

基于模糊模型的非线性离散时间系统辨识： 算法与性能分析¹⁾

万峰 孙优贤

(浙江大学工业控制技术国家重点实验室现代控制工程研究所 杭州 310027)
(E-mail: eefwan@ee.ust.hk)

摘要 讨论使用模糊系统方法辨识非线性离散时间系统时, 模糊系统模型的构造、逼近性质以及模型参数的自适应调整算法. 研究了该辨识方案的有关性能, 对模糊模型的参数误差和辨识误差进行了分析, 并给出了模糊模型参数的估计值收敛到其真实值所需的持续激励条件.

关键词 模糊系统模型, 非线性系统辨识, 性能分析, 持续激励条件, 参数收敛性
中图分类号 TP273

Nonlinear Discrete-Time System Identifications Based on Fuzzy Models: Algorithms and Performance Analyses

WAN Feng SUN You-Xian

(Institute of Modern Control Engineering, National Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027)
(E-mail: eefwan@ee.ust.hk)

Abstract This paper investigates the constructions, the approximation properties, and the parameter tuning algorithms of fuzzy system models in identification of nonlinear discrete-time systems with fuzzy system approach. The performance of the identification scheme is analyzed, including the parameter error and prediction error of the fuzzy model. The persistent excitation conditions are established, under which the parameters of the fuzzy system model converge to their true values.

Key words Fuzzy system model, nonlinear system identification, performance analysis, persistent excitation condition, parameter convergence

1 引言

近年来, 有关模糊系统理论与应用的研究取得了令人瞩目的成果^[1~5]. 然而与某些传统领域相比, 模糊系统的设计与分析还缺乏一套系统而有效的方法, 以及完整严密的理论

1) 香港研究资助局 (HKUST 6061/99) 资助

Supported by Hong Kong Research Grants Council under the grant (HKUST 6061/99)

收稿日期 2003-07-15 收修改稿日期 2004-06-29

Received July 15, 2003; in revised form June 29, 2004

作为指导和支持. 对模糊系统的构造与结构特性和逼近精度、学习算法及性能等进行深入分析, 建立一套完善的理论体系是模糊理论研究者的一项重要任务^[6].

模糊系统方法在复杂非线性系统建模中已被广泛应用^[1~5, 7~9]. 与其它方法相比, 模糊系统模型能够较好地利用人类语言知识, 易于表达、理解和实现, 并且可以方便地选择初始参数以加速辨识算法的收敛. 但目前对基于模糊系统模型的非线性系统辨识算法和性能的研究与分析, 对比分段多项式、样条函数和小波级数等仍有较大差距. 另一方面, 在辨识中模糊系统往往被当作黑盒, 其内部特殊的结构和性质未被充分考虑而导致保守的辨识或控制方案以及相应的性能分析.

本文讨论使用模糊系统对非线性离散时间系统进行辨识时, 模糊系统模型的构造、逼近性质和模型参数的自适应调整算法及其有关性能, 着重研究模糊模型的参数误差、辨识误差, 以及模糊系统模型中参数的收敛性.

2 基于模糊模型的非线性系统辨识

考虑由如下非线性自回归移动平均 (N-ARMA) 模型表示的单输入 - 单输出 (SISO) 非线性离散时间系统

$$y(t+1) = g(y(t), \dots, y(t-n+1); u(t), \dots, u(t-m+1)) \quad (1)$$

式中系统输入 $u \in R_u \subset R$, 系统输出 $y \in R_y \subset R$, R_u 及 R_y 设为有界紧集, n 和 m 分别表示输入和输出的阶次. 记

$$\mathbf{x}(t) = [y(t), \dots, y(t-n+1); u(t), \dots, u(t-m+1)]^T = [x_1, \dots, x_s]^T \quad (2)$$

为 $g(\cdot)$ 的输入矢量, 其维数 $s = n + m$, 即 $\mathbf{x} \in X \subset R^{n+m}$. $g(\mathbf{x})$ 为定义在 X 上描述系统动态特性的未知非线性函数. 并假定对 \mathbf{x} 各分量都连续且偏导数有界. N-ARMA 模型比状态空间模型所能表示的范围要窄, 但足以表示大部分的实际系统^[10-11].

2.1 模糊系统模型的构造方法

用以描述未知函数 $g(\mathbf{x})$ 的模糊系统模型为

$$\hat{y}(t+1) = f(y(t), \dots, y(t-n+1); u(t), \dots, u(t-m+1)) \quad (3)$$

式中 $f(\cdot)$ 为下面将要构造的模糊系统, 其输出 $\hat{y}(t+1)$ 也就是对被控对象输出 $y(t+1)$ 的预测. 以下构造一个使用乘积推理机、单值模糊产生器、中心 - 平均模糊消除器和三角形隶属度函数的标准 Mamdani 型模糊系统 $f(\cdot)$.

步骤 1. 记 $X = [\alpha_1, \beta_1] \times \dots \times [\alpha_s, \beta_s]$. 对于每个 $j (j = 1, 2, \dots, s)$, 在 $[\alpha_j, \beta_j]$ 上定义 N_j 个模糊集, 并选用以下的三角形隶属度函数

$$\mu_{A_j^1}(x_j) = \mu(x_j; \gamma_j^1, \gamma_j^1, \gamma_j^2) \quad (4a)$$

$$\mu_{A_j^{N_j}}(x_j) = \mu(x_j; \gamma_j^{N_j-1}, \gamma_j^{N_j}, \gamma_j^{N_j}) \quad (4b)$$

$$\mu_{A_j^r}(x_j) = \mu(x_j; \gamma_j^{r-1}, \gamma_j^r, \gamma_j^{r+1}) \quad (4c)$$

其中 $r = 2, 3, \dots, N_j - 1$, $\alpha_j = \gamma_j^1 < \gamma_j^2 < \dots < \gamma_j^{N_j} = \beta_j$, 且

$$\mu(z; a, b, c) = \begin{cases} (z-a)/(b-a), & \text{if } a \neq b \text{ and } z \in [a, b] \\ (c-z)/(c-b), & \text{if } b \neq c \text{ and } z \in [b, c] \\ 1, & \text{if } a = b \text{ and } z \leq a \\ 1, & \text{if } b = c \text{ and } z \geq c \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

步骤 2. 构造 $N = \prod_{j=1}^s N_j$ 条模糊 IF-THEN 规则如下:

$$\text{Rule }^{i_1 \dots i_s} : \text{IF } x_1 \text{ is } A_1^{i_1} \text{ and } \dots \text{ and } x_s \text{ is } A_s^{i_s}, \text{ THEN } y \text{ is } B^{i_1 \dots i_s} \quad (6)$$

这里 $i_1 = 1, 2, \dots, N_1; \dots; i_s = 1, 2, \dots, N_s$; 模糊集 $B^{i_1 \dots i_s}$ 中心点为可变参数 $\theta^{i_1 \dots i_s}$.

步骤 3. 使用规则 (6), 并选取乘积推理机、单值模糊产生器和中心 - 平均模糊消除器建立的 Mamdani 型模糊系统 $f(\mathbf{x})$. 该模糊系统可表示为

$$f(\mathbf{x}) = \frac{\sum_{i_1=1}^{N_1} \dots \sum_{i_s=1}^{N_s} \theta^{i_1 \dots i_s} [\prod_{j=1}^s \mu_{A_j^{i_j}}(x_j)]}{\sum_{i_1=1}^{N_1} \dots \sum_{i_s=1}^{N_s} [\prod_{j=1}^s \mu_{A_j^{i_j}}(x_j)]} = \sum_{i_1=1}^{N_1} \dots \sum_{i_s=1}^{N_s} \theta^{i_1 \dots i_s} [\prod_{j=1}^s \mu_{A_j^{i_j}}(x_j)] \quad (7)$$

三角形隶属度函数 (4) 和 (5) 的选用使得上面第一个等式右边的分母项恒等于 1, 因此第二个等号成立.

注 1. 隶属度函数的选择很多, 例如高斯 (Gaussian) 型隶属度函数和梯形隶属度函数等. 三角形隶属度函数最大的优点是简单, 不仅计算容易而且概念上也很简洁^[12].

2.2 模糊系统模型的逼近性质

下面的引理揭示了上节所述的 Mamdani 型模糊系统逼近未知非线性函数的能力.

引理 1^[1,2,8]. 对于任何给定的紧集 X 上的连续函数 $g(\mathbf{x})$ 以及任意常数 $\varepsilon > 0$, 存在一个式 (7) 所表示的标准 Mamdani 型模糊系统, 使得不等式

$$\| f(\mathbf{x}) - g(\mathbf{x}) \|_{\infty} < \varepsilon \quad (9)$$

成立, 其中 X 上的无穷范数 $\| * \|_{\infty}$ 定义为 $\| p(\mathbf{x}) \|_{\infty} = \sup_{\mathbf{x} \in X} |p(\mathbf{x})|$.

引理 2^[13,2]. 记 $h_j = \max_{1 \leq r \leq N_j-1} |\gamma_j^{r+1} - \gamma_j^r|$.

1) 如果 $g(\mathbf{x})$ 在 X 上连续可导, 那么

$$\| f(\mathbf{x}) - g(\mathbf{x}) \|_{\infty} \leq \sum_{j=1}^s \left\| \frac{\partial g(\mathbf{x})}{\partial x_j} \right\|_{\infty} h_j \quad (10)$$

2) 如果 $g(\mathbf{x})$ 在 X 上二次连续可导, 那么

$$\| f(\mathbf{x}) - g(\mathbf{x}) \|_{\infty} \leq \frac{1}{8} \sum_{j=1}^s \left\| \frac{\partial^2 g(\mathbf{x})}{\partial x_j^2} \right\|_{\infty} h_j^2 \quad (11)$$

注 2. 引理 1 是万能逼近定理的一个特例, 它为前面所构造的模糊系统用于建模和控制提供了理论上的保证, 即只要有足够多的模糊规则, 一致逼近的精度就可以足够高. 因此, 模糊辨识模型 (3) 是式 (1) 所示的被控对象的一个合理表示. 引理 2 则给出了 $g(\mathbf{x})$ 导数的上界已知时, 上述模糊系统较为量化的逼近性能. 另外, 因 h_j 往往较小, 故不等式 (11) 右边的误差二阶界要比不等式 (10) 右边的误差一阶界小. 由于 $g(\mathbf{x})$ 通常为未知函数, $\left\| \frac{\partial g(\mathbf{x})}{\partial x_j} \right\|_{\infty}$ 和 $\left\| \frac{\partial^2 g(\mathbf{x})}{\partial x_j^2} \right\|$ 一般也是未知的, 引理 2 实际运用上较为困难, 但有一定的理论指导意义.

2.3 模糊系统模型的参数调整算法

将模糊系统 $f(\mathbf{x})$ 改写为下面对参数线性的形式:

$$f(\mathbf{x}) = \phi^T(\mathbf{x})\boldsymbol{\theta} \quad (12)$$

其中 $\boldsymbol{\theta}$ 是 N 个参数 $\theta^{i_1 \dots i_s}$ 合在一起的矢量形式, 即

$$\boldsymbol{\theta} = [\theta^{1 \dots 1}, \dots, \theta^{1 \dots 1 \dots N_s}, \dots, \theta^{N_1 \dots N_2 \dots 1}, \dots, \theta^{N_1 \dots N_2 \dots N_s}]^T \quad (13)$$

$\phi(\mathbf{x})$ 为对应的模糊基函数

$$\phi(\mathbf{x}) = [\phi^{1 \dots 1}(\mathbf{x}), \dots, \phi^{1 \dots 1 \dots N_s}(\mathbf{x}), \dots, \phi^{N_1 \dots N_2 \dots 1}(\mathbf{x}), \dots, \phi^{N_1 \dots N_2 \dots N_s}(\mathbf{x})]^T \quad (14)$$

这里 $\phi^{i_1 \dots i_s}(\mathbf{x}) = \prod_{j=1}^s \mu_{A_j^{i_j}}(x_j)$.

由引理 1 知, 使用足够多的模糊集, 模糊系统能够以任意精度逼近非线性函数 $g(\cdot)$, 而当模糊集不够时上述逼近可能产生误差. 为此, 定义以下最佳参数 $\boldsymbol{\theta}^0$ (参数真实值)

$$\boldsymbol{\theta}^0 = \arg \min_{\boldsymbol{\theta}} \{ \sup_{\boldsymbol{\theta}} |f(\mathbf{x}|\boldsymbol{\theta}) - g(\mathbf{x})| \} \quad (15)$$

它表示 $f(\mathbf{x}|\boldsymbol{\theta}^0)$ 是对 $g(\mathbf{x})$ 逼近得最好的模糊系统. 定义逼近误差 $w(\mathbf{x})$ 为

$$w(\mathbf{x}) = f(\mathbf{x}|\boldsymbol{\theta}^0) - g(\mathbf{x}) \quad (16)$$

那么

$$y(t+1) = \phi^T(\mathbf{x}(t))\boldsymbol{\theta}^0 + w(\mathbf{x}(t)) \quad (17)$$

令 $\boldsymbol{\theta}(t)$ 为 t 时刻 $\boldsymbol{\theta}^0$ 的估计. 以下讨论几种常用的参数辨识算法.

算法 1. 带死区的投影算法

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1) + \frac{I(t)\phi(\mathbf{x}(t-1))}{\delta - \phi^T(\mathbf{x}(t-1))\phi(\mathbf{x}(t-1))} e(t) \quad (18)$$

式中 $\mathbf{x}(t)$, $\boldsymbol{\theta}(t)$ 和 $\phi(\mathbf{x}(t))$ 由式 (2), (13) 和 (14) 定义, δ 为一小正数, $\boldsymbol{\theta}(0)$ 为给定的初始值. 模型预报误差 $e(t)$ 定义为

$$e(t) = y(t) - \phi^T(\mathbf{x}(t-1))\boldsymbol{\theta}(t-1) \quad (19)$$

指示函数 $I(t)$ 为

$$I(t) = \begin{cases} 1, & \text{if } |e(t)| > 2\Delta \\ 0, & \text{otherwise} \end{cases} \quad (20)$$

其中常数 Δ 决定死区大小.

带死区的投影算法简单, 因而易于计算. 当存在测量误差、建模误差或计算机舍入误差时, 这种算法有较好的鲁棒性, 因为在预测误差小于上述扰动噪声信号时, 调整算法会被自动关掉. 对基于模糊系统模型的辨识而言, 如果在构造模糊系统时没有定义足够的模糊集, 那么不论如何调节模糊系统的参数, 辨识误差都不可能减至很小. 在这种情况下, 根据引理 2, 可将死区设定为

$$\Delta \geq \sum_{j=1}^s \left\| \frac{\partial g(\mathbf{x})}{\partial x_j} \right\|_{\infty} h_j \quad (21)$$

从而在当预测误差小于建模误差时关掉参数辨识算法. 投影算法的缺点是收敛速度通常较慢. 相比而言, 下面的正交投影算法和最小二乘算法收敛速度则快得多.

算法 2. 正交投影算法

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1) + \frac{P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))}{\boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))} e(t) \quad (22)$$

$$P(t) = P(t-1) - \frac{P(t-1)\boldsymbol{\phi}(\mathbf{x}(t-1))\boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-1)}{\boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-1)\boldsymbol{\phi}(\mathbf{x}(t-1))} \quad (23)$$

其中初始值 $\boldsymbol{\theta}(0)$ 给定, $P(0) = I$, $e(t)$ 定义同式 (19). 对于式 (22) 中分母项 $\boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))$ 可能出现为零的情况, 可在分母项引入 $\delta > 0$, 使得

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1) + \frac{P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))}{\delta + \boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))} e(t) \quad (24)$$

算法 3. 最小二乘算法

式 (24) 中当 $\delta = 1$ 时, 即为最小二乘算法:

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1) + \frac{I(t)P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))}{1 + \boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-2)\boldsymbol{\phi}(\mathbf{x}(t-1))} e(t) \quad (25)$$

$$P(t) = P(t-1) - \frac{P(t-1)\boldsymbol{\phi}(\mathbf{x}(t-1))\boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-1)}{1 + \boldsymbol{\phi}^T(\mathbf{x}(t-1))P(t-1)\boldsymbol{\phi}(\mathbf{x}(t-1))} \quad (26)$$

其中初始值 $\boldsymbol{\theta}(0)$ 给定, $P(0)$ 为一正定矩阵, $e(t)$ 定义同式 (19). 正交投影算法收敛最快, 但对噪声相当敏感. 相反, 最小二乘算法克服噪声影响能力最强. 对于这两种参数辨识算法, 也可采用如算法 1 的死区开关技术以增强其鲁棒性.

3 性能分析

3.1 模糊模型辨识误差分析

考虑模糊系统模型 (7) 带参数 $\boldsymbol{\theta}$ 的表示式 $f(\mathbf{x}|\boldsymbol{\theta})$. $g(\mathbf{x})$ 表示未知的被控对象 (1). 由引理 2 可知, 逼近误差 $w(\mathbf{x})$ 满足下面的不等式

$$|w(\mathbf{x})| \leq \sup |f(\mathbf{x}|\boldsymbol{\theta}^0) - g(\mathbf{x})| \leq \sum_{j=1}^s \left\| \frac{\partial g}{\partial x_j} \right\|_{\infty} h_j \quad (27)$$

式中最右边项的大小可在模糊系统设计的过程中加以控制. 模型辨识误差为 $e(t+1) = y(t+1) - \hat{y}(t+1)$, 则有

$$\begin{aligned} e(t+1) = y(t+1) - \hat{y}(t+1) &= g(\mathbf{x}(t)) - f(\mathbf{x}(t)|\boldsymbol{\theta}(t)) = \\ &= (f(\mathbf{x}(t)|\boldsymbol{\theta}^0) + w(t)) - f(\mathbf{x}(t)|\boldsymbol{\theta}(t)) = \boldsymbol{\phi}^T(\mathbf{x}(t))(\boldsymbol{\theta}^0 - \boldsymbol{\theta}(t)) + w(t) \end{aligned} \quad (28)$$

建模的最终目标是使模型的输出 $\hat{y}(t+1)$ 与对象的输出 $y(t+1)$ 尽可能地接近. 由上式可知, $e(t+1)$ 包括两部份: 1) 参数估计 $\boldsymbol{\theta}(t)$ 与最佳参数 $\boldsymbol{\theta}^0$ 的差, 称为参数误差, 在 $e(t+1)$ 中的影响可用 $\boldsymbol{\phi}^T(\mathbf{x}(t))(\boldsymbol{\theta}^0 - \boldsymbol{\theta}(t))$ 表示; 2) 最佳模糊模型 $f(\mathbf{x}(t)|\boldsymbol{\theta}^0)$ 与未知对象 $g(\mathbf{x})$ 的差即逼近误差 $w(\mathbf{x})$, 也称为结构误差. 2.3 节的参数调整算法和 2.1 节的模糊系统设计的目标, 就是要分别将参数误差和结构误差减到最小.

以带死区的投影算法为例, 下面考虑基于模糊模型的非线性系统辨识方案的性能.

定理 1. 对于被控对象 (1), 如果使用模糊系统模型 (3) 以及 2.3 节的带死区的投影算法 (18)~(20), 并选择死区 $\Delta = \sum_{j=1}^s \left\| \frac{\partial g(\mathbf{x})}{\partial x_j} \right\| h_j$, 则

1) 参数误差 $\tilde{\boldsymbol{\theta}}(t) = \boldsymbol{\theta}(t) - \boldsymbol{\theta}^0$ 满足

$$|\tilde{\boldsymbol{\theta}}(t)| \leq |\tilde{\boldsymbol{\theta}}(t-1)| \leq |\tilde{\boldsymbol{\theta}}(0)|, \quad t \geq 1 \quad (29)$$

2) 辨识误差 $e(t+1) = y(t+1) - \hat{y}(t+1)$ 满足

$$\limsup_{t \rightarrow \infty} |e(t+1)| \leq 2\Delta \quad (30)$$

证明. 由式 (27) 及 $\Delta = \sum_{j=1}^s \left\| \frac{\partial g(\mathbf{x})}{\partial x_j} \right\| h_j$ 知

$$|w(\mathbf{x})| \leq \sup |g(\mathbf{x}) - f(\mathbf{x}|\boldsymbol{\theta}^0)| \leq \sum_{j=1}^s \left\| \frac{\partial g}{\partial x_j} \right\|_{\infty} h_j = \Delta \quad (31)$$

下面的证明参照文献 [14] 的思路, 易证

1) $\|\tilde{\boldsymbol{\theta}}(t)\|^2$ 是一个非增数列且下界为零, 故式 (29) 成立;

$$2) \lim_{t \rightarrow \infty} \frac{I(t-1)[e^2(t) - 4\Delta^2]}{\delta + \boldsymbol{\phi}^T(\mathbf{x}(t-1))\boldsymbol{\phi}(\mathbf{x}(t-1))} = 0 \quad (32)$$

从三角形隶属度函数的定义 (4), (5) 和回归向量的表达式 (14) 可见, 在域 X 上 $\boldsymbol{\phi}(\mathbf{x}(t-1))$ 的 Euclidean 长度总是有限的, 因而上式的分母 $\delta + \boldsymbol{\phi}^T(\mathbf{x}(t-1))\boldsymbol{\phi}(\mathbf{x}(t-1))$ 有界, 故式 (30) 成立. 证毕.

注 3. 定理 1 表明辨识方案可确保参数误差总是非增的, 且预测误差可控制在 2Δ 的范围内. 因为通过设计足够多的模糊集可使引理 2 中的 h_j 任意地小, 所以参数误差也可以控制得很小, 但这并不保证参数的估计 $\boldsymbol{\theta}(t)$ 收敛到最优值 $\boldsymbol{\theta}^0$.

注 4. 容易验证, 带死区的正交投影算法和最小二乘算法也有类似的性能.

3.2 模糊模型参数收敛性分析

本节讨论使用模糊模型 (3) 和正交投影算法 (22) 和 (23) 来辨识非线性系统 (1) 时, 模糊模型中参数的收敛性.

下面的引理给出了模糊模型参数的估计值收敛到其真实值所需的持续激励条件.

引理 3^[14]. 对于某个 $k \geq \prod_{j=1}^s N_j$, 如果

$$\text{rank}[\phi(x(1)), \dots, \phi(x(k))] = \prod_{j=1}^s N_j = \text{dimension of } \theta^0 \tag{33}$$

那么由正交投影算法 (22) 和 (23) 求出的参数估计 $\theta(t)$ 在 k 步内收敛到真实值, 即 $\theta(k) = \theta^0$. 式 (33) 称为正交投影算法的持续激励条件. 为叙述简明起见, 下面以输入向量 $x(t)$ 为二维为例, 给出上述辨识问题的持续激励条件, 但思路可推广到更一般的情况.

定理 2. 考虑使用图 1 所示模糊集的模糊模型. 称 $(\gamma_1^i, \gamma_2^j) (i = 1, \dots, N_1, j = 1, \dots, N_2)$ 为回归向量 $\phi(x)$ 的中心点. 如果对于某个 $k \geq N_1 N_2, x(1), \dots, x(k)$ 包含这 $N_1 N_2$ 个中心点, 那么持续激励条件 (33) 成立, 因而模糊模型参数的估计收敛到其真实值.

证明. 由图 1 及回归向量的定义式 (14) 可见, 如果 $x(t) = (\gamma_1^i, \gamma_2^j) (i \text{ 和 } j \text{ 给定})$, 那么 $\phi(x)$ 中只有第 $j + (i - 1)N_2$ 个分量的值为 1 而其它分量皆为 0, 即

$$\phi(x(t)) = \phi(\gamma_1^i, \gamma_2^j) = (0 \dots 0 1 0 \dots 0)^T \tag{34}$$

其中 1 在第 $j + (i - 1)N_2$ 个位置处. 如果对于某个 $k \geq N_1 N_2$, 输入信号 $x(1), \dots, x(k)$ 包含这 $N_1 N_2$ 个中心点 $(\gamma_1^i, \gamma_2^j) (i = 1, \dots, N_1, j = 1, \dots, N_2)$, 那么 $N_1 N_2 \times k$ 矩阵包含有一个 $N_1 N_2 \times N_1 N_2$ 的单位矩阵, 因此持续激励条件 (33) 成立. 证毕.

二维情况下模糊系统的中心点如图 1 所示.

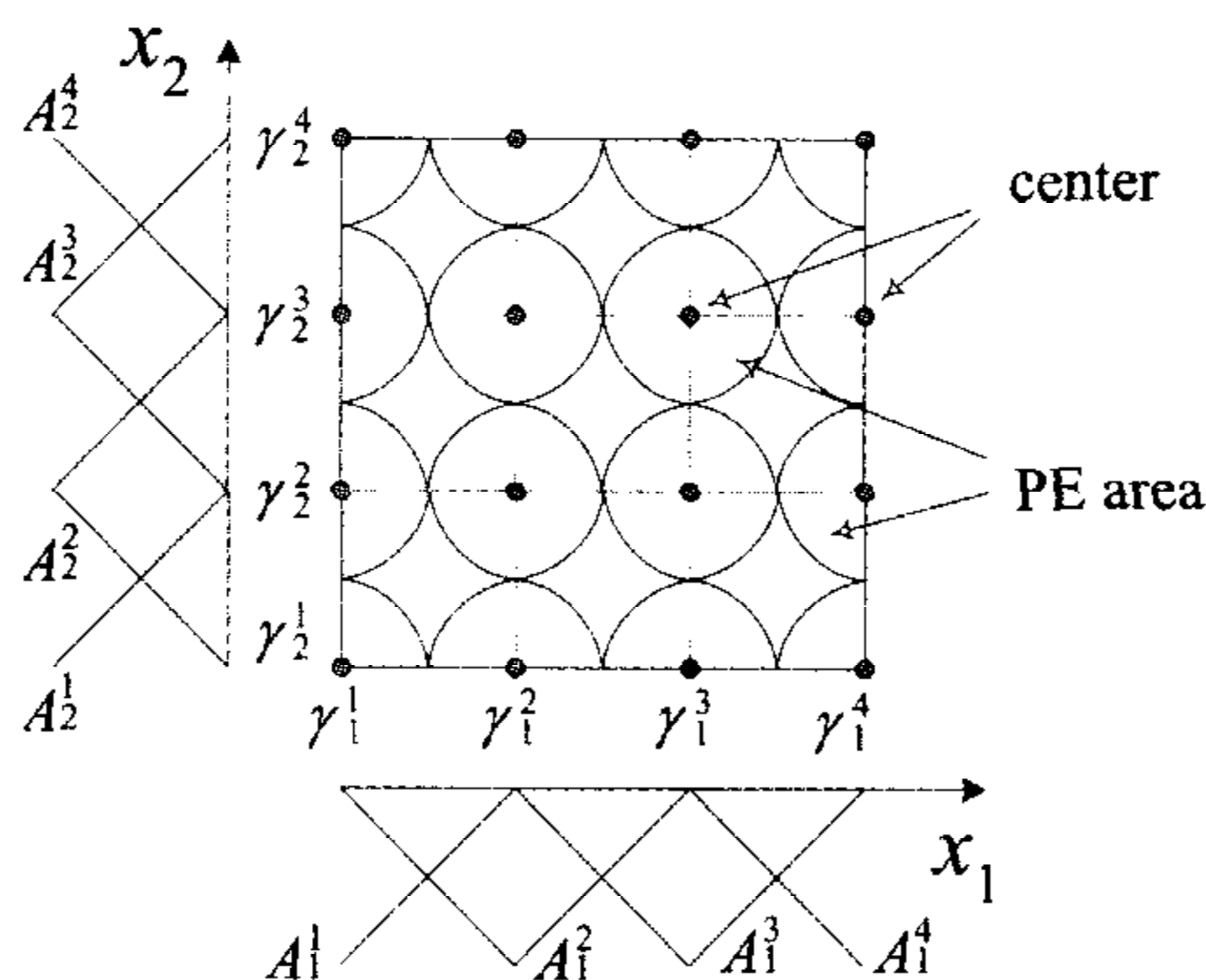


图 1 模糊模型回归向量的中心点和持续激励区域的示意图

Fig. 1 Centers and PE areas of a two-dimensional fuzzy system model

定理 2 要求模糊系统的输入向量 $x(t)$ 经过所有中心点. 从实际应用的角度来说, 这并不现实. 下面的定理表明, 如果模糊模型所有的中心点附近的某个区域内至少有一个输入向量, 那么持续激励条件就可满足. 从而得到更易实现的持续激励条件.

引理 4^[15]. 方阵 $A \in R^{n \times n}$ 为非奇异阵, 如果它是对角优势的, 即

$$|a_{jj}| > \sum_{i=1, i \neq j}^n |a_{ij}|, \text{ for } j = 1, 2, \dots, n \tag{35}$$

其中 a_{ij} 表示 A 位置 (i, j) 上的元素.

运用引理 4 可将定理 2 中的条件放松, 如下面的定理所示.

定理 3. 考虑使用图 1 所示模糊集的模糊模型. 定义以 (γ_1^i, γ_2^j) 为中心的持续激励区域为

$$PE_{(\gamma_1^i, \gamma_2^j)} = \{(x_1, x_2) | \mu_{A_1^i}(x_1)\mu_{A_2^j}(x_2) > 0.5\} \quad (36)$$

如果对于某个 $k \geq N_1 N_2$, 在以模糊模型的中心点 (γ_1^i, γ_2^j) 为中心的所有的持续激励区域 $PE_{(\gamma_1^i, \gamma_2^j)}$ 内都能找到 $x(1), \dots, x(k)$ 中的至少一个输入向量, 那么 $\theta(k) = \theta^0$.

证明. 对于 $\mathbf{x}(t) = (x_1, x_2)$, 回归向量 $\phi(\mathbf{x})$ 中最多只有 4 个分量的值不为零. 它们是 $\mu_{A_1^i}(x_1)\mu_{A_2^j}(x_2)$, $\mu_{A_1^i}(x_1)(1-\mu_{A_2^j}(x_2))$, $(1-\mu_{A_1^i}(x_1))\mu_{A_2^j}(x_2)$ 和 $(1-\mu_{A_1^i}(x_1))(1-\mu_{A_2^j}(x_2))$, 其中 $\mu_{A_1^i}(x_1)\mu_{A_2^j}(x_2)$ 落在 $\phi(\mathbf{x})$ 中 $j+(i-1)N_2$ 处. 由于这 4 个分量之和为 1 且 $\mu_{A_1^i}(x_1)\mu_{A_2^j}(x_2) > 0.5$, 故 $\mu_{A_1^i}(x_1)\mu_{A_2^j}(x_2)$ 大于其它 3 个分量的和. 因为每一持续激励区域 $PE_{(\gamma_1^i, \gamma_2^j)}$ (对于所有的 i 和 j) 都含有至少一个输入向量, 所以矩阵 $[\phi(x(1)), \dots, \phi(x(k))]$ 包含有一个对角优势 $N_1 N_2 \times N_1 N_2$ 的单位矩阵. 因此根据引理 3 和引理 4, 持续激励条件 (33) 成立, $\theta(k) = \theta^0$. 证毕.

有关模糊模型参数收敛性较为详细的分析以及持续激励输入信号的设计参见文献 [16,17].

4 仿真研究

本节通过使用模糊模型对离散时间非线性系统进行辨识的一个例子, 研究模型辨识精度及参数收敛性.

例. 辨识以二阶非线性移动平均 (N-MA) 模型表示的非线性系统

$$y(t+1) = \frac{u(t)}{1 + u^2(t) + u^2(t-1)} \quad (37)$$

模糊模型为 $\hat{y}(t+1) = f(u(t), u(t-1))$, 其中 $\hat{y}(t+1)$ 为对系统输出 $y(t+1)$ 的预测, $f(\cdot)$ 为一个由乘积推理机、单值模糊产生器、中心-平均模糊消除器和三角形隶属度函数 (4) 与 (5) 构成的 Mamdani 型模糊系统. 定义 5×5 个模糊集均匀地覆盖模糊系统输入的定义域 $[-1, 1] \times [-1, 1]$. 设计输入信号 u 为 $\{-1, -1, -0.5, -1, 0, -1, 0.5, -1, 1, -1, 1, -0.5, 0.5, -0.5, 0, 0.5, -0.5, -0.5, -1, -0.5, -1, 0, 0.5, 0, 0, 0, 0.5, 0, 1, 0, 1, 0.5, 0.5, -0.5, 0, 0.5, -0.5, 0.5, -1, 0.5, -1, 1, -0.5, 1, 0, 1, 0.5, 1, 1, 1\}$.

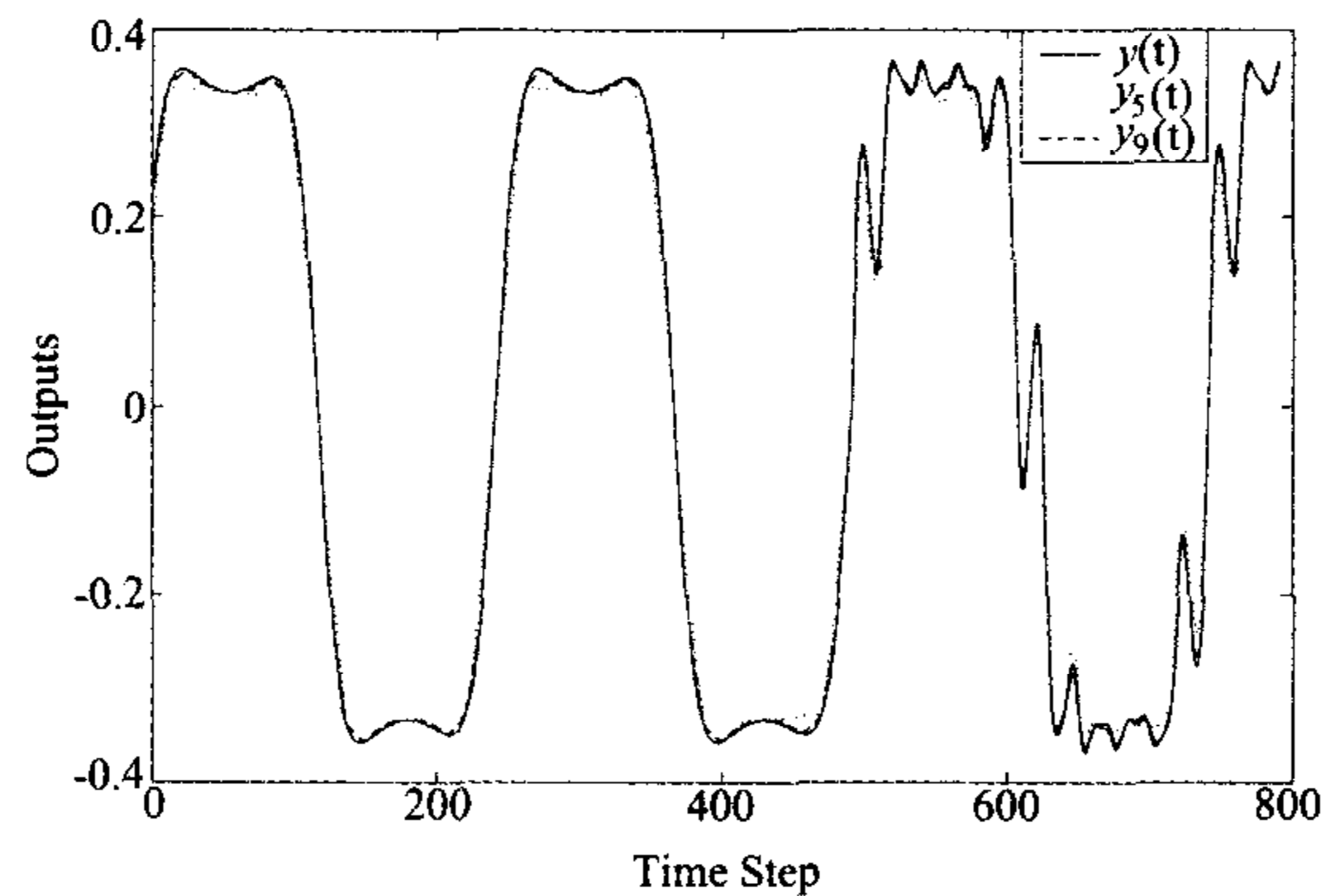
根据定理 2, 可以验证上述输入信号满足持续激励条件, 故辨识得到的模糊模型参数的估计值将收敛到其真实值.

利用正交投影算法 (22), (23) 和上述输入信号以及相应的系统输出, 对模糊模型的参数进行辨识. 辨识结束后可以看到, 模糊模型的参数 θ 为 $[-0.3333, -0.4444, -0.5, -0.4444, -0.3333, -0.2222, -0.3333, -0.4, -0.3333, -0.2222, 0, 0, 0, 0, 0.2222, 0.3333, 0.4, 0.3333, 0.2222, 0.3333, 0.4444, 0.5, 0.4444, 0.3333]^T$.

对于输入信号

$$u(t) = \begin{cases} \sin(2\pi t/250), & 1 \leq t < 500 \\ 0.8 \sin(2\pi t/250) + 0.2 \sin(2\pi t/25), & 500 \leq t \leq 800 \end{cases} \quad (38)$$

图 2 给出了辨识得到的模糊模型和原系统的响应. 模糊模型预测的平均绝对误差为 0.017. 由图可见, 尽管只有 25 个输入输出数据对, 辨识得到的模糊模型已经能够很好地表示原来的系统. 容易检验, 辨识后模糊模型的参数非常接近其真实值. 另外, 由引理 2 及定理 1 可知, 使用更多的模糊集可提高模糊模型的逼近精度. 作为对比, 图 2 也给出了使用 9×9 个模糊集时模糊模型的输出. 此时模糊模型预测的平均绝对误差为 0.004.



系统真实输出 ($y(t)$) (Outputs of the plant $y(t)$), 模型预测输出 $y_5(t)$: 5x5 个模糊集 (fuzzy models ($y_5(t)$: 5x5 fuzzy sets), $y_9(t)$: 9x9 个模糊集 ($y_9(t)$: 9x9 fuzzy sets)

图 2 仿真结果

Fig. 2 Simulation results

5 结束语

本文讨论了基于模糊模型的离散时间非线性系统辨识方法. 研究了模糊系统的构造、模糊模型的逼近性质以及模糊模型中参数的自适应调整算法, 分析了辨识方案中模糊模型的参数误差和辨识误差, 并得出了模糊模型参数的估计值收敛到其真实值所需的持续激励条件.

应当指出, 本文的工作只是文献 [6] 提出的未来研究方向 2 和 3 上的一个引子. 对于不同类型模糊系统的结构特性及逼近精度的分析, 适用于模糊系统的不同学习算法的研究, 算法的收敛性及学习结束后模糊系统性能的分析, 都需要更进一步的努力, 以通过严格的理论分析与证明, 最终建立一套完善的理论体系.

致谢 第一作者感谢导师香港科技大学王立新博士的指导.

References

- 1 Wang L X. Adaptive Fuzzy Systems and Control: Design and Analysis. Englewood Cliffs, N.J.: Prentice Hall, 1994
- 2 Wang L X. A Course in Fuzzy Systems and Control. Englewood Cliffs, N.J.: Prentice Hall, 1997
- 3 Yen J, Langari R. Fuzzy Logic: Intelligence, Control and Applications. Englewood Cliffs, N.J.: Prentice Hall, 1999
- 4 Passino K M, Yurkovich S. Fuzzy Control. California: Addison-Wesley, 1998
- 5 Ying H. Fuzzy Control and Modeling: Analytical Foundations and Applications. New York: IEEE Press, 2000

- 6 Wang L X. Fuzzy systems: Challenges and chance—My experiences and perspectives. *Acta Automatica Sinica*, 2001, **27**(4): 585~590 (in Chinese)
- 7 Wang L X. Analysis and design of fuzzy identifiers of nonlinear dynamic systems. *IEEE Transactions on Automatic Control*, 1995, **40**(1): 11~23
- 8 Wang L X, Mendel J M. Fuzzy basis functions, universal approximation, and orthogonal least squares learning. *IEEE Transactions on Neural Networks*, 1992, **3**(5): 807~814
- 9 Hellendoorn H, Driankov D (Eds.). *Fuzzy Model Identification: Selected Approaches*. Berlin: Springer-Verlag, 1997
- 10 Chen S, Billings S A. Representations of nonlinear systems: The NARMAX model. *International Journal of Control*, 1989, **49**(3): 1013~1032
- 11 Sales R, Billings S A. Self-tuning control of nonlinear ARMAX models. *International Journal of Control*, 1990, **51**(4): 753~769
- 12 Pedrycz W. Why triangular membership functions? *Fuzzy Sets and Systems*, 1994, **64**(1): 21~30
- 13 Zeng X J, Singh M G. Approximation accuracy analysis of fuzzy systems as function approximators. *IEEE Transactions on Fuzzy Systems*, 1996, **4**(1): 44~63
- 14 Goodwin G C, Sin K S. *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, N.J.: Prentice-Hall, 1984
- 15 Horn R A, Johnson C R. *Matrix Analysis*. Cambridge: Cambridge University Press, 1990
- 16 Wan F, Wang L X, Sun Y X. On the persistent excitation conditions for adaptive fuzzy systems in nonlinear identifications. In: *Proceedings of 39th IEEE Conference on Decision and Control*, Sydney, Australia. Piscataway, N.J.: IEEE Press, 2000. 858~863
- 17 Wan F, Wang L X, Sun Y X, Zhu H Y. Generating persistently exciting inputs for nonlinear dynamic system identification using fuzzy models. In: *Proceedings of 10th IEEE Conference on Fuzzy Systems*, Melbourne, Australia. Piscataway, N.J.: IEEE Press, 2001. 505~508

万 峰 浙江大学与香港科技大学联合培养博士。目前研究兴趣为模糊系统与神经网络、非线性控制、自适应控制及过程控制等。

(WAN Feng Received the Ph.D. degree from Hong Kong University of Science and Technology. His research interests include fuzzy systems and neural networks, nonlinear control, adaptive control and process control.)

孙优贤 教授、博士生导师、中国工程院院士。长期从事过程控制理论与应用、容错控制理论与应用、鲁棒控制理论与应用以及造纸过程的模型化和计算机控制等方面的研究。

(SUN You-Xian Professor and academician of Chinese Academy of Engineering. His research interests include process control, fault tolerant control, robust control and papermaking process modeling and control.)