outputs : measured (black), high-order model (blue), and low-order model (red).

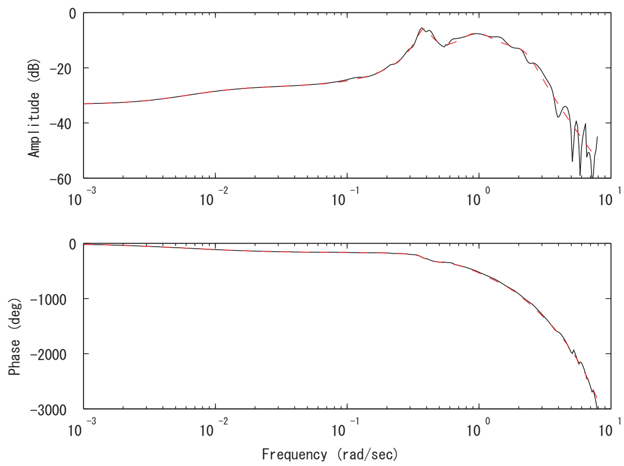

Fig.4 The frequency responses high-order model (black) and the reduced model (red)

5.2 Case study

The input and output data which were obtained by system identification test for a real system, are shown in Fig.3. This system had characteristics that the operating frequency range is broad. Thus, the frequency region was divided and 100th and 200th ARX model were estimated for low (1Hz) and high (10Hz) frequencies, respectively. The frequency responses of these models are combined and reduced by the proposed method. The reduced model was 8th order model whose frequency response was shown in Fig.4. Then, we compared measured output with simulated output of high-frequency model and reduced model, which are shown in Fig.5. A fitting rates of the high-order ARX model and the reduced model were 50% and 43%, respectively. It is considered that the proposed method works well for the real system.

6. CONCLUSION

We proposed the new method of model reduction for the minimization problem of asymptotic likelihood function. This method has two advantages. The first is that the method can identify the continuous-time model which is not affected by sampling interval. So, the method can be used for systems which have broad frequencies of interest. The second is that the asymptotic method assures the maximum likelihood estimator. However, this can be interpreted that if PSO does not converge to the global minimum, the estimate is not maximum likelihood estimator. In fact, the ordinary PSO does not always converge to the global minimum. So, when we use this approach in practice, we must confirm that the frequency response of the estimated model sufficiently fits the frequency response of the high-order model.

REFERENCES

[1] L.Ljung and Z.D.Yuan : Asymptotic properties of black-box identification of transfer function, IEEE Trans on Automatic Control, Vol.AC-30, No.6, pp.514-530 (1985)
[2] B.Wahlberg : Model reduction of high-order estimated models : Asymptotic ML approach, Int.J. Control, Vol.49, No.1, pp.169-192 (1989)
[3] Y.Zhu : Multivariable System Identification for Process Control, Elsevier Science (2001)
[4] J.Kennedy and R.Eberhart : Particle swarm optimization, Proceedings of IEEE International Conference on Neural Networks, pp. 1942-1948 (1995)
[5] T.Wada and T.Sugie : System identification using frequency responce based on particle swarm optimization, Proceeding of the 52nd Annual Conference of the ISCIE, Kyoto (2008)