

复杂系统模糊建模的模糊树方法*

毛剑琴¹ 岳玉芳¹ 张建刚¹ 代冀阳¹ 李幼平²

(1. 北京航空航天大学第七研究室·北京, 100083; 2. 中国工程物理研究院·成都, 610003)

摘要: 对含高度非线性的复杂系统的模糊建模提出了一种二叉线性模糊树方法. 证明了对 n 维空间中任一闭集上有限样本集或连续函数, 总存在模糊树模型以任一精度逼近之. 模糊树模型在一定程度上模拟了对复杂问题进行分层、分段简化决策的思维过程, 故对其进一步的研究具有科学与应用价值.

关键词: 模糊建模; 二叉树; 隶属函数

文献标识码: A

Fuzzy-Tree Method for Complex System Fuzzy Modeling

MAO Jianqin¹, YUE Yufang¹, ZHANG Jiangan¹, DAI Jiyang¹ and LI Youping²

(1. The Seventh Division, Beijing University of Aeronautics & Astronautics · Beijing, 100083, P. R. China;

2. Academy of Engineering Physics · Chengdu, 610003, P. R. China)

Abstract: A binary linear fuzzy tree structure approach, i. e., FT model, is proposed for complex nonlinear system fuzzy modeling. It is proved that for any continuous function of $L^2(\Omega^n)$, there always exists an FT model to approximate it arbitrarily well. In some sense, FT model simulates the layered decision-making and piece-wise linearized processing procedure for solving complex problems.

Key words: fuzzy modeling; binary tree; membership function

1 引言(Introduction)

对难以找到一个全局的函数建模的复杂系统, 通常采用的一类方法是将输入空间划分后, 在每个子空间内求其相对简单的局部映射关系. 文献[1]中基于网格法的 Takagi-Sugeuo (T-S) 模型, 是一种有效的方法. 但当输入空间维数增加时, 非线性优化所需的时间长. Jyh-Shing Roger Jang 基于模糊神经网络的 ANFIS 模型^[2]明显地改进该模型的精度和计算效率, 但仍难以处理由于高维输入而引起的规划条数(网络数目)“爆炸”问题. Gelfand 提出了一种二叉树结构的分块线性模型^[3], 其优点是对输入空间的维数不敏感, 但在相邻子空间分界处映射函数会出现不连续或不光滑. 本文采用二叉模糊树结构对输入空间进行逐次划分; 在各子空间中采用多个线性模型的非线性加权逼近; 从而在整个输入空间上形成了一个连续、光滑的非线性映射函数.

2 闭集的模糊树结构划分 (Structural partition for closed sets of a fuzzy tree)

定义 1^[3] 二叉树模型由正整数构成的有限非空集 T 和两个函数序列 $l(\cdot)$ 和 $r(\cdot)$ 构成, 且满足:

- 1) $\forall t \in T, l(t) = r(t) = 0$ 或 $l(t) > t, r(t) > t$ 成立;
- 2) $\forall t \in T$, 除了 T 中最小的整数, 必存在唯一 $s \in T$, 使得 $t = l(s)$ 或 $t = r(s)$;

T 中的最小元素称为根节点, 记作 $\text{root}(T)$. 对于节点 $t \in T, l(t)$ 和 $r(t)$ 分别表示由 t 生成的左右节点. 若节点 t 不是父节点, 即 $l(t) = r(t) = 0$, 则称作终节点. $\bar{T} \subset T$ 表示 T 的终节点集合.

引理 1 对 n 维空间中任一给定闭集 χ 二叉树模型的全部终节点形成对给定闭集 χ 的划分. 即每一终节点对应的元素 $x \in \chi$ 属于同一个等价类.

证 将给定闭集 χ 赋予二叉树的根节点 $\text{root}(T)$. 对 χ 进行第一次划分后产生两个子集 χ_1 和 χ_2 , 它们分别对应 $l(\text{root}(T)) \in T$ 和 $r(\text{root}(T)) \in T$. χ 中的任一元素必属于且仅属于这两个子集之一. 即每一个子集中的元素属于同一个等价类. 继而对这两个子集分别进一步划分, 直至全部的终节点. 全部终节点形成集合 $\bar{T}$. 根据有限覆盖定理, 对于给定的闭集, $\bar{T} \subset T$ 是存在的, 并且 χ 中的元素必属于且仅属于某个 $\bar{t} \in \bar{T}$ 的子集. 即全部的终节点形成对 χ 的一个划分.

定义 2 模糊二叉树模型由二叉树模型和隶属度函数序列 $\{\mu_j(x)\}, j \in T$ 构成.

引理 2 对于 n 维空间中任一给定闭集 χ , 模糊二叉树模型的全部终节点形成对给定闭集 χ 的一个模糊划分. 即每一终节点对应的元素属于同一个模糊等价类.

* 基金项目: 北京市自然科学基金(4992007)和博士点基金(2000000625)资助项目
收稿日期: 1998-10-28; 收修改稿日期: 2001-05-09.

证 类似引理 1.

图 1 示意地描述了模糊二叉树对给定二维闭集的划分. 具体算法见第 3 节. 已有的网格法等多数是沿输入的各维方向进行划分而一次形成的划分, 而模糊树模型不论维数多少, 每次只将一个子空间分成两个更小的子空间, 因此逐次形成一个划分; 进而, 这种划分还可自适应地在对应函数关系复杂的输入空间作细划分, 而在其它部分作粗划分, 从而为提高精度和计算效率开辟了新途径.

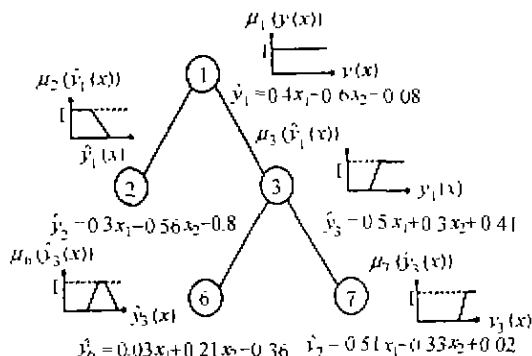


图 1(a) 模糊二叉树示意图

Fig. 1(a) Sketch of fuzzy binary tree

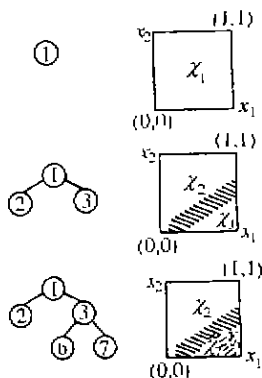


图 1(b) 模糊二叉树对二维闭集的划分

Fig. 1(b) Partition for two dimensional closed sets of a fuzzy tree

3 模糊二叉树(Fuzzy binary tree)

由于多输出问题可化为多个单输出问题, 这里只考虑一维输出问题. 令样本集 $\Omega = \{(x^i, y^i), i = 1, \dots, L\}$, $x^i \in X \subset \mathbb{R}^n$, $y^i \in E \subset \mathbb{R}$. X 称为输入空间, 是 n 维实空间的一个闭集. E 称为输出空间, $|y^i| < \infty$.

令 $x^i = (1, x_1^i, \dots, x_n^i)^T \in X_1 \subset \mathbb{R}^{n+1}$ 为输入向量, 对应每个节点 $t \in T$, 有最小二乘系数构成的向量 $c_t = [c_t(0), \dots, c_t(n)]^T$, 隶属函数 $\mu_t(x)$, 输出的估计 $\hat{y}_t^i = c_t^T x^i$ 及辅助的隶属函数 $\hat{\mu}_t(x)$. 同时每个节点对应输入空间的一个模糊子空间, 记为 X_t . 子节点对应的模糊子空间在父节点所对应的模糊子空间内划分. 辅助隶属函数用来描述子节点如何对父节点所对应的模糊子空间进行划分. 故除根节点外, 有递推关系:

$$\mu_t(x) = \mu_{p(t)}(x) \hat{\mu}_t(x). \quad (1)$$

其中 $\mu_t(x) \in [0, 1]$, $\hat{\mu}_t(x) \in [0, 1]$.

模糊树模型的输出可由所有终节点的输出的反模糊化

得到

$$\hat{y} = \sum_{t \in T} \mu_t(x) \hat{y}_t / \sum_{t \in T} \mu_t(x). \quad (2)$$

t 的子节点 $l(t)$, $r(t)$ 的辅助隶属函数为

$$\hat{\mu}_{l(t)}(x) = \frac{1}{1 + \exp[-\alpha_{l(t)}(c_{l(t)}^T x - \theta_{l(t)})]}, \quad (3a)$$

$$\hat{\mu}_{r(t)}(x) = \frac{1}{1 + \exp[-\alpha_{r(t)}(c_{r(t)}^T x - \theta_{r(t)})]}. \quad (3b)$$

其中 $\alpha_{l(t)} \in \mathbb{R}$, $\alpha_{r(t)} \in \mathbb{R}$, $\theta_t \in \mathbb{R}$, θ_t 为

$$\theta_{l(t)} = \theta_{r(t)} = \frac{\sum_{i=1}^L \mu_t(x^i) \hat{y}_t^i}{\sum_{i=1}^L \mu_t(x^i)} = \frac{\sum_{i=1}^L \mu_t(x^i) c_{l(t)}^T x^i}{\sum_{i=1}^L \mu_t(x^i)}. \quad (4)$$

二叉线性模糊树对输出的逼近误差为

$$\begin{aligned} e &= \sum_{i=1}^L \frac{1}{2} [y^i - \hat{y}_t^i]^2 = \\ &= \sum_{i=1}^L \frac{1}{2} \left[y^i - \frac{\sum_{t \in T} \mu_t(x^i) \hat{y}_t^i}{\sum_{t \in T} \mu_t(x^i)} \right]^2 = \\ &= \sum_{i=1}^L \frac{1}{2} [y^i - \frac{\sum_{t \in T} \mu_t(x^i) c_t^T x^i}{\sum_{t \in T} \mu_t(x^i)}]^2 = \\ &= \sum_{i=1}^L \frac{1}{2} [y^i - c_T^T \bar{x}^i]^2. \end{aligned} \quad (5)$$

向量 c_t 和参数 α 的选择要使估计误差 e 取极小. 为此, 可由最小二乘求得 c_T

$$c_T = (X^T X)^{-1} X^T Y. \quad (6)$$

$X = [\bar{x}^1 \dots \bar{x}^L]^T$, $Y = [y^1 \dots y^L]^T$. 至于 α_t , $t \in \bar{T}$ 的确定, 可采用已有的优化方法, 找到使输出误差全局最小的 α , 恕不赘述. 作为例子, 只考虑梯度下降方法:

$$\alpha_t(j+1) = \alpha_t(j) - p \frac{\partial e}{\partial \alpha_t}, \quad (7)$$

$$\frac{\partial e}{\partial \alpha_t} = \sum_{i=1}^L \frac{\partial e^i}{\partial \alpha_t} = \sum_{i=1}^L (\hat{y}^i - y^i) * \frac{\partial \hat{y}(x)}{\partial \mu_t(x)} \bigg|_{x^i} * \frac{\partial \mu_t(x)}{\partial \alpha_t} \bigg|_{x^i},$$

$$\frac{\partial \hat{y}(x)}{\partial \mu_t(x)} = [\{c_t\}^T x - \hat{y}(x)] / \sum_{t \in T} \mu_t(x),$$

$$\frac{\partial \mu_t(x)}{\partial \alpha_t} = \mu_{p(t)}(x) * [1 - \hat{\mu}_t(x)] * \hat{\mu}_t(x) * [(c_{p(t)}^T x - \theta_t)].$$

当然, 梯度法不能保证找到全局最小, 然而若满足于局部极小, 该方法不失为一个可选择的简单途径.

二叉模糊树生成新的叶节点的过程为: 考虑反模糊化公式(2), 定义

$$e_t = \sum_{i=1}^L [\mu_t(x^i) (y^i - \hat{y}_t^i)]^2, \quad (8)$$

$$e_t^i = \sum_{i=1}^L \left[\mu_t(x^i) (y^i - \frac{\hat{\mu}_{l(t)}(x^i) \hat{y}_{l(t)}^i + \hat{\mu}_{r(t)}(x^i) \hat{y}_{r(t)}^i}{\hat{\mu}_{l(t)}(x^i) + \hat{\mu}_{r(t)}(x^i)}) \right]^2. \quad (9)$$

当

$$e_t > e_t^i \quad (10)$$

成立,则对节点 t 进行划分,否则不再划分,节点 t 成为终节点.

须指出,式(6)的求解可通过递推最小二乘方法、奇异值分解方法等进行;虽然随着子空间的增加,式(6)求解也会出现维数过大难于求解的问题,但是与 ANFIS 因为规则划分过多造成的难于处理高维输入问题并非同一问题.据 Matlab 的能力,模糊二叉树能处理至少 100 维输入的建模问题.

4 二叉线性模糊树对有限样本集或有界连续函数逼近的收敛性 (Convergence of linear binary fuzzy tree's approximation to finite sample set or bounded continuous function)

定理 令 χ 为 n 维实空间中的闭集, $\Omega = \{(x^i, y^i), i = 1, \dots, L\}$, $x^i \in \chi \subset \mathbb{R}^n$, $y^i \in E \subset \mathbb{R}$; E 为一维实空间中的闭集.则必定可由(2)所描述的二叉线性模糊树模型以任意精度逼近之.

证 首先考虑(3)中 $a_{l(t)} \rightarrow \infty$ 及 $a_{r(t)} \rightarrow \infty$ 的情形,即对应图 1 中输入空间的划分不存在模糊区,是所谓的“硬划分”.由(4)可得

$$\theta_{l(t)} = \theta_{r(t)} = \bar{y}. \quad (11)$$

这时有

$$\hat{\mu}_{l(t)}(x) = \begin{cases} 1, & g_t(x) \geq 0, \\ 0, & g_t(x) < 0, \end{cases} \quad (12)$$

$$\hat{\mu}_{r(t)}(x) = \begin{cases} 0, & g_t(x) \geq 0, \\ 1, & g_t(x) < 0. \end{cases} \quad (13)$$

这里

$$g_t(x) = c_t x - \bar{y}_t. \quad (14)$$

$g_t(x) = 0$ 为 n 维空间中的超平面.

二叉树是有向树,据有向树是无回路的有向图的性质,

加之(10)和(5)自然保证了 $e_k = \sum_{i=1}^L e^i$ 逼近误差的单调下降性.由(12),(13)和(6)可知,此时,求解 c_t 简化为

$$x_t c_t = y_t. \quad (15)$$

其中

$$x_t = \begin{bmatrix} x_{t1}^1 \\ \vdots \\ x_{tm}^1 \end{bmatrix}_{m \times n}, \quad y_t = \begin{bmatrix} y_t^1 \\ \vdots \\ y_t^m \end{bmatrix}_{m \times 1}.$$

$\{(x_{t1}^1, y_t^1), \dots, (x_{tm}^1, y_t^m)\} \in \Omega_t, \Omega_t \subset \Omega, t \in \bar{T}$, 即终节点 t 是由 m 个输入/输出样本构成的样本子集 Ω_t .

当 $m > n$ 时,

$$y_t \notin R(x_t). \quad (16)$$

其中 $R(x_t)$ 为矩阵 x_t 的值域.

而 c_t 为式(6)的最小二乘解,其残差

$$\|y_t - \hat{y}_t\| = \|y_t - c_t x_t\| > 0. \quad (17)$$

由于 Ω 为有限个数的样本集, E 为 n 维空间中的闭集.根据有限覆盖定理,对输入空间不断的划分,最终总存在有界的有限非空的 $\bar{T}$, 使一些 $t \in \bar{T}$ 对应的 Ω_t 中有

$$\text{rank } x_t = m = n. \quad (18)$$

而另一些 $t \in \bar{T}$ 的 Ω_t 中有

$$\text{rank } x_t = m < n. \quad (19)$$

若(18)满足,则逼近误差为 0,即

$$e_h^t = \sum_{i \in \Omega_t} \frac{1}{2} \|y_i^t - \hat{y}_i^t\|^2 = 0.$$

对于(19)的出现,可分成两种情形讨论.不妨设 Ω_t 为 $l(s)$ 所对应的输入输出样本子集,而 Ω_{t+1} 则对应 $r(s)$.

第一种情形: Ω_t 出现(19),而 Ω_{t+1} 满足(18).这时总可以在 Ω_{t+1} 中找到至少 $n - m$ 个样本对 $\{(x^k, y^k)\} \in \Omega_{t+1}, kq = 1, \dots, n - m$, 加上 Ω_t 中的样本对满足(10),同时还满足误差

$$e_h^t = e_h^{t+1} = \sum_{i \in \Omega_t \cup \Omega_{t+1}} \frac{1}{2} \|y_i^t - \hat{y}_i^t\|^2 = 0. \quad (20)$$

在选取 Ω_{t+1} 中的 $n - m$ 个样本时,以离 $g_s(x^t)$ 最近的样本为选择优先级最高, $g_s(x^t)$ 为终节点 t 的父节点 s 所形成的划分 $r(s)$ 和 $l(s)$ 的 n 维空间的超平面.

第二种情形: Ω_t 与 Ω_{t+1} 同时出现(19).这时可以参照第一种情形,在 Ω_{t+1} 中找到足够的样本对补充 Ω_t 并在 Ω_t 中找到足够的样本对补充 Ω_{t+1} , 使之均满足(18).由于父节点 Ω_s 一定满足(16)才具备产生子节点的条件,所以具有足够多的样本对使 Ω_t 与 Ω_{t+1} 能互相补充,以使均满足(18)及(20).当 Ω_t 为 $r(s)$ 时可采用类似作法.

因而,在“硬划分”时,逼近误差 e 是单调下降并具有下确界为 0,即对任意给定的 $\varepsilon > 0$,总存在有限非空 $\bar{T}$, 使得

$$e_h = \sum_{i=1}^L e_h^i < \varepsilon.$$

进一步考虑“软划分”情况,即在(13)中 $a_{l(t)}$ 及 $a_{r(t)}$ 取有界实数,即输入空间的划分具有模糊区.当 $n = 1$ 时,图 2 中直线取 l_2, l_3 示意了一种“硬划分”的逼近,而 s_1, s_2 则示意了取不同 $a_{l(t)}$ 及 $a_{r(t)}$ 时用曲线段代替直线段“软划分”逼近.若模糊树构成过程中 $a_{l(t)}$ 和 $a_{r(t)}$ 由梯度下降与模拟退火相结合的方法来优化,能保证在给定节点时逼近误差最小化.故总有

$$0 \leq e \leq e_h.$$

其中 e_h 为 $a_{l(t)} \rightarrow \infty$ 及 $a_{r(t)} \rightarrow \infty$ 时的逼近误差,而 e 为辅助隶属函数经过优化后所得到的逼近误差.

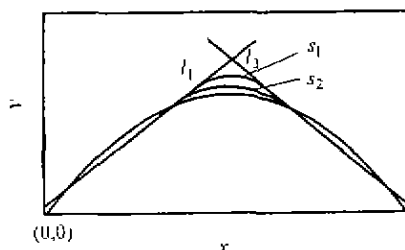


图 2 $\alpha_{l(t)}$ 和 $\alpha_{r(t)}$ 的选择对逼近的影响

Fig 2 Effect of the selection of $\alpha_{l(t)}$ and $\alpha_{r(t)}$ on approximation

以上已证明了 e 的上界 e_h 是单调下降收敛到 0 的哥西序列,故 e 也必定是单调下降极限为 0. 即对给定 $\varepsilon > 0$, 对 n 维

空间中任一闭集上的有限样本集一定存在有界的 $\bar{T}$, 使(2)形的二叉线性模糊树对输出样本集的逼近误差 $e = \sum_{x \in \bar{T}} e_i < \varepsilon$.

证毕

推论 令 $y = f(x)$ 为 n 维空间中一闭集 X 上的有界连续函数, 其中 $x \in X \subset \mathbb{R}^n, y \in E \subset \mathbb{R}$, 则必定可由(2)所描述的模糊树模型以任意精度逼近之.

证 由于 n 维数空间上的闭集 X 具有紧致性, 故总可取有限集 $\{x^i\}, \forall i = 1, \dots, L$, 使有界序列 $\{y^i = f(x^i)\}, \forall i = 1, \dots, L$ 是列紧的. 即当 $L \rightarrow \infty$ 时, 有 $\{y^i = f(x^i)\} \rightarrow y = f(x)$.

又根据上述定理, (2)所描述的模糊树模型可以任意逼近 n 维空间一闭集的有限样本集, 故可得出(2)所描述的模糊树也可以任意逼近闭集上的有界连续函数. 证毕.

5 仿真结果与比较(Simulation result and comparison)

例 将模糊树模型用于图3所示标准例题麦克-格拉斯混沌时间序列的预测^[2]. $\tau > 17$, 数据 101 ~ 210 用作样本(训练数据), 数据 601 ~ 1110 用作检验数据. 本例取 $x = [1, x(t-18), x(t-12), x(t-6), x(t)]$, $y = x(t+6)$, $\alpha_i, i \in T$, 初值取为 0.3. 经过迭代优化, 使得 $\alpha_i = 7.24$.

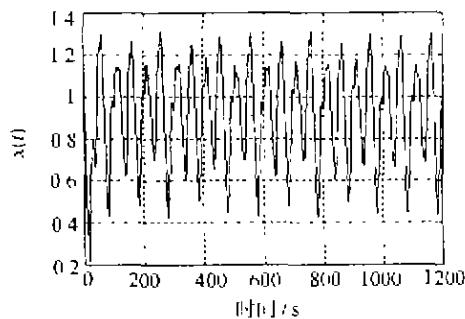


图3 一部分麦克-格拉斯混沌时间序列
Fig. 3 Parts of Mackey-Glass chaotic time series

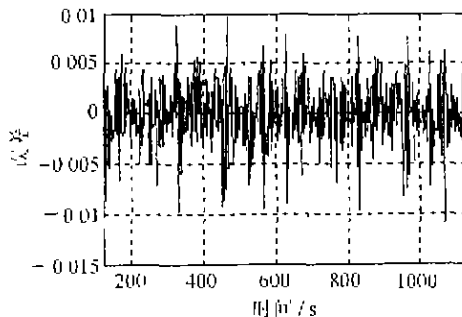


图4 模糊树模型的预测误差
Fig. 4 Prediction error of a fuzzy tree model

图4给出了模糊树模型的预测误差曲线. 表1给出了与其它模型比较的辨识结果. 仿真结果表明, 与 NN 模型相比, 模糊树模型具有计算量小, 精度高的特点, 与 ANFIS 模型相比, 基本达到了相同的误差水平; 然而, 如使用说明中指出, ANFIS 模型只能解决输入维数小于等于 5 的问题, 模糊树模型则具有对于输入空间维数不敏感的优点.^[4]

表1 模糊树模型与其它模型结果比较

Table 1	Result comparisons of fuzzy tree and other models			
	NN 模型	ANFIS 模型	模糊树模型	
样本数	500	500	500	
输入空间维数	4	4	4	
训练均方误差	0.0106	0.0020	0.0027	
预测均方误差	0.0107	0.0019	0.0026	
规则数目	21 个神经元	16	31 个节点	(16 条规则)

6 结论(Conclusion)

本文所提出的二叉线性模糊树模型, 可以对 n 维空间任一闭集上的有限样本集进行任意精度的逼近. 该方法具有计算速度快, 精度高, 对输入空间维数不敏感等优点. 模糊树模型在一定程度上模拟了对复杂问题进行分层分段简化决策的思维过程, 故对其进一步的研究具有理论与应用价值.

参考文献(References)

- [1] Takagi T and Sugeno M. Fuzzy identification of systems and applications to modelling and control [J]. IEEE Trans. Systems, Man and Cybernetics, 1985, 15(1): 116-132
- [2] Jyh-Shung Roger Jang. ANFIS: adaptive-network-based fuzzy inference system [J]. IEEE Trans. Systems, Man, and Cybernetics, 1993, 23(3): 665-685
- [3] Gelfand S B and Ravishanker C S. A tree-structured piecewise linear adaptive filter [J]. IEEE Trans. Information Theory, 1993, 39(6): 1907-1922
- [4] Zhang Jiangang, Mao Jianqin, Xia Tian, et al. Fuzzy tree model and its application in complex system identification [J]. Acta Automatica Sinica, 2000, 26(3): 378-381 (in Chinese)

本文作者简介

毛剑琴 1940 年生. 1964 年毕业于北京大学数学力学系. 1982 年获工学硕士学位. 1985 年获工学博士学位. 现任北京航空航天大学第七研究室教授, 博士生导师, 中国自动化学会理事副秘书长, IEEE 控制系统学会北京分会主席, IEE 会上. 主要研究领域为控制理论与控制工程. 具体是: 系统的智能辨识与参数的智能化估计; 智能控制及其应用; 鲁棒控制及其应用等. Email: durmao@public.bia.net.cn

张建刚 1971 年生. 1993 年毕业于北京航空航天大学自动控制系, 获工学学士学位. 1999 年获北京航空航天大学工学博士学位. 现在摩托罗拉公司任职. 研究领域为控制理论与控制工程, 研究方向为小波理论在系统辨识中的应用, 模糊辨识, 系统智能化辨识与参数的智能化估计等.

岳玉芳 1974 年生. 博士生. 研究方向是计算智能, 模糊建模与控制. 数据具有不确定性的鲁棒建模, 盲信号分离等. Email: yueyufang@263.net

代翼阳 1966 年生. 博士. 研究方向是鲁棒控制理论及应用, 计算智能及其在建模与控制中的应用等.

李幼平 1938 年生. 中国工程院院士. 中国工程物理研究院研究员. 博士生导师. 研究方向是通信与电子系统.