Error Analysis in Nonlinear System Identification Using Fuzzy System*

WANG Shi-tong YU Dong-jun

(Department of Computer Science East China Shipbuilding Institute Zhenjiang 212003)
E-mail: zjstwang@public.zj.js.cn

Abstract In this paper, the numerical approximation characteristics of fuzzy system are discussed, and the influence of approximation error and initial state error on fuzzy system are analyzed. Finally, an important conclusion is obtained that under some conditions, fuzzy system output differs little from that of the actual system.

Key words Universal approximation, fuzzy system, system identification, approximation error.

The fuzzy sets theory was introduced by Zadeh in 1965. Since then, fuzzy sets theory has been widely used in many fields. Nonlinear system identification[1~3] using fuzzy system is one of the most important application branches. Many researchers have studied this topic, but fewer have studied error analysis in nonlinear system identification[1,2,4,5] using fuzzy system. This paper will discuss this question.

Generally, a fuzzy system consists of a set of fuzzy rules. In this paper, we consider an MISO (multi input single output) fuzzy system as the following:

Rule 1. IF x_1 is A_1^1 AND x_2 is A_2^1 AND ... x_n is A_n^1, THEN y is B^1 ELSE

Rule 2. IF x_1 is A_1^2 AND x_2 is A_2^2 AND ... x_n is A_n^2, THEN y is B^2 ELSE

.....

Rule N. IF x_1 is A_1^N AND x_2 is A_2^N AND ... x_n is A_n^N, THEN y is B^N ELSE

Fact: x is A'_1 AND x_2 is A'_2 AND ... x_n is A'_n

Conclusion: y is B'

According to Dr. L. X. Wang, using max produce inference and centroid defuzzification method, the final numerical output can be described as follows:

$$\hat{y}=\frac{\int \mu_{B'}(y)ydy}{\int \mu_{B'}(y)dy}, \tag{1}$$

where $$\mu_{B'}(y)=\bigvee_{x_1,x_2,\ldots,x_n}\left[\left(\prod_{i=1}^{n}\mu_{A'_i}(x_i)\right)\cdot\left(\bigvee_{j=1}^{N}\left(\prod_{i=1}^{n}\mu_{A_i^j}(x_i)\right)\cdot\mu_{B^j}(y)\right)\right].$$

In practice, especially in control applications, we consider that the output fuzzy sets B^j are singletons β^j, i.e.,

---

* This research is supported by the National Natural Science Foundation of China (国家自然科学基金, No. 69673001). **WANG Shi-tong** was born in 1964. He is the head and professor of Department of Computer Science of East China Shipbuilding Institute. His research interests include artificial intelligence, fuzzy mathematics, and neural network. **YU Dong-jun** was born in 1973. He is a Ph. D. student now. His research interests include fuzzy system, artificial intelligence, and neural network.

Manuscript received 1998-07-06, accepted 1998-08-25.

$$\mu_{B^j}(y)=\begin{cases}1, & \text{if } y=\beta^j\\ 0, & \text{otherwise}\end{cases},\quad j=1,2,\dots,N, \tag{2}$$

then we have

$$\mu_{B'}(y)=\begin{cases}\prod_{i=1}^{n}\mu_{A_i^j}(x_i), & \text{if } y=\beta_j\\ 0, & \text{oterwise}\end{cases},\quad j=1,2,\dots,N;i=1,2,\dots,n. \tag{3}$$

Thus the final output can be rewritten as follows:

$$\hat{y}=\frac{\sum_{j=1}^{N}\beta^j\left(\prod_{i=1}^{n}\mu_{A_i^j}(x_i)\right)}{\sum_{j=1}^{N}\left(\prod_{i=1}^{n}\mu_{A_i^j}(x_i)\right)}. \tag{4}$$

Let $\Phi_j(\vec{x})=\dfrac{\prod_{i=1}^{n}\mu_{A_i^j}(x_i)}{\sum_{k=1}^{N}\left(\prod_{i=1}^{n}\mu_{A_i^k}(x_i)\right)}$, $j=1,2,\dots,N$, where $\vec{x}=(x_1,x_2,\dots,x_n)^T$. Thus Eq. (4) can be rewritten as

$$\hat{y}=\hat{f}(\vec{x})=\sum_{j=1}^{N}\beta^j\Phi_j(\vec{x}). \tag{5}$$

Dr. L. X. Wang has proved Eq. (5) is a universal approximator.

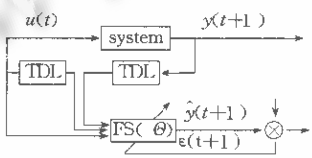

Fig. 1 Structure of nonlinear system identification using fuzzy system (FS)

In real applications, we can use a fuzzy system to identify the real system (see Fig. 1, $u(t)$ is the activation function, TDL and Θ are the time delay logic and the parameter of fuzzy system, respectively). When the fuzzy system approximates the real system well enough, it can be applied independently. This method has been widely used, but there still are several questions to be studied, for example, (a) the numerical approximation characteristics of fuzzy system; (b) the influence of approximation error and initial state error on fuzzy system.

The rest of this paper is organized as follows. Section 1 studies the numerical approximation characteristics of fuzzy system. Section 2 studies the influence of approximation error and initial state error on fuzzy system. Section 3 deals with simulation analysis. Section 4 concludes this paper.

1 Numerical Approximation Characteristics of Fuzzy System

As we all know, in the application when we use samples to train a fuzzy system, it is a procedure of ascertaining the parameters of the fuzzy system, and it is also a procedure of interpolation. During the application, first of all, one should ascertain the number of rules. Is there a theorem to direct the selection of the number of rules? Theorem 1 answers this question.

Theorem 1. Give N samples $\{\vec{x}_i,y_i^d\}$, $\vec{x}_i\in R^n$, $y_i^d\in R$, $i=1,2,\dots,N$. $\vec{x}_i\neq\vec{x}_j$ when $i\neq j$, then there must exist a fuzzy system (as described by Eq. (5)) which contains N rules satisfying

$$\hat{y}_i=\hat{f}(\vec{x})_i=y_i^d, \tag{6}$$

where $\hat{y}_i$ is the actual output of the fuzzy system, y_i^d is the desired output.

Proof. We have defined $\Phi_j(\vec{x})=\dfrac{\sum_{i=1}^{n}\mu_{A_i^j}(x_i)}{\sum_{k=1}^{N}\left(\prod_{i=1}^{n}\mu_{A_i^k}(x_i)\right)}$, $j=1,2,\dots,N$, $\hat{y}=\sum_{j=1}^{N}\beta^j\Phi_j(\vec{x})$. We use Guassian membership function, i.e.,

$$\mu_{A_i^j}(x_i)=\exp\left(-\frac{(x_i-a_i^j)^2}{2(b_i^j)^2}\right).$$

We can easily see that the following two assertions hold:

Assertion 1. $\{\Phi_j(\vec{x})\}$ $(j=1,2,\ldots,N)$ is a partition of unity and linearly independent.

Assertion 2. Let $S_N(\vec{x}_i)\triangleq[\Phi_1(\vec{x}_i),\Phi_2(\vec{x}_i),\ldots,\Phi_N(\vec{x}_i)]$, $i=1,2,\ldots,N$, then it is linearly independent too.

Given N samples $\{\vec{x}_i,y_i^d\}$, $i=1,2,\ldots,N$, we can use the following matrix equation to uniformly represent the relationship between the inputs and the outputs in the fuzzy system.

$$G_k\cdot\begin{pmatrix}\beta^1\\ \beta^2\\ \cdots\\ \beta^k\end{pmatrix}=\begin{pmatrix}\Phi_1(\vec{x}_1) & \Phi_2(\vec{x}_1) & \ldots & \Phi_k(\vec{x}_1)\\ \Phi_1(\vec{x}_2) & \Phi_2(\vec{x}_2) & \ldots & \Phi_k(\vec{x}_2)\\ \cdots & \cdots & \cdots & \cdots\\ \Phi_1(\vec{x}_N) & \Phi_2(\vec{x}_N) & \ldots & \Phi_k(\vec{x}_N)\end{pmatrix}\cdot\begin{pmatrix}\beta^1\\ \beta^2\\ \cdots\\ \beta^k\end{pmatrix}=\begin{pmatrix}y_1^d\\ y_2^d\\ \cdots\\ y_N^d\end{pmatrix}, \tag{7}$$

where k indicates the number of rules. According to assertions 1 and 2, when $k=N$, G_k is full rank, i.e. $(\beta^1,\beta^2,\ldots,\beta^k)^T$ in Eq. (7) must have unique solution, and this unique solution is a function with G_k as its parameters. As a_i^j,b_i^j, which make G_k be full rank, are not unique, so the solution $(\beta^1,\beta^2,\ldots,\beta^k)$ in Eq. (7) is not unique.

From above we can draw the following conclusion. When the number of rules equals is that of samples, we will be able to realize accurate interpolation. So the number of rules should not exceed that of samples, otherwise it may cause too much training, even oscillation, thus reducing the generalization ability of fuzzy system. This theorem provides the criterion for the selection of rules.

2 Influence of Approximation Error and Initial State Error on Fuzzy System

2.1 Fuzzy identification of deterministic nonlinear system

A deterministic SISO (single input single output) discrete time system can be formalized as:

$$y(t+1)=f(y(t),\ldots,y(t-n_y+1),u(t),\ldots,u(t-n_u)), \tag{8}$$

where $u(t)\in U\subset R$, $y(t)$ are the input and output of the system respectively. n_u+1, n_y are their maximum delay time, respectively. Initial state value of the system at time 0 is

$$Y_0=[y(0),\ldots,y(-n_y+1),u(0),\ldots,u(-n_u)]^T, \tag{9}$$

when $t\leqslant 0$, $y(t)=y_{-t}\in Y\subset R^{n_y}$, $f(\cdot)$ is continuous function. The structure of fuzzy identification of a nonlinear system can be described as Fig. 1.

Suppose we have got the fuzzy system of the actual system:

$$\hat{y}(t+1)=\hat{f}(y(t),\ldots,y(t-n_y+1),u(t),\ldots,u(t-n_u),\Theta), \tag{10}$$

where Θ is the parameter of the fuzzy system.

Dr. L. X. Wang has proved a fuzzy system can approximate any continuous function defined on compact set to any prescribed accuracy. That is to say, given $\forall\ \varepsilon_f>0$, $u(\cdot)\in U$, $y_{-t}\in Y$, we can find Θ^*, which satisfies the following inequality:

$$\max\left\|\begin{array}{l}f(y(t),\ldots,y(t-n_y+1),u(t),\ldots,u(t-n_u))\\ -\hat{f}(y(t),\ldots,y(t-n_y+1),u(t),\ldots,u(t-n_u),\Theta^*)\end{array}\right\|<\varepsilon_f. \tag{11}$$

Generally, after Θ^* is found, the fuzzy system can be independently applied instead of the actual system, that is,

$$\hat{y}(t+1)=\hat{f}(\hat{y}(t),\ldots,\hat{y}(t-n_y+1),u(t),\ldots,u(t-n_u),\Theta^*) \tag{12}$$

and the initial state value of fuzzy system $\hat{Y}_0=[\hat{y}(0),\dots,\hat{y}(-n_y+1),u(0),\dots,u(-n_u)]^T$.

Now, there are two questions to be considered. (1) Obviously, there exists approximation error between $\hat{f}(\cdot)$ and $f(\cdot)$. (2) The initial state values of $\hat{f}(\cdot)$ and $f(\cdot)$ in actual applications can not be strictly equal, which may affect the value of $\hat{y}(t)$ at time t. As both (1) and (2) exist, then is the fuzzy system still effective?If so, what are conditions? We will discuss these questions in the following section.

2.2 The influence of approximation error and initial state error on fuzzy system

Let state variables $x_i(t)=y(t-n_y+i)$, $x_{n_y+j}=u(t-n_u+j-1)$, $(i=1,2,\dots,n_y;\ j=1,2,\dots,n_u)$, respectively. Then the actual system (8) can be rewritten as the following state equation and output equation:

$$\begin{cases} x(t+1)=F(x(t),u(t)) \\ y(t)=x_{n_y}(t) \end{cases}, \tag{13}$$

where $F(x(t),u(t))=[x_2(t),\dots,x_{n_y}(t),f(x(t),u(t)),x_{n_y+2}(t),\dots,u(t)]^T$, $x(0)=Y_0$. Let $\hat{x}_i(t)=\hat{y}(t-n_y+i)$, $\hat{x}_{n_y+j}(t)=u(t-n_u+j-1)$, $(i=1,2,\dots,n_y;\ j=1,2,\dots,n_u)$, then we can get state equation and output equation of (12) as follows:

$$\begin{cases} \hat{x}(t+1)=\hat{F}(\hat{x}(t),u(t),\Theta^*) \\ \hat{y}(t)=\hat{x}_{n_y}(t) \end{cases}, \tag{14}$$

where $\hat{F}(\hat{x}(t),u(t))=[\hat{x}_2(t),\dots,\hat{x}_{n_y}(t),\hat{f}(\hat{x}(t),u(t)),\hat{x}_{n_y+2}(t),\dots,u(t)]^T$, $\hat{x}(0)=\hat{Y}_0$.

Suppose that the actual system satisfies the condition described below:

Condition:

The partial derivative of any variable of the continuous function

$$f(y(t),\dots,y(t-n_y+1),u(t),\dots,u(t-n_u));R^{n_y-n_u+1}\to R$$

in defined space satisfies

$$L=\max\left[\left|\frac{\partial f}{\partial y(t-n_y+i)}\right|,\left|\frac{\partial f}{\partial u(t-n_u+j-1)}\right|\right]<\infty, \tag{15}$$

where $i=1,\dots,n_y$; $j=1,2,\dots,n_u+1$; $t\in[0,M]$; $u(t)\in U$, $y_{-i}\in Y$. M is a finite positive integer. Then the influence of approximation error and initial state error on fuzzy system can be described using the following theorem.

Theorem 2. Given $\forall\ \varepsilon>0$, if for the actual system (8) and fuzzy system (12),

(1) the actual system satisfies the above condition;

(2) the initial state error between the actual system (8) and the fuzzy system (12) satisfies $\max\|Y_0-\hat{Y}_0\|<r$ (r is defined in Eq. (18)), then there must exist Θ^* (parameter of fuzzy system) which satisfies

$$\max\|y(t)-\hat{y}(t)|<\varepsilon,$$

where $t\in[0,M]$.

Proof. From Eqs. (13) and (14), we can get that to any given $t\in[0,M]$, the following inequality holds:

$$\begin{aligned} \|x(t)-\hat{x}(t)\| &= \|F(x(t-1),u(t-1))-\hat{F}(\hat{x}(t-1),u(t-1),\Theta^*)\| \\ &\leqslant \|F(x(t-1),u(t-1))-F(\hat{x}(t-1),u(t-1)\| \\ &\quad+\|F(\hat{x}(t-1),u(t-1))-\hat{F}(\hat{x}(t-1),u(t-1),\Theta^*)\| \end{aligned} \tag{16}$$

According to the mean value theorem, there exists

$$\bar{x}(t-1)=\lambda\cdot x(t-1)+(1-\lambda)\cdot\hat{x}(t-1),\ (0<\lambda<1),$$

which makes the right-hand side of the inequality above satisfy the following equality:

$$\text{the right-hand side}=\left\|\frac{\partial F}{\partial x}\Big|_{\bar{x}(t-1)}(x(t-1)-\hat{x}(t-1))\right\|+\|\bar{F}(t-1)\|, \tag{17}$$

where
$$\frac{\partial F}{\partial x}=\begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \cdots & \cdots & 0 \\ \cdots & & & & & & \cdots \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \cdots & \frac{\partial f}{\partial x_{n_y+1}} & \cdots & \frac{\partial f}{\partial x_{n_y+n_u}} \\ \cdots & & & & & & \cdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$
$$F(t-1)=[0,\ldots,0,f(\hat{x}(t-1),u(t-1))-\hat{f}(\hat{x}(t-1),u(t-1),\Theta^*),0,\ldots,0)^T.$$

Since $f(\cdot)$ satisfies the condition, so $\frac{\partial F}{\partial x}|_{\bar{x}(t-1)}$ is finite. We take matrix norm as $\|\cdot\|_F$, vector norm as $\|\cdot\|_2$ (these two norms are compatible), thus we have
$$\frac{\partial F}{\partial x}|_{\bar{x}(t-1)}\leqslant\sqrt{n_y+n_u-2+(n_y+n_u)L^2}\triangleq L_m.$$

To $\|\bar{F}(t-1)\|$, given $\forall\ \varepsilon_f>0$, there must exist Θ^* that satisfies the following inequality:
$$\|f(\hat{x}(t-1),u(t-1))-\hat{f}(\hat{x}(t-1),u(t-1),\Theta^*)\|<\varepsilon_f.$$
So, $\|\bar{F}(t-1)\|_2<\varepsilon$ holds. As $\|\cdot\|_F$ is compatible with $\|\cdot\|_2$, combined with Eq. (17), we have
$$\|x(t)-\hat{x}(t)\|<\left\|\frac{\partial F}{\partial x}\right\|_{\bar{x}(t-1)}\cdot\|x(t-1)-\hat{x}(t-1)\|+\|F(t-1)\|$$
$$<\|L_m\|\cdot\|x(t-1)-\hat{x}(t-1)\|+\varepsilon_f.$$
The following inequality can be deduced by analogy:
$$\|x(t)-\hat{x}(t)\|<\sum_{n=0}^{t-1}L_m^n\cdot\varepsilon_f+L_m^t\|x(0)-\hat{x}(0)\|.$$
Given $\forall\ \varepsilon>0$, when we select the parameters of fuzzy system such that the following inequality holds
$$\varepsilon_f<\left(\sum_{n=0}^{t-1}L_m^n\right)^{-1}\cdot\varepsilon,$$
then the initial state error satisfies
$$r=\|x(0)-\hat{x}(0)\|<\frac{1}{L_m^t}\left(\varepsilon-\sum_{n=0}^{t-1}L_m^n\varepsilon_f\right). \tag{18}$$

We must have $\|x(t)-\hat{x}(t)\|<\varepsilon$, i.e. $\max\|y(t)-\hat{y}(t)\|=\|x_{n_y}(t)-\hat{x}_{n_y}(t)\|<\varepsilon$. This inequality along with the fact that t is an arbitrary value in $[0,M]$ proves theorem 2.

Theorem 2 indicates that if the fuzzy system approximates the actual system well enough, even though there exists initial state error between the fuzzy system and the actual system, this fuzzy system can also work well.

3 Simulation and Analysis

Example 1. We use fuzzy system to identify the following nonlinear system:
$$y(t+1)=\frac{y(t)y(t-1)(y(t)+2.5)}{1+y^2(t)+y^2(t-1)}+u(t).$$
Activation function $u(t)=\sin\frac{2t\pi}{25}$.

Here, we use TSK (Takagi-Sugeno) fuzzy system to identify the nonlinear system. TSK fuzzy system also consists of three parts: fuzzification, fuzzy inference, defuzzification. The obvious feature of this system is the fact that the output variable in the conclusion part of each fuzzy rule is the linear combination of input variables in the condition part of each rule. The rule in TSK fuzzy system can be expressed as:
$$R^i:\text{IF } x_1 \text{ IS } A_1^i \text{ AND } x_2 \text{ IS } A_2^i \text{ AND}\ldots\text{AND } x_n \text{ IS } A_n^i,\ \text{THEN } y^i=p_0^i+p_1^ix_1+\ldots+p_n^ix_n.$$

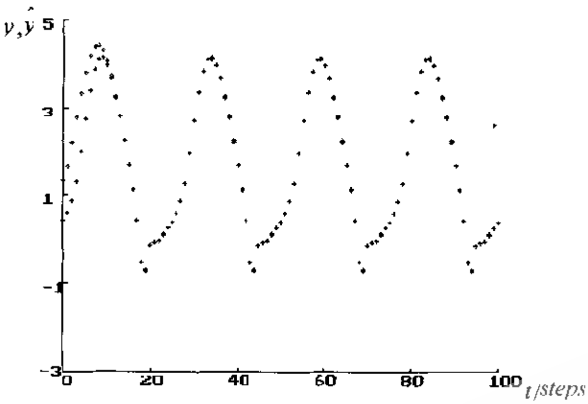

Fig. 2 The hollow dots——the actual system.
Initial state values are $y(-1)=1.5$, $y(0)=0.8$
The solid dots——the fuzzy system.
Initial state values are $\hat{y}(-1)=1.5$, $\hat{y}(0)=0.0$

We divide the fuzzy subsets of $u(t)$, $y(t)$, $y(t-1)$ into {NS, ZO, PB}. NS, ZO, PB represent Negative small, Zero, Positive big, respectively. Gaussian membership function is used, i.e. $\mu_{A_j^i}(x_j)=\exp\left(-\frac{(x_j-a_j^i)^2}{2(b_j^i)^2}\right)$. Learning algorithm is based on BP. After 200 iterations, the mean variance of all samples is 0.000320.

After the TSK fuzzy system is built, it can be independently applied. Figure 2 shows the curves of the actual system and the fuzzy system at different initial states, respectively.

From Fig. 2, we can draw the following conclusion. Though there exists approximation error between the fuzzy system and the actual system, so long as the approximation error is small enough, even if the initial state values of the actual system and the fuzzy system are different (satisfying (2) in theorem 2), the fuzzy system can still approximate the actual system very well. This proves theorem 2 in practice.

4 Conclusion

In this paper, we have discussed the numerical approximation characteristics of fuzzy systems and pointed out that the number of rules of a fuzzy system should not exceed that of samples we can get. We have also investigated the influence of approximation error and initial state error on fuzzy system and get a conclusion that under some conditions, fuzzy system output differs little from the output of the actual system. In the end, through simulation, we prove the correctness of the theorems described in this paper.

References

1 Lin C T. Neural Fuzzy Systems. New York: Prentice-Hall Press, 1997. 481～490

2 Pao Xiaohong *et al*. Model error analysis in nonlinear system identification using neural networks (I). Control and Decision, 1997,12(5):20～25

3 Wang Shi-tong *et al*. Fuzzy neural system and its application. Beijing: Publishing House of Beijing University of Aeronautics and Astronautics, 1996. 200～208
(王士同等. 模糊神经系统及其应用. 北京:北京航空航天大学出版社,1996. 200～208)

4 Freeman J. Learning and Generalization in RBFN. Journal of Neural Computation, 1995,7(3):32～36

5 Niyogi D, Girosi F. On the relationship between generalization error, hypothesis complexity and sample complexity for RBFN. AI Laboratory, MIT, 1994. 103～105

非线性系统的模糊辨识误差分析

王士同　於东军

(华东船舶学院计算机系　镇江　212003)

摘要　该文讨论了模糊系统的数字逼近特性,同时分析了逼近误差和初始状态误差对模糊系统的影响. 最后得出一个重要结论:在一定条件下,模糊系统的输出与实际系统相差不大.

关键词　全局逼近,模糊系统,系统辨识,逼近误差.

中图法分类号　TP18