A Multi-model Identification Algorithm Based on Weighted Cost Function and Application in Thermal Process[1)]

XUE Zhen-Kuang LI Shao-Yuan

(*Institute of Automation, Shanghai Jiaotong University, Shanghai* 200030)

(E-mail: syli@sjtu.edu.cn)

Abstract Most existing algorithms for identifying multi-model system are based on minimizing the square of bias between global outputs of the actual system and the identified model, but the resultant model lacks of robustness. In order to solve this problem, this paper considers some other algorithms in which local models are identified independently and presents a multi-model identification algorithm based on weighted cost function, which uses the idea of local weighted regression and local approximation while keeps the model structure of global identification algorithm. The result of application to a 300MW unit boiler superheater illustrates that the multi-model generated by the proposed algorithm has better trade-off between global fitting and local interpretation.

Key words Multi-model, identification, nonlinear, local model network, thermal process

1 Introduction

Local model network (LMN) is a typical multi-model used for modeling nonlinear systems, in which global outputs are described as the sum of local ones multiplied by their corresponding basis functions[1]. Usually modeling LMN can be divided into two tasks: first to find optimal number and position of basis functions using clustering algorithm, then to identify the parameters of local models. The clustering algorithms used for modeling LMN include fuzzy mean (FCM) clustering, satisfactory fuzzy clustering (SFC)[2] and so on. In this paper we will discuss identifying problem for the parameters of local models.

Most existing algorithms for identifying LMN model are based on minimizing the square of bias between global outputs of the actual system and the identified model. Because of using linear models to approximate nonlinearity of the actual system in neighbor of operating points, the structure mismatch will result in "oscillating" local models and make the resultant LMN model generalize poorly[3]. Some other algorithms identify the parameter of each local model independently to guarantee individual local models to be local approximations of the underlying system, but the resultant LMN model is difficult to have good global fitting[3].

In order to solve the above problems, this paper presents a multi-model identification algorithm based on weighted cost function, which uses the idea of local weighted regression and local approximation, and keeps the model structure of global identification algorithm. The result of application to a 300MW unit boiler superheater illustrates the LMN model generated by the proposed identification algorithm has better trade-off between global fitting and local interpretation.

2 The multi-model system based on LMN

The multi-model system based on LMN can be described as[1]

$$\hat{y}(k) = \sum_{i=1}^{n} \rho_i(\boldsymbol{\phi}(k)) \cdot \hat{f}_i(\boldsymbol{\varphi}(k)) \tag{1}$$

where $\hat{y}(k) \in R$ is the model output at time k, ρ_i are basis functions which depend on some scheduling vector $\boldsymbol{\phi}(k) \in R^{n_\phi}$, $\hat{f}_i$ are local models which depend on the regressive vector $\boldsymbol{\varphi}(k) \in R^{n_\varphi}$. The number of basis functions is given by n.

1) Supported by National Natural Science Foundation of P. R. China (60474057) and partly by the Key Technology and Development of Shanghai Science and Technology Department (04DZ11008) and the Specialized Research Fund for the Doctoral Program of Higher Education of P. R. China (20020248028)

Received August 14, 2003; in revised form June 17, 2004

Usually, Gaussian bell is chosen as basis function, which can be described as

$$\tilde{\rho}_i(\boldsymbol{\phi}) = \exp\left[\frac{(\boldsymbol{\phi}-\boldsymbol{c}_i)^{\mathrm{T}}(\boldsymbol{\phi}-\boldsymbol{c}_i)}{s_i^2}\right] \tag{2}$$

where $\boldsymbol{c}_i \in R^{n_\phi}$ are center vectors and s_i determine the width of the bell.

To achieve a partition of unity of the scheduling space, the basis functions have to be normalized, *i.e.*,

$$\rho_i(\boldsymbol{\phi}) = \frac{\tilde{\rho}_i(\boldsymbol{\phi})}{\sum\limits_{i=1}^{n}\tilde{\rho}_i(\boldsymbol{\phi})} \tag{3}$$

Because MIMO model can always be transformed into multiple MISO models, here we choose local models $\hat{f}_i$ in form of MISO discrete time CARIMA model, which can be described as

$$A_i(z^{-1})y(k) = \boldsymbol{B}_i\boldsymbol{u}(k-1) \tag{4}$$

where $\begin{cases} A_i(z^{-1}) = 1 + a_{i1}z^{-1} + \cdots + a_{in_A}z^{-n_A} \\ \boldsymbol{B}_i(z^{-1}) = \boldsymbol{B}_{i0} + \boldsymbol{B}_{i1}z^{-1} + \cdots + \boldsymbol{B}_{in_B}z^{-n_B} \end{cases}$, $\boldsymbol{B}_{ij} \in R^{1\times M}(j = 0,\cdots,n_B)$ are row vectors and M is the dimension of inputs, n_A, n_B are the order of model output and inputs, respectively.

Substituting (4) into (1), we can get a different form of LMN model as

$$\hat{y}(k) = \sum_{i=1}^{n}\rho_i(\boldsymbol{\phi}(k))\boldsymbol{\varphi}^{\mathrm{T}}(k)\cdot\boldsymbol{\theta}_i \tag{5}$$

where

$$\boldsymbol{\varphi}^{\mathrm{T}}(k) = [-y(k-1) \quad \cdots \quad -y(k-n_A) \quad \boldsymbol{u}^{\mathrm{T}}(k-1) \quad \cdots \quad \boldsymbol{u}^{\mathrm{T}}(k-n_B-1)]$$
$$\boldsymbol{\theta}_i = [a_{i1} \quad \cdots \quad a_{in_A} \quad \boldsymbol{B}_{i0} \quad \cdots \quad \boldsymbol{B}_{in_B}]^{\mathrm{T}}$$

3 Multi-model identification algorithm

The problem of multi-model identification is to determine the parameters n, $\boldsymbol{c}_i$, s_i and $\boldsymbol{\theta}_i$ in (5). To avoid identifying all parameters at one operation using complex nonlinear optimization algorithm, the parameters n, $\boldsymbol{c}_i$, s_i can be determined beforehand using clustering algorithm or prior knowledge[2]. Therefore, multi-model identification for remaining parameters $\boldsymbol{\theta}_i$ is reduced to a linear optimization problem, which can be solved by effective linear learning algorithm[4].

3.1 Previous algorithm for multi-model identification

Most existing algorithms identify the parameters of all local models at the same time by minimizing the square of bias between global outputs of the actual system and the identified multi-model. Because of using linear models to approximate the nonlinearity of the actual system in neighbor of operating points, the structure mismatch will result in "oscillating" local models and make the resultant LMN model generalize poorly[3]. The following global least squares cost function can be used.

$$J_G(\boldsymbol{\Theta}) = \sum_{k=1}^{N}(y(k) - \hat{y}(k,\boldsymbol{\Theta}))^2 = (\boldsymbol{Y} - \boldsymbol{\Psi\Theta})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{\Psi\Theta}) \tag{6}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1^{\mathrm{T}} \quad \boldsymbol{\theta}_2 \quad \cdots \quad \boldsymbol{\theta}_n^{\mathrm{T}}]$, $\boldsymbol{Y} = [y(1) \quad y(2) \quad \cdots \quad y(N)]^{\mathrm{T}}$

and $\boldsymbol{\Psi} = \begin{bmatrix} \rho_1(1)\boldsymbol{\varphi}^{\mathrm{T}}(1) & \rho_2(1)\boldsymbol{\varphi}^{\mathrm{T}}(1) & \cdots & \rho_n(1)\boldsymbol{\varphi}^{\mathrm{T}}(1) \\ \rho_1(2)\boldsymbol{\varphi}^{\mathrm{T}}(2) & \rho_2(2)\boldsymbol{\varphi}^{\mathrm{T}}(2) & \cdots & \rho_n(2)\boldsymbol{\varphi}^{\mathrm{T}}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1(N)\boldsymbol{\varphi}^{\mathrm{T}}(N) & \rho_2(N)\boldsymbol{\varphi}^{\mathrm{T}}(N) & \cdots & \rho_n(N)\boldsymbol{\varphi}^{\mathrm{T}}(N) \end{bmatrix}$

Some other algorithms identify the parameters of each local model independently to guarantee individual local models to be local approximations of the underlying system, but the resultant LMN model is difficult to have good global fitting[3]. The local least squares cost function for th local model can be written as follows.

$$J_i(\boldsymbol{\theta}_i) = \sum_{k=1}^{N_i}\rho_i(k)(y_i(k) - \hat{y}_i(k,\boldsymbol{\theta}_i))^2 \tag{7}$$

In local cost function, each data sample is used to identify only one local model that is chosen in terms of maximal value of basis function. We can rewrite the local cost function in a different form as

$$J_L(\boldsymbol{\Theta}) = (\boldsymbol{Y} - \boldsymbol{\Psi}'\boldsymbol{\Theta})^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{Y} - \boldsymbol{\Psi}'\boldsymbol{\Theta}) \tag{8}$$

where $\boldsymbol{\Psi}' = \begin{bmatrix} \tau_1(1)\boldsymbol{\varphi}^{\mathrm{T}}(1) & \tau_2(1)\boldsymbol{\varphi}^{\mathrm{T}}(1) & \cdots & \tau_n(1)\boldsymbol{\varphi}^{\mathrm{T}}(1) \\ \tau_1(2)\boldsymbol{\varphi}^{\mathrm{T}}(2) & \tau_2(2)\boldsymbol{\varphi}^{\mathrm{T}}(2) & \cdots & \tau_n(2)\boldsymbol{\varphi}^{\mathrm{T}}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \tau_1(N)\boldsymbol{\varphi}^{\mathrm{T}}(N) & \tau_2(N)\boldsymbol{\varphi}^{\mathrm{T}}(N) & \cdots & \tau_n(N)\boldsymbol{\varphi}^{\mathrm{T}}(N) \end{bmatrix}$, $\tau_j(\cdot) = \begin{cases} 1, & \text{if } j = \arg\max_i \rho_i(\cdot) \\ 0, & \text{else} \end{cases}$

$\boldsymbol{\Theta}$, $\boldsymbol{Y}$ are the same as (6), $\boldsymbol{Q} = diag(\max_i \rho_i(1) \quad \max_i \rho_i(2) \quad \cdots \quad \max_i \rho_i(N))$.

3.2 Multi-model identification algorithm based on weighted cost function

After fully considering the advantage and disadvantage of global and local cost functions, we can see that complementation for each other exists, so this paper presents a multi-model identification algorithm based on weighted cost function, which uses the idea of local weighted regression and local approximation in local cost function, and keeps the model structure in global cost function. The resultant LMN model has better trade-off in terms of global fitting and local interpretation. The weighted cost function is defined as

$$J_W = \alpha J_G + (1-\alpha)J_L \tag{9}$$

where α is a constant satisfying $0 \leqslant \alpha \leqslant 1$. The value of α can be determined according to user's need. If user prefers to a better global performance, he can choose α with a larger value, or vice versa.

Substituting (6) and (8) into (9), we can get

$$J_W(\boldsymbol{\Theta}) = \alpha(\boldsymbol{Y} - \boldsymbol{\Psi}\boldsymbol{\Theta})^{\mathrm{T}}(\boldsymbol{Y} - \boldsymbol{\Psi}\boldsymbol{\Theta}) + (1-\alpha)(\boldsymbol{Y} - \boldsymbol{\Psi}'\boldsymbol{\Theta})^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{Y} - \boldsymbol{\Psi}'\boldsymbol{\Theta}) \tag{10}$$

Minimizing the weighted cost function with $\partial J_W(\boldsymbol{\Theta})/\partial\boldsymbol{\Theta} = 0$, we can get

$$\boldsymbol{\Theta} = (\alpha\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi} + (1-\alpha)\boldsymbol{\Psi}'^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\Psi}')^{-1}(\alpha\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{Y} + (1-\alpha)\boldsymbol{\Psi}'^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{Y}) \tag{11}$$

3.3 Multi-model validation

The global performance of multi-model can be measured by standard global output deviation, which is described as

$$\varepsilon = \frac{1}{N}\sum_{k=1}^{N}(y(k) - \hat{y}(k))^2 \tag{12}$$

where $y(k), \hat{y}(k)$ are global outputs of the actual system and the resultant multi-model, respectively.

The transfer function of MISO system can be described as

$$Y(s) = G_1(s)U_1(s) + G_2(s)U_2(s) + \cdots + G_{n_u}(s)U_{n_u}(s)$$

The local performance of each local model can be measured by frequency domain identification error, which is defined as[5]

$$E = \max_{\omega\in[0,\omega_c]}\left\{\sum_{i=1}^{n_u}\left|\frac{\hat{G}_i(j\omega) - G_i(j\omega)}{G_i(j\omega)}\right| \times 100\%\right\} \tag{13}$$

where $G_i(j\omega)$ and $\hat{G}_i(j\omega)$ are frequency responses of the actual system and the resultant multi-model, respectively, $\omega_c = \max_i(\omega_{ci})$ for $i = 1, 2, \cdots, n_u$, with ω_{ci} satisfying $\angle G_i(j\omega) = -\pi$.

4 Application to thermal process

We will apply the proposed multi-model identification algorithm to a 300MW unit boiler superheater steam temperature controlled system. In this system, each superheater process is equipped with an attemperator device (spray water injection at the inlet) for controlling outlet temperature of superheater steam, as shown in Fig. 1. The first attemperator is used to roughly regulate outlet steam temperature of the secondary superheater, while the final controlling performance of outlet steam temperature (of third superheater) is mainly determined by the secondary attemperator. We just consider the secondary attemperator controlled process (including secondary attemperator and third superheater) and take it as example in this paper. The process variable to be controlled is the steam temperature at the outlet of the third superheater, and the manipulated variable is the command of spray water valve in the secondary attemperator.

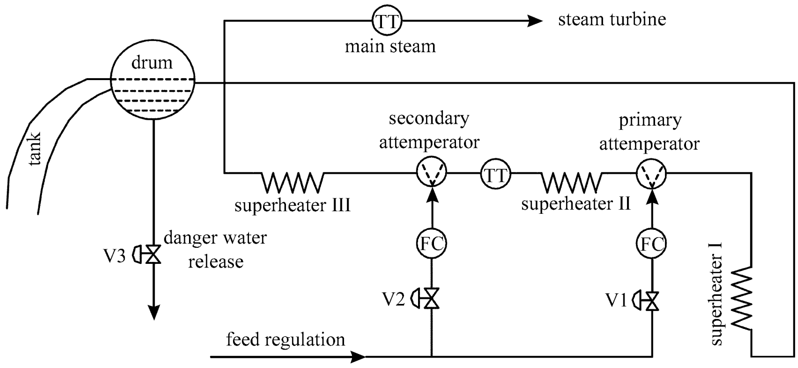

Fig. 1 The framework of 300MW unit boiler

Load cycling operation of thermal power plants leads to changes in operating point right cross the whole operating range[6]. This results in nonliearity dynamic property in the secondary attemperator controlled process. This paper uses multi-model to account for nonlinearity, and generates LMN model for controlled process. Taking into consideration a priori knowledge about process dynamics and the identified LMN model, the load level of power plant is used as scheduling vector[6] and pairs of Gaussian parameters $[c_i, s_i]$ are chosen as [180,20], [200,20], [220,20], [240,20], [260,20], [280,20], [300,20]. The sampling time is $T_0 = 5s$ and the value of α is chosen as $\alpha = 0.7$. The parameters of local models are identified using multi-model identification algorithm based on global cost function, local cost function and weighted cost function, respectively. The resultant local models identified by the proposed algorithm are listed in Table 1.

Table 1 Local models under different load levels identified by the proposed algorithm

Load Level (MW)	Local models identified by the proposed algorithm ($\alpha = 0.7$)
180	$\dfrac{-0.8902\times10^{-5}z^2-3.3174\times10^{-5}z-0.8691\times10^{-5}}{z^3-2.8498z^2+2.7068z-0.85688}$
200	$\dfrac{-1.9921\times10^{-5}z^2-8.0717\times10^{-5}z-1.7621\times10^{-5}}{z^3-2.8137z^2+2.6394z-0.8255}$
220	$\dfrac{-3.087\times10^{-5}z^2-11.294\times10^{5}z-2.694\times10^{-5}}{z^3-2.7833z^2+2.582z-0.7983}$
240	$\dfrac{-3.049\times10^{-5}z^2-11.896\times10^{-5}z-2.8218\times10^{-5}}{z^3-2.762z^2+2.5429z-0.7804}$
260	$\dfrac{-2.8884\times10^{-5}z^2-10.445\times10^{-5}z-2.5504\times10^{-5}}{z^3-2.7398z^2+2.5021z-0.7616}$
280	$\dfrac{-2.4262\times10^{-5}z^2-9.4259\times10^{-5}z-2.0605\times10^{-5}}{z^3-2.7124z^2+2.4525z-0.7392}$
300	$\dfrac{-3.6582\times10^{-5}z^2-13.614\times10^{-5}z-3.0831\times10^{-5}}{z^3-2.6718z^2+2.3803z-0.7072}$

The global performances of LMN models identified by different algorithms can be calculated using (12), and the results are listed in Table 2. The LMN model identified by multi-model identification algorithm based on global cost function (global algorithm) has better global fitting for identification data, but generalizes poorly for validation data, while the one identified by multi-model identification algorithm based on weighted cost function (weighted algorithm) obtains obvious improvement in this aspect, and has good global fitting for both identification data and validation data.

Table 2 Standard global output deviation for LMN models identified by different identification algorithms

	Global algorithm	Local algorithm	Weighted algorithm
ε (for identification data)	8.93×10^{-4}	4.55×10^{-1}	2.98×10^{-3}
ε (for validation data)	3.89×10^{-1}	7.72×10^{-1}	3.12×10^{-3}

Using (13), frequency domain identification errors for local models under different load levels identified by global algorithm and weighted algorithm respectively can be calculated, as shown in Table 3. Global algorithm can not guarantee all resultant local models to be local approximations of the underlying system, even results in large deviation between some local models and corresponding local properties of underlying system. The local models identified by weighted algorithm have good local performance under all load levels. This further illustrates the validity of the proposed multi-model identification algorithm.

Table 3 Frequency domain identification errors for local models under different load levels

	180MW	200MW	220MW	240MW	260MW	280MW	300MW
E (global algorithm)	37.13%	4.09%	7.26%	16.04%	14.51%	13.33%	39.26%
E (weighted algorithm)	7.15%	3.40%	3.69%	1.84%	0.71%	5.30%	5.84%

5 Conclusion

The multi-model system based on LMN describes nonlinear dynamic system as the sum of local linear models multiplied by their corresponding basis functions. To control the underlying system effectively, it is important for LMN model to possess good global fitting and local interpretation. After fully considering the advantage and disadvantage of existing algorithms, we present a multi-model identification algorithm based on weighted cost function, which uses the idea of local weighted regression and local approximation, and keeps the model structure of global identification algorithm. The result of application to a 300MW unit boiler superheater illustrates the validity of the proposed algorithm.

References

1 Johansen T A, Foss B A. A NARMAX model representation for adaptive control based on local models. *Modeling, Identification and Control*, 1992, **13**(1): 25~39

2 Li Ning, Li Shao-Yuan, Xi Yu-Geng. Multiple model predictive control for MIMO systems. *Acta Automatica Sinica*, 2003, **29**(4): 516~523

3 Murry-Smith R, Johansen T A. Multiple model approaches to modeling and control. London: Taylor and Francis, 1997

4 Boukhris A, Mourot G, Ragot J. Non-linear dynamic system identification: a multi-model approach. *International Journal of Control*, 1999, **72**(7/8): 591~604

5 Wang Qing-Guo, Zhang Yong. Robust identification of continuous systems with dead-time from step responses. *Automatica*, 2001, **37**(3): 377~390

6 Prasad G, Swidenbank E, Hogg B W. A local model networks based multivariable long-range predictive control strategy for thermal power plants. *Automatica*, 1998, **34**(10): 1185~1204

XUE Zhen-Kuang Received his master degree from Jiangsu University in 2001. Now he is a Ph. D. candidate in the Institute of Automation at Shanghai Jiaotong University. His research interest includes multi-model modeling and control, and predictive control.

LI Shao-Yuan Received his Ph. D. degree from Nankai University in 1997. Now he is a professor in the Institute of Automation at Shanghai Jiaotong University. His research interests include fuzzy systems and nonlinear system control.