

文章编号: 1000-8152(2008)05-0837-08

T-S模糊广义系统的逼近性

马俊峰, 张庆灵

(东北大学 系统科学研究所, 辽宁 沈阳 110004)

摘要: 本文研究T-S模糊广义系统的逼近性, 给出了T-S模糊广义系统的逼近性定理. 证明其可以以任意的精度逼近一类广泛存在的非线性广义系统. 还将MISO(多输入单输出)情况推广到MIMO(多输入多输出)的情况. 在逼近性定理的基础上, 利用神经网络的方法对非线性广义系统建模, 给出了神经网络的结构及学习算法. 本文共提出了两种神经网络的训练策略, 对各自的优点与不足给出了分析, 最后用数值例子验证了算法的有效性.

关键词: T-S模糊广义系统; 逼近性; 非线性广义系统; 神经网络

中图分类号: TP273 **文献标识码:** A

Approximation property of T-S fuzzy singular systems

MA Jun-feng, ZHANG Qing-ling

(Institute of System Sciences, Northeastern University, Shenyang Liaoning 110004, China)

Abstract: This article focuses on the approximation problem of T-S fuzzy singular systems. An approximation theorem of T-S fuzzy singular systems is proposed. It demonstrates that T-S fuzzy singular systems could approximate to a wide class of nonlinear singular systems with arbitrarily high accuracy. The MISO (multi-input single-output) situation is extended to MIMO (multi-input multi-output) situation. Based on this approximation theorem and using neural network method, the model of T-S fuzzy singular systems can be constructed along with its learning algorithm. Two neural network training strategies are also proposed and their advantages and disadvantages are analyzed respectively. Finally, a numerical example is given to illustrate the validity of the algorithms.

Key words: T-S fuzzy singular systems; approximation property; nonlinear singular systems; neural network

1 引言(Introduction)

自从1965年L. A. Zadeh提出了模糊集的概念以来^[1], 模糊控制在理论上和实际应用当中都取得了快速的发展. 特别是1985年, Takagi, T和Sugeno, M建立了后件为线性方程的T-S模糊系统^[2], 把整个非线性系统看作是多个局部线性系统的模糊逼近, 更适合于我们的研究. 1992年, 王立新首先证明了采用乘积推理、中心平均加权反模糊化、高斯型隶属度函数的模糊系统可以以任意精度逼近任意一个紧集上的连续函数^[3], 从而说明模糊系统是通用的逼近器. 随后针对不同的模糊系统, 有多人证明了这条性质^[4~7]. 其中, T-S模糊系统的逼近性也得到了证明^[8]. 这样对于正常的非线性系统来说, 我们都可以找到一个模糊系统去进行任意精度的逼近.

1999年和2000年, Taniguchi, T等人把正常T-S模糊系统模型推广到一般情况, 提出了T-S模糊广义系统模型^[9,10], 可用来处理非线性广义系统的问题. 而相对于正常系统的逼近定理, T-S模糊广义系统的逼

近定理却没有证明. 在本文中将会看到, T-S模糊广义系统同样是一个通用的逼近器, 它可以以任意精度逼近在紧集上定义的连续函数.

对于给定的一个非线性广义系统来说, 要使用T-S模糊广义系统的控制理论进行控制^[11], 首先必须对系统进行建模, 由本文中将要证明的T-S模糊广义系统的逼近定理知, 对于一类非线性广义系统这种模型是一定存在的. 这里我们利用神经网络的方法对系统进行建模, 利用神经网络快速、准确的参数辨识能力, 可以得到T-S模糊广义系统的模型. 这种网络通常被称为模糊神经网络(FNN), 它将模糊逻辑和神经网络有机的结合在一起, 充分利用了二者的优点, 在各种实际问题中效果较好, 被广泛的应用于各个领域.

2 主要结论(Main results)

2.1 问题描述(Problem formulation)

首先简单介绍T-S模糊广义系统的描述, 设 R_i 为第 i 条规则, 则T-S模糊广义系统模型可以描述为:

收稿日期: 2006-08-20; 收修改稿日期: 2007-12-28.

基金项目: 国家自然科学基金资助项目(60574011); 东北大学流程工业综合自动化教育部重点实验室资助项目.

R_i : if $\zeta_1(t)$ is M_{i1} and $\zeta_2(t)$ is $M_{i2} \cdots$ and $\zeta_P(t)$ is M_{iP} , then

$$\begin{aligned} E\dot{x}(t) &= A_i x(t) + B_{1i} u(t) + B_{2i} w(t), \\ y(t) &= C_i x(t), i = 1, 2, \cdots, N. \end{aligned}$$

其中: M_{ip} 是模糊集, N 是模糊规则数, $x(t) \in \mathbb{R}^K$ 是状态变量, $u(t) \in \mathbb{R}^L$ 是控制输入向量, $w(t) \in \mathbb{R}^L$ 是干扰输入向量, $y(t) \in \mathbb{R}^M$ 是受控输出向量, 一般来说 E 是奇异矩阵, $A_i \in \mathbb{R}^{K \times K}$, $B_{1i}, B_{2i} \in \mathbb{R}^{K \times L}$, $C_i \in \mathbb{R}^{M \times K}$, $\zeta_1(t), \zeta_2(t), \cdots, \zeta_P(t)$ 是前件变量, 令 $\zeta(t) = [\zeta_1(t), \zeta_2(t), \cdots, \zeta_P(t)]^T$, 可选择 $\zeta(t) = x(t)$, $P = K$, 并假设它们与输入变量 $u(t)$ 无关, 在这种假设下采用平行分布补偿算法(PDC)^[12]设计状态反馈控制器时更易于处理. T-S模糊广义系统的全局模型为:

$$\begin{aligned} E\dot{x} &= \sum_{i=1}^N \lambda_i(\zeta) [A_i x + B_{1i} u + B_{2i} w], \\ y &= \sum_{i=1}^N \lambda_i(\zeta) C_i x, \end{aligned} \quad (1)$$

其中:

$$\lambda_i(\zeta) = \frac{\mu_i(\zeta)}{\sum_{i=1}^N \mu_i(\zeta)}, \quad \mu_i = \prod_{p=1}^P \mu_{ip}(\zeta_p), \quad \sum_{i=1}^N \lambda_i(\zeta) = 1.$$

它能够有效的逼近一类广泛存在的非线性广义系统. 考虑下面形式的非线性广义系统模型:

$$\begin{aligned} E\dot{x}(t) &= f(x(t), u(t)) + Bw(t), \\ y(t) &= g(x(t), u(t)), \end{aligned} \quad (2)$$

其中: 状态变量 $x(t) \in \mathbb{R}^n$, 输入变量 $u(t) \in \mathbb{R}^m$, 输出变量 $y(t) \in \mathbb{R}^l$, 干扰变量 $w(t) \in \mathbb{R}^m$, $\text{rank } E = r$, $f(x(t), u(t)), g(x(t), u(t))$ 由变量 $x(t), u(t)$ 的连续函数组成. 设 P, Q 是可逆矩阵, $QEP = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$,

做变量替换 $x = P\tilde{x}$, $\tilde{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$, 则系统(2)等价地转化为下面的微分代数方程的形式:

$$\begin{cases} \dot{x}_1(t) = \tilde{f}_1(x_1(t), x_2(t), u(t)) + B_1 w(t), \\ 0 = \tilde{f}_2(x_1(t), x_2(t), u(t)) + B_2 w(t), \\ y(t) = \tilde{g}(x_1(t), x_2(t), u(t)), \end{cases} \quad (3)$$

其中 $x_1(t) \in \mathbb{R}^r$, $x_2(t) \in \mathbb{R}^q$, $r + q = n$, 是状态变量, $\tilde{f}_1(x_1, x_2, u), \tilde{f}_2(x_1, x_2, u)$ 由变量 x_1, x_2, u 的连续函数组成.

本文要研究的问题就是找到形式如(1)的T-S模糊广义系统来逼近型如(2)或(3)的非线性广义系统. 在第2.2节中, 首先在理论上证明, 对于一类广泛存在的型如(2)或者(3)的非线性广义系统模型, 都可以找到一个型如(1)的T-S模糊广义系统模型, 使得它们

可以以任意的精度逼近. 在第3部分, 我们将利用神经网络的方法, 找到一个这样的模型.

2.2 T-S模糊广义系统的逼近性(Approximation property of T-S fuzzy singular systems)

在给出逼近定理之前, 首先给出一些必要的符号定义. 令集合 $U \subset \mathbb{R}^n$ 是一个紧集, $C(U)$ 表示所有定义在紧集 U 上的连续函数组成的集合, 即 $C(U) = \{f(X) : X \in U, \text{且} f(X) \text{是} X \text{的连续函数}\}$. 设集合 Γ 是定义在紧集 U 上的一组实连续函数组成的集合, 记 $\Gamma = \{f_i(X), X \in U, i = 1, 2, \cdots, q\}$, 定义解集 $\Omega = \{X : f_i(X) = 0, f_i(X) \in \Gamma, i = 1, 2, \cdots, q, X \in U\}$.

引理 1 Ω 是 \mathbb{R}^n 上的有界闭集.

证 根据解集 Ω 中含有解的个数的多少, 可分为3种情况进行讨论:

当 Ω 是空集时, 把空集认为是一种特殊的有界闭集, 此时结论成立. 当 Ω 是有限集合时, 显然有限集合必是有界闭集, 此时结论也正确.

当 Ω 是由无穷多元素组成的集合时, 由于 Γ 是定义在紧集 U 上的, 所以集合 $\Omega \subset U$, 又因为 U 是 \mathbb{R}^n 上的紧集, 它必是有界的, 所以它的子集 Ω 也是有界的.

根据Bolzano-Weierstrass致密性定理, 有界的无限集合 Ω 中至少有一个聚点. 设 X_0 是集合 Ω 的任意一个聚点, 那么显然 X_0 也是 U 的聚点, 又因为 U 是紧集, 因此它的聚点 $X_0 \in U$.

在以 X_0 为中心、1为半径的邻域 $\rho(X_0, X) < 1$ 中选择一点 X_1 , 满足 $X_1 \in \Omega$, 且 $X_1 \neq X_0$, $\rho(X_0, X)$ 表示 X_0 与 X 的Euclid距离, 根据聚点的定义这种点是一定存在的, 记 $\delta_1 = \rho(X_0, X_1)$, 再在邻域 $\rho(X_0, X) < \frac{1}{2}\delta_1$ 中选择一点 X_2 , 满足 $X_2 \in \Omega$, 且 $X_2 \neq X_0$, 记 $\delta_2 = \rho(X_0, X_2)$, 一般地, 在邻域 $\rho(X_0, X) < \frac{1}{2^n}\delta_n$ 中选取一点 $X_n \in \Omega$, 且 $X_n \neq X_0$, 记 $\delta = \rho(X_0, X_n)$, 由此过程, 根据归纳法可以在 Ω 中找到一个数列 $\{X_n\}$, 使其收敛于 X_0 , 记为 $\lim_{n \rightarrow \infty} X_n = X_0$.

对于集合中 $\forall f_i(X) \in \Gamma, i = 1, 2, \cdots, q$, 由于它是紧集 U 上的连续函数, 又有 $X_0 \in U$, 因此 f_i 在 X_0 处是连续的, 对于收敛于 X_0 的数列 $\{X_n\} \subset U$, 可知

$$\lim_{n \rightarrow \infty} f_i(X_n) = f_i\left(\lim_{n \rightarrow \infty} X_n\right) = f_i(X_0),$$

因为对于数列 $\{X_n\}$ 中的点, 总有 $f_i(X_n) = 0$, 因此 $f_i(X_0) = 0$, 所以聚点 $X_0 \in \Omega$.

由 X_0 的任意性, Ω 中的任意一个聚点都属于它本身, 根据闭集的定义, 可知 Ω 是 \mathbb{R}^n 上的闭集.

证毕.

一般的, MIMO T-S模糊系统可以表示成:

$$Y(z(t)) = \sum_{i=1}^N \lambda_i(z(t))Y_i(z(t)), \quad (4)$$

其中 $z(t) \in V$, V 是 \mathbb{R}^n 上的紧集, $Y_i(z(t))$ 是 $z(t)$ 的线性函数组成的向量函数. 这里采用单点模糊化、乘积推理和中心平均加权反模糊化方法构造的 T-S 模糊系统, $Y_i(z(t))$ 在模糊系统中代表第 i 条规则对应的输出向量, $\lambda_i(z(t))$ 表示第 i 条规则的输出权重. 记所有建立在紧集 V 上型如(4)的模糊系统的全体组成的空间为 Ξ_V .

注 1 这里的 MIMO 系统指的是模糊系统的输入空间中有多多个输入变量, 输出空间中有多多个输出变量. 与通常我们在动态系统中讲的多输入多输出系统不同, 那里的多输入表示的是多个控制变量, 比如系统(2)中的 $u(t)$, 多输出表示多个输出变量, 比如系统(2)中的 $y(t)$. 本文提到的 MIMO 系统都与这里的理解一样.

注 2 (4)是 MIMO T-S 模糊系统的一般模型, 当表示 T-S 模糊广义系统时, 部分方程的输出恒为零, 即

$$Y_1(z(t)) = \sum_{i=1}^N \lambda_i(z(t))Y_{1i}(z(t)),$$

$$0 = \sum_{i=1}^N \lambda_i(z(t))Y_{2i}(z(t)), \quad (5)$$

记(5)中零输出方程的解集为 \emptyset .

MIMO 系统(4)可以分解为多个如下形式 MISO 模糊系统^[3]:

$$y(z(t)) = \sum_{i=1}^N \lambda_i(z(t))y_i(z(t)), \quad (6)$$

其中 $y_i(z(t))$ 表示 $y(z(t))$ 对应于第 i 条规则时的输出, $y \in \mathbb{R}$. 记建立在某个紧集 U 上所有型如(6)的模糊系统的全体构成的空间为 Ψ_U .

在 \mathbb{R}^n 上定义两个函数之间的距离

$$d_\infty(y_1 - y_2) = \sup_{z \in \mathbb{R}^n} (|y_1 - y_2|),$$

则 (Ψ_U, d_∞) 构成一个度量空间^[3]. 类似地, 在 \mathbb{R}^n 上定义两个函数向量之间的距离

$$D_\infty(Y_1 - Y_2) = \max_{i=1}^n \sup_{z \in \mathbb{R}^n} (|Y_{1i} - Y_{2i}|),$$

其中 Y_{1i}, Y_{2i} 表示 Y_1, Y_2 的第 i 个分量, 容易验证这个距离定义满足非负性、对称性和三点不等式, 因此 (Ξ_V, D_∞) 也构成度量空间.

引理 2 ^[5] 对于任意给定的定义在紧集 $U \subset \mathbb{R}^n$ 上的实连续函数 g , 以及任意给定的 $\varepsilon > 0$, 存在 $f \in \Psi_U$, 使得 $d_\infty(y(z) - g(z)) < \varepsilon$.

由引理2知, (Ψ_U, d_∞) 在 $(C(U), d_\infty)$ 上稠密.

在系统(3)中, 由于干扰输入是线性的, 不失一般性, 假设 $w(t) = 0$. 表示其中代数方程为 $\Theta_b(z(t)) =$

$\tilde{f}_2(z(t))$, 微分与输出方程为

$$\Theta_a(z(t)) = \begin{bmatrix} \tilde{f}_1(z(t)) \\ \tilde{g}(z(t)) \end{bmatrix}, \quad z(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ u(t) \end{bmatrix},$$

假设 Θ_a, Θ_b 作为函数本身都是定义在紧集 U 上的. 但由于代数约束 $\Theta_b = 0$ 的存在, 定义 $\Omega = \{z : \Theta_b(z) = 0, z \in U\}$, Ω 是 $\Theta_b(z)$ 的零解空间, 实际上函数 Θ_a, Θ_b 都是在论域 Ω 上取值. 设 $\theta(z)$ 是定义在集合 Ω 上非线性广义系统(3)中任意一个实连续函数, 即

$$\theta(z) \in \{\Theta_{a1}, \Theta_{a2}, \dots, \Theta_{a(r+l)}, \Theta_{b1}, \Theta_{b2}, \dots, \Theta_{bq}\},$$

其中 $\Theta_{ai}, i = 1, 2, \dots, (r+l)$ 表示 Θ_a 的第 i 个分量; $\Theta_{bj}, j = 1, 2, \dots, q$ 表示 Θ_b 的第 j 个分量.

定理 1 对于上述定义的 $\theta(z)$, 以及任意给定的 $\varepsilon > 0$, 存在 $f \in \Psi_\Omega$, 使得

$$d_\infty(\theta(z) - f(z)) < \varepsilon.$$

证 由引理1可知, 上面定义的 Ω 是紧集 U 上的有界闭集. 作为代数约束的解空间, Ω 可被认为是一个无限集合, 这样整个的微分代数系统才有意义. 根据 Heine-borel 有限覆盖定理易知, \mathbb{R}^n 中有界闭集与紧集是一致的(紧集通常被认为是无限集合才有意义), 因此集合 Ω 就是 \mathbb{R}^n 上的紧集. 对于定义在紧集 U 上非线性广义系统(3), 由于代数约束的存在, 其中每一个函数 $\theta(z)$ 都是在紧集 Ω 上取值的.

可以证明, $\theta(z)$ 在 Ω 上还是连续的. 对于任意的 $z_0 \in \Omega$, 仿照引理1的证明过程, 可以在紧集 Ω 中构造一个数列 $\{z_n\}$, 使得 $z_n \rightarrow z_0$. 由于 $\theta(z)$ 是紧集 U 上的连续函数, $\Omega \subset U$, 因此 $\theta(z)$ 在点 z_0 处是连续的, 对于 Ω 中收敛于 z_0 的点列 $\{z_n\}$, 必有

$$\lim_{n \rightarrow \infty} \theta(z_n) = \theta(\lim_{n \rightarrow \infty} z_n) = \theta(z_0),$$

因此, $\theta(z)$ 在 Ω 上是连续的.

再由引理2知, 紧集 Ω 上的连续函数 $\theta(z)$ 一定可以在 Ψ_Ω 中找到一个函数 $f(z)$, 使其以任意给定的精度 $\varepsilon > 0$ 逼近 $\theta(z)$, 即 $d_\infty(\theta(z) - f(z)) < \varepsilon$. 证毕.

注 3 T-S 模糊广义系统的逼近问题我们也可以从另一个角度来考虑, 假设集合 U, Ω 以及函数 $\theta(z)$ 的定义与前面一致. 函数 $\theta(z)$ 在不考虑代数约束的情况下, 作为定义在紧集 U 上的连续函数, 由引理2可知, 可以在 Ψ_U 中找到一个模糊系统 $f(z)$, 使得 $\sup_{z \in U} (|\theta(z) - f(z)|) < \varepsilon$. 那么, 当函数 $\theta(z)$ 考虑代数约束时, 即 $\theta(z)$ 在 Ω 上取值, 由定理1的证明知道, 集合 Ω 是集合 U 上的一个紧集, 因此

$$\sup_{z \in \Omega} (|\theta(z) - f(z)|) \leq \sup_{z \in U} (|\theta(z) - f(z)|) < \varepsilon,$$

则

$$d_\infty(\theta(z) - f(z)) = \sup_{z \in \Omega} (|\theta(z) - f(z)|) < \varepsilon,$$

这说明在 Ω 上的任意一个像 $\theta(z)$ 这样的连续函数都可以在 Ψ_Ω 中找到一个模糊系统 $f(z)$ 使它们以任意的精度逼近,也可说成 (Ψ_Ω, d_∞) 在 $(C(\Omega), d_\infty)$ 上是稠密的.

考虑代数方程 $\Theta_{bi}(z) = 0, i = 1, 2, \dots, q$, 既然可以找到 $f_i(z) \in \Psi_\Omega$ 以任意精度逼近0, 那么解集 $\Omega_i = \{z : \Theta_{bi}(z) = 0, z \in U\}$ 与 $\Phi_i = \{z : f_i(z) = 0, z \in U\}$ 有什么关系呢? 下面给出分析.

假设 1 对于 $\forall z_0 \in \Omega_i, \delta > 0$, 在邻域 $\rho(z, z_0) < \delta$ 中可以找到 $z^{(1)}$ 和 $z^{(2)}$, 满足 $\Theta_{bi}(z^{(1)})\Theta_{bi}(z^{(2)}) < 0$.

引理 3 $\forall \varepsilon > 0, \exists f_i(z) \in \Psi_{\Omega_i}$, 满足:

- 1) $d_\infty(f_i(z)) < \varepsilon$;
- 2) $\forall z_1 \in \Phi_i, \exists z_2 \in \Omega_i$, 使得 $\rho(z_1, z_2) < \varepsilon$;
- 3) $\forall z_3 \in \Omega_i, \exists z_4 \in \Phi_i$, 使得 $\rho(z_3, z_4) < \varepsilon$.

证 1) 显然 $\forall \varepsilon_1 > 0, \exists f_i^{(1)}(z) \in \Psi_{\Omega_i}$, 使 $d_\infty(f_i^{(1)}(z) - \Theta_{bi}(z)) = d_\infty(f_i^{(1)}(z)) < \varepsilon_1$.

2) $\forall \varepsilon_2 > 0$, 设 $\Lambda(z') = \{z : \rho(z, z') < \varepsilon_2, z' \in \Omega_i\}$, 令 $W = \bigcup_{z' \in \Omega_i} \Lambda(z')$, 则 W 是一个开集, 它的补集 $\bar{W} = U - W$ 是一个闭集. 令 $d = \inf_{z \in \bar{W}} |\Theta_{bi}(z)|$, 根据引理2, $\exists f_i^{(2)}(z) \in \Psi_U$, 使 $d_\infty(f_i^{(2)}(z) - \Theta_{bi}(z)) < d$. 可以证明 $\forall z \in \Phi_i, \Phi_i = \{z : f_i^{(2)}(z) = 0, z \in U\}$, 必有 $z \in W$; 否则, 若 $\exists z_0 \in \Phi_i$, 满足 $z_0 \in \bar{W}$, 则必有 $|\Theta_{bi}(z_0)| = |f_i^{(2)}(z_0) - \Theta_{bi}(z_0)| > d$, 矛盾.

因此 $\forall z_1 \in \Phi_i, \exists z_2 \in \Omega_i$, 使得 $\rho(z_1, z_2) < \varepsilon_2$.

3) 根据前面假设, $\forall z' \in \Omega_i, \forall \varepsilon_3 > 0$, 在邻域 $M(z') = \{z : \rho(z, z') < \varepsilon_3\}$ 中, 可找到 $z^{(1)}, z^{(2)}$, 满足

$$\Theta_{bi}(z^{(1)})\Theta_{bi}(z^{(2)}) < 0.$$

不妨设

$$\Theta_{bi}(z^{(1)}) > 0, \Theta_{bi}(z^{(2)}) < 0,$$

令

$$d(z') = \min(|\Theta_{bi}(z^{(1)})|, |\Theta_{bi}(z^{(2)})|),$$

根据引理2, $\exists f_{ij}(z) \in \Psi_U$, 使得 $d_\infty(f_{ij}(z) - \Theta_{bi}(z)) < d(z')$, 且函数 $f_{ij}(z)$ 满足 $f_{ij}(z^{(1)}) > 0, f_{ij}(z^{(2)}) < 0$. 根据闭区域连续函数的性质, 一定存在 $z'' \in M(z')$, 满足 $\rho(z', z'') < \varepsilon_3$, 使得 $f_{ij}(z'') = 0$, 因此有 $z'' \in \Phi_i$. 取 $d' = \inf_{z' \in \Omega} d(z')$, 则 $\exists f_i^{(3)}(z) \in \Psi_U$, 满足 $d_\infty(f_i^{(3)}(z) - \Theta_{bi}(z)) < d'$. 因此可知, $\forall z_3 \in \Omega_i, \exists z_4 \in \Phi_i$, 使得 $\rho(z_3, z_4) < \varepsilon_3$.

综上1), 2), 3), $\exists \varepsilon = \min(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, 可找到 $f_i(z)$, 同时满足1, 2, 3. 证毕.

注 4 对于某些特殊的代数约束方程, 例如线性方程来说, 可以找到模糊系统使 Φ_i 与 Ω_i 相同. 对于大多数的非

线性约束方程, 由引理3知, 可以找到模糊系统使它们的约束域以任意的精度逼近.

注 5 引理3的假设条件是为了避免某些特殊的情况. 以二维情况为例, 如图1所示, 当 $y = \Theta_{bi}(z)$ 与 $y = 0$ 相切时, 若找到的逼近函数是 $y_2 = f_2(z)$, 在 $\Theta_{bi} = 0$ 的解 z_0 附近可找到满足 $f_2(z_1) = 0, f_2(z_2) = 0$ 的 z_1, z_2 逼近 z_0 ; 但当找到的逼近函数是 $y_1 = f_1(z)$ 时, z_0 附近无满足 $f_1(z) = 0$ 的点去逼近 z_0 , 此时不能说它们的解集逼近.

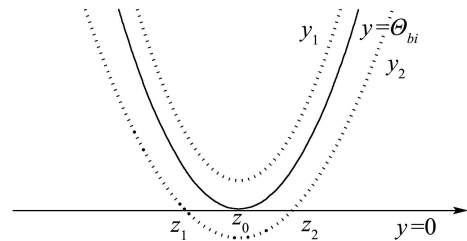


图 1 相切时的图示

Fig. 1 Illustration of tangency case

前面讨论了MISO的情况, 但在实际应用中MIMO更普遍, 而且作为微分代数系统本身, 它的模糊系统模型也必然是多输出的, 因此有必要研究多输出情况. 定理1证明了MISO系统可以在 Ψ_Ω 中找到一个型如(6)的模糊系统进行任意精度的逼近. 但是这样找到的MISO模糊系统之间的规则一般是不同的, 那么是否可以找到公共的规则, 使得在这组规则下, 可以建立一个如(4)的MIMO T-S模糊广义系统模型, 使其逼近广义非线性系统(3)呢, 下面的定理将要证明可以找到一个这样的模型.

定理 2 对于任意型如(3)的非线性广义系统, 当它们的代数方程满足假设条件时, 都可以找到一个型如(5)的T-S模糊广义系统, 使它们相互逼近.

证 设系统(3)共有 M 个输出, $M = r + l + q$. 对于每个输出 $\Theta_{ai}(z), i = 1, 2, \dots, r + l$, 由定理1知道, 可以找到 $y_i(z) \in \Psi_\Omega$, 对 $\forall \varepsilon > 0$, 使得 $d_\infty(\Theta_{ai}(z) - y_i(z)) < \varepsilon/2$.

由模糊系统的构造知, y_i 必是 U 上的连续函数, 因此对于 $\forall \tilde{z} \in \Omega$, 一定可以找到 $\delta(\tilde{z}, \varepsilon) > 0$, 对所有满足 $\rho(z, \tilde{z}) < \delta(\tilde{z}, \varepsilon)$ 的 z , 都有 $|y_i(z) - y_i(\tilde{z})| < \varepsilon/2$, 令 $\Delta_i = \{\delta(\tilde{z}, \varepsilon), \forall \tilde{z} \in \Omega\}, \eta_i = \inf_{z \in \Omega} \Delta_i$, 可以证明 $\eta_i > 0$. 令 $\eta = \min\{\eta_1, \eta_2, \dots, \eta_{r+l}, \varepsilon\}$. 对于每个代数方程 $\Theta_{bi}(z), i = r + l + 1, r + l + 2, \dots, M$, 令 $\Omega_i = \{z : \Theta_{bi}(z) = 0, z \in U\}$, 由引理3, 对于给定的精度 η , 可以找到 $y_i(z) \in \Psi_{\Omega_i}$, 令 $\Phi_i = \{z : y_i(z) = 0, z \in U\}$, 则约束域 Φ_i, Ω_i 相互逼近. 由于 $\Omega = \bigcap_{i=r+l+1}^M \Omega_i$, 令 $\Phi = \bigcap_{i=r+l+1}^M \Phi_i$, 则 Φ, Ω 相互逼近. 设 $Y_1 = [y_1 \ y_2 \ \dots \ y_{r+l}]^T, Y_2 = [y_{r+l+1} \ y_{r+l+2} \ \dots \ y_M]^T = 0$, 则

$$\begin{bmatrix} Y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \sum_{j_1=1}^{N_1} \lambda_{1j_1}(z)y_{1j_1}(z) \\ \sum_{j_2=1}^{N_2} \lambda_{2j_2}(z)y_{2j_2}(z) \\ \vdots \\ \sum_{j_M=1}^{N_M} \lambda_{Mj_M}(z)y_{Mj_M}(z) \end{bmatrix} = \frac{\sum_{j_1=1}^{N_1} \cdots \sum_{j_M=1}^{N_M} \mu_{1j_1} \cdots \mu_{Mj_M} \begin{bmatrix} y_{1j_1} \\ y_{2j_2} \\ \vdots \\ y_{Mj_M} \end{bmatrix}}{\sum_{j_1=1}^{N_1} \cdots \sum_{j_M=1}^{N_M} \mu_{1j_1} \cdots \mu_{Mj_M}}, \quad (7)$$

其中: $N_i, i = 1, 2, \dots, M$ 表示第*i*方程规则条数. (7)是一个型如(5)的MIMO T-S模糊广义系统模型. 对于任意给定的精度 $\varepsilon > 0$, 根据上面模型的建立过程可知, $\forall z \in \Omega, \exists \tilde{z} \in \Phi$, 使得 $\rho(z, \tilde{z}) < \varepsilon$, 且

$$\begin{aligned} D_\infty(\Theta_a(z) - Y_1(\tilde{z})) &= \\ \max_{i=1}^{r+l} (d_\infty(\theta_i(z) - y_i(\tilde{z}))) &< \\ \max_{i=1}^{r+l} (d_\infty(\theta_i(z) - y_i(z)) + d_\infty(y_i(z) - y_i(\tilde{z}))) &< \\ \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{aligned}$$

同样, $\forall z \in \Phi, \exists \tilde{z} \in \Omega$, 使得 $\rho(z, \tilde{z}) < \varepsilon$, 且

$$D_\infty(\Theta_a(\tilde{z}) - Y_1(z)) < \varepsilon,$$

可以认为两个系统相互逼近.

证毕.

作为一种特殊情况, 当 $N_1 = N_2 = \dots = N_M = N, \mu_{1j}(z) = \mu_{2j}(z) = \dots = \mu_{Mj}(z) = \mu_j(z), j = 1, 2, \dots, N$ 时, 即每个输出都有相同的规则时, 上述结果(7)可以转化为简单的形式:

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \frac{\sum_{j=1}^N \mu_j(z) [y_{1j} \quad y_{2j} \cdots y_{Mj}]^T}{\sum_{j=1}^N \mu_j(z)}.$$

在模糊系统的实际应用当中, 对于MIMO情况也是这样来建立模糊系统的, 定理2给出的结论是这种T-S模糊广义系统的存在性, 在应用中还是要找到对每个输出都相同的公共规则建立模糊系统.

3 利用神经网络建立T-S模糊广义系统模型(Constructing T-S fuzzy singular system model using neural network)

在第2部分中, 证明了T-S模糊广义系统的逼近性, 理论上保证了这种系统可以以任意的精度逼近

一类非线性广义系统. 在这节中, 将使用神经网络的方法, 利用它准确、快速的参数辨识能力, 找到一个这样的T-S模糊广义系统, 使其逼近一个给定的非线性广义系统模型. 其中模糊集的隶属度函数选择高斯函数 $\mu(z) = \exp(-\frac{(z-c)^2}{\sigma^2})$.

3.1 神经网络模型(Neural network model)

T-S模糊广义系统的神经网络模型共四层, 如图2所示. 其中, $i = 1, 2, \dots, N, N$ 是规则个数, $j = 1, 2, \dots, M, M$ 是输出变量个数, $k = 1, 2, \dots, K, K$ 是网络输入变量个数.

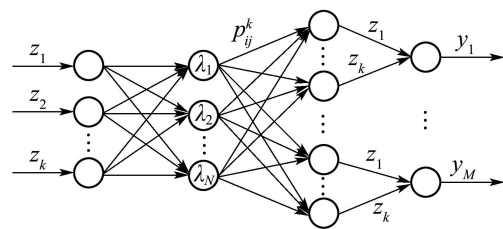


图2 模糊神经网络结构

Fig. 2 Structure of fuzzy neural network

第1层是输入层, 对输入的数据进行模糊化处理, 采用的是单值模糊化;

第2层是规则层, 计算每条规则的输出, 并进行归一化处理;

第3层是连接层, 计算输出空间变量之间线性关系的系数 p_{ij}^k ;

第4层为输出层, 连接权是输入变量, 最后可以得到系统的输出

$$y_j = \sum_{i=1}^N \lambda_i \sum_{k=1}^K p_{ij}^k z_k = \sum_{k=1}^K \sum_{i=1}^N \lambda_i p_{ij}^k z_k, \quad (8)$$

λ_i 的定义见式(1).

3.2 学习算法(Learning algorithm)

建立T-S模糊广义系统的模型需要确定的参数有: 规则数 N , 隶属度函数的中心 c_{ik} , 宽度 σ_{ik} , 以及第3层的连接权 p_{ij}^k . 选择输出误差

$$E = \sum_{j=1}^M (y_j - y_{Ej})^2. \quad (9)$$

y_j 是第*j*个输出变量的网络输出, y_{Ej} 是对应原系统的参考输出.

策略1 网络直接训练方法.

采用定理1的构造思想, 直接在论域 Ω 上来构造模糊系统, 相应的神经网络的学习过程分为3步:

第1步 数据获取, 利用数值计算方法解系统(3)中的代数方程, 可利用MATLAB命令;

第2步 对第1步中得到的输入空间的数据进行聚类, 这里的聚类算法选择模糊c-均值算法(FCM), 得到初始的规则数、隶属度函数的中心. 为了避免

聚类时类中心很接近的情况,这会导致宽度很小,出现不正常的输出,影响权值计算,这里宽度用(0,1]之间的随机数据进行初始化.试验表明,这种方法与以类中心的最小距离做为宽度的方法相比,效果更好.求初始权值时,利用最小二乘法,输入第*i*条规则的中心,调整权值,使得参考输出是系统取这个中心时的输出

$$p_{ij}^k(t+1) = p_{ij}^k(t) - \alpha \frac{\partial E}{\partial p_{ij}^k} = p_{ij}^k(t) - \alpha (y_j(c_i) - y_{Ej}(c_i)) c_{ik},$$

这个过程的初始权值 $p_{ij}^k(0)$ 取(0,1)间的随机值.

第3步 利用BP网络的误差逆传播算法对隶属函数的中心 c_{ik} 、宽度 σ_{ik} 、第3层连接权 p_{ij}^k 进行进一步的修改.利用梯度下降算法,带入式(8)(9)并求导,可以得到各参数的修正律:

$$\begin{aligned} p_{ij}^k(t+1) &= p_{ij}^k(t) - \alpha \frac{\partial E}{\partial p_{ij}^k} = p_{ij}^k(t) - \alpha (y_j - y_{Ej}) \mu_i(z) z_k, \\ c_{ik}(t+1) &= c_{ik}(t) - \beta \frac{\partial E}{\partial c_{ik}} = c_{ik}(t) - \beta \sum_{j=1}^M (y_j - y_{Ej}) (p_{ij} z - y_j) \frac{\lambda_i(z) (z_k - c_{ik})}{\sigma_{ik}^2}, \\ \sigma_{ik}(t+1) &= \sigma_{ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{ik}} = \sigma_{ik}(t) - \gamma \sum_{j=1}^M (y_j - y_{Ej}) (p_{ij} z - y_j) \frac{\lambda_i(z) (z_k - c_{ik})^2}{\sigma_{ik}^3}, \end{aligned} \quad (10)$$

其中 α, β, γ 是收敛系数,它们的大小影响到网络收敛的速度和精度,需要根据经验以及多次试验进行调整.

优点:利用了定理1的模糊系统构造思想,直接应用代数方程约束 Ω 中的数据,利用这样的数据训练网络更有针对性,收敛速度快、精度高.

不足:对于只含有一个代数方程约束的情况非常有效,但当含有多个代数方程时,由于与它们同解的方程有很多,它们在集合 Ω 上都为零,无法把它们区分开.为了解决这个问题,笔者提出了下面这两种阶段的网络训练方法.

策略2 两阶段网络训练方法.

以注3中提到的构造思想为基础,首先在第1阶段不考虑代数方程的约束,把它们当成一组非线性函数,利用神经网络的方法进行逼近.可以得到模糊系统的一个总体的结构模型,但此时的代数方程部分不满足精度要求.第2阶段,在约束集 Ω 中选取若干数据对第1阶段得到的网络进行进一步的优化,可以得到最后的T-S模糊广义系统的模型.

第1阶段

第1步 数据获取,直接在函数本身的定义域 U 上产生随机数据;

第2步和第3步与第1种策略当中的第2,3步相同.

第2阶段

第1步 数据获取,与策略1当中的数据的获取做法相同;

第2步 不需要初始化,直接利用第1步产生的数据对第1阶段产生的初始模型进行训练,学习算法与(10)中相同.

优点:解决了第1种策略中存在的问题,通过第1阶段的训练可以得到一个初始的模型,对应的代数方程中的函数在 U 上取值,避免了在 Ω 上取值都为零的情况.

不足:第1阶段不是在非线性广义系统本身的约束域 Ω 上取值,而是在函数本身的定义域 U 上来划分模糊系统的输入空间,这样就扩大了论域,得到的模糊规则数必有冗余,使规则数增多,计算量增大,训练时间增长.

4 数值例子(Numerical example)

考虑下面的非线性广义系统

$$\begin{aligned} \dot{x}_1 &= x_1^2 + \sin(x_2/2), \\ 0 &= e^{x_1} + x_2, \end{aligned} \quad (11)$$

本文要做的就是找到一个T-S模糊广义系统模型,使其逼近上面给出的非线性模型.在这个模型中,假设 $x_1 \in [-1, 1], x_2 \in [-\pi, \pi]$,在这个区域内观察逼近的效果.对应于前面的论述,这里函数本身的定义域 U 指的就是区域 $[-1, 1] \times [-\pi, \pi]$,而非线性系统本身的约束域 Ω 指的是

$$\Omega = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : 0 = e^{x_1} + x_2, x_1 \in [-1, 1], x_2 \in [-\pi, \pi] \right\}.$$

下面利用神经网络的方法找到T-S模糊广义系统的模型,首先采用第1种训练策略.在约束域 Ω 上产生100组数据对网络进行训练,取收敛系数 $\alpha = 0.1, \beta = 0.01, \gamma = 0.01$.利用MATLAB编写聚类程序,选择规则数 $c = 4$,得到隶属度函数的中心和宽度的网络初始参数.编写神经网络学习程序调整权值,根据式(10)的修正律公式修改参数.经过388的循环,均方根误差可以达到0.010,这个精度通过修改收敛系数和循环次数可进一步提高.简记

$$c_i = (c_{i1}, c_{i2}), \sigma_i = (\sigma_{i1}, \sigma_{i2}), p_i = \begin{bmatrix} p_{i1}^1 & p_{i1}^2 \\ p_{i2}^1 & p_{i2}^2 \end{bmatrix}.$$

则各条规则的隶属度函数的中心为

$$\begin{aligned} c_1 &= (0.3484, -1.4250), c_2 = (-0.2186, -0.4684), \\ c_3 &= (-0.7418, -0.4772), c_4 = (0.8158, -2.2591). \end{aligned}$$

宽度:

$$\sigma_1 = (0.5274, 2.0084), \sigma_2 = (0.4998, 0.2102),$$

$$\sigma_3 = (0.3287, 1.5805), \sigma_4 = (0.6167, 2.7288);$$

第3层连接权:

$$p_1 = \begin{bmatrix} -0.0368 & 0.5752 \\ -0.0107 & -0.0344 \end{bmatrix}, p_2 = \begin{bmatrix} 0.8140 & 0.3521 \\ 0.1581 & -0.1137 \end{bmatrix},$$

$$p_3 = \begin{bmatrix} -1.0274 & 0.1379 \\ -0.0235 & 0.0010 \end{bmatrix}, p_4 = \begin{bmatrix} 0.4510 & 0.0790 \\ 0.4486 & 0.1731 \end{bmatrix}.$$

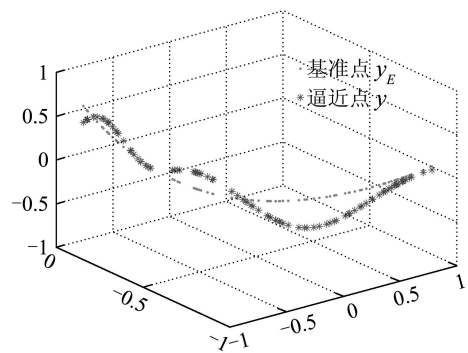
由求出的 c_i 和 σ_i 可以得到模糊规则中模糊集 M_{ik} 对应的隶属度函数 $\mu_{ik}(x) = \exp(-\frac{(x - c_{ik})^2}{\sigma_{ik}^2})$, 利用连接权 p_i , 可以得到模糊规则后件的变量系数 A_i, B_{1i} 和 B_{2i} . 这样可以得到 $N = 4$ 条模糊规则:

R_i : if x_1 is M_{i1} and x_2 is M_{i2} , then

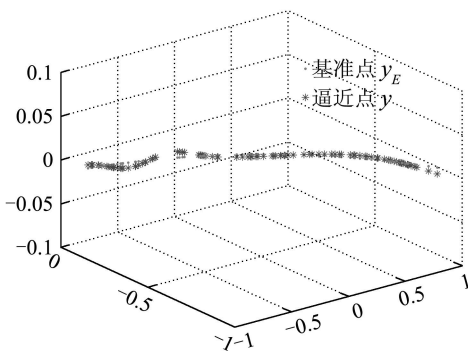
$$x_i = p_{i1}^1 x_1 + p_{i1}^2 x_2,$$

$$0 = p_{i2}^1 x_1 + p_{i2}^2 x_2, i = 1, 2, 3, 4. \quad (12)$$

同时可以建立T-S模糊广义系统的模型, 把得到的模型与给定的非线性系统(11)进行比较, 观察逼近的效果, 如图3所示.



(a) 微分方程



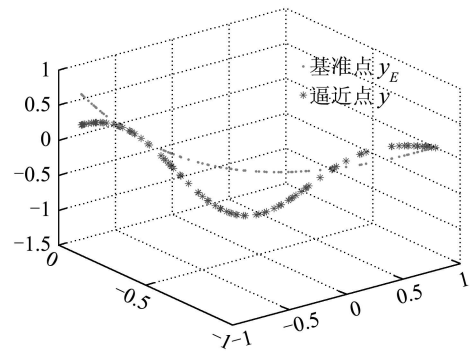
(b) 代数方程

图 3 直接训练方法得到的逼近效果

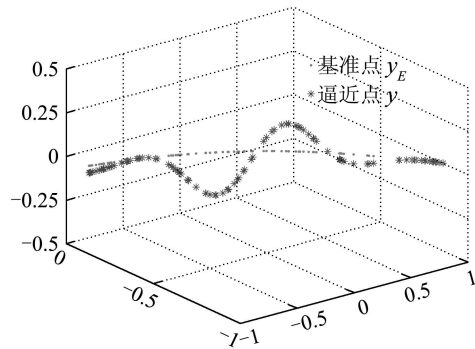
Fig. 3 Approximation results using direct training method

此例中含有一个代数方程, 适合应用策略1, 只用了4条规则经过388步就达到了满意的精度.

采用两阶段网络训练策略, 选择收敛系数 $\alpha = 0.1, \beta = 0.01, \gamma = 0.01$. 在第1阶段, 利用聚类程序, 选择规则数 $c = 7$, 得到隶属度函数的中心和宽度的网络初始参数. 利用网络学习程序调整权值, 经过500步的循环, 均方根误差可达到0.0891. 在约束域 Ω 中取100个点检验得到的网络模型的逼近效果. 对比结果如图4所示, 均方根误差达到0.5903, 显然与第1种方法的结果相比, 虽然规则数目增多, 学习时间增长, 但是逼近效果并不理想. 需要利用约束域当中的数据进行第2阶段的训练.



(a) 微分方程



(b) 代数方程

图 4 经过第1阶段得到的逼近效果

Fig. 4 Approximation results after the first stage training

第2阶段, 在约束域 Ω 上产生100组数据, 利用网络学习程序, 对第1阶段的网络继续训练, 经过87步, 均方根误差达到0.010, 这个精度可以进一步提高. 此时各条规则的隶属度函数的中心为

$$c_1 = (-0.0909, -2.0041), c_2 = (0.6727, 1.1233),$$

$$c_3 = (-0.4548, -1.3049), c_4 = (-0.5214, -1.8288),$$

$$c_5 = (0.3494, 1.8446), c_6 = (0.5007, -0.4049),$$

$$c_7 = (0.1848, -0.9593).$$

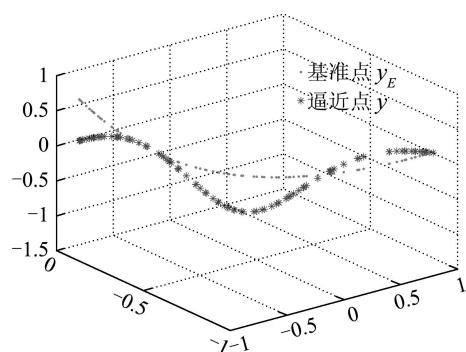
宽度:

$$\begin{aligned}\sigma_1 &= (0.5975, 0.8709), \sigma_2 = (0.6913, 1.3171), \\ \sigma_3 &= (0.7399, 1.0021), \sigma_4 = (1.0416, 0.8227), \\ \sigma_5 &= (0.7918, 0.9974), \sigma_6 = (0.9694, 0.4686), \\ \sigma_7 &= (0.8071, 0.7141);\end{aligned}$$

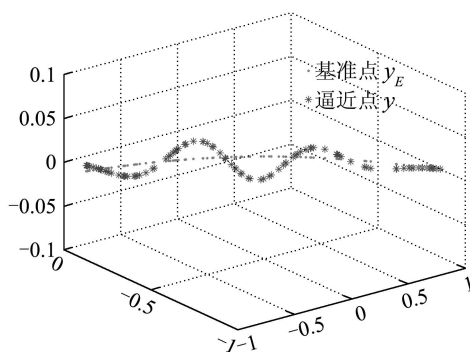
第3层连接权:

$$\begin{aligned}p_1 &= \begin{bmatrix} 0.3692 & 0.6022 \\ 1.3878 & 0.7580 \end{bmatrix}, p_2 = \begin{bmatrix} 0.9951 & 0.8853 \\ 1.1059 & 3.2815 \end{bmatrix}, \\ p_3 &= \begin{bmatrix} -2.2931 & 1.1836 \\ -0.8398 & 0.3583 \end{bmatrix}, p_4 = \begin{bmatrix} 0.2275 & -0.3147 \\ 1.5766 & 0.3992 \end{bmatrix}, \\ p_5 &= \begin{bmatrix} -1.0750 & 0.2750 \\ -0.6595 & 0.9211 \end{bmatrix}, p_6 = \begin{bmatrix} 0.5290 & 0.2903 \\ 1.2199 & -2.3801 \end{bmatrix}, \\ p_7 &= \begin{bmatrix} 0.6385 & 0.2399 \\ 1.2444 & -0.2682 \end{bmatrix}.\end{aligned}$$

类似于策略1中模糊规则(12), 这里可以得到 $N = 7$ 条模糊规则. 同样可以得到T-S模糊广义系统的模型, 逼近效果如图5所示, 比第1阶段的结果明显好出很多. 但是与第1种直接训练方法的结果相比, 规则数要多了3条, 这说明有些规则是冗余的, 这个方法的训练时间也更长.



(a) 微分方程



(b) 代数方程

图5 经过第2阶段得到的逼近效果

Fig. 5 Approximation results after the second stage training

5 结论(Conclusion)

本文证明了T-S模糊广义系统的逼近定理, 把MISO推广到MIMO情况, 并且利用神经网络采用两种方法对非线性广义系统进行建模, 给出了算法, 通过例子验证了算法的有效性. 但是在T-S模糊广义系统的逼近问题上仍存在问题没有解决, 例如非线性广义系统的脉冲问题, 是否能被得到的T-S模糊广义系统所逼近, 因为T-S模糊广义系统在控制的时候一定要考虑脉冲的问题. 这个问题有待进一步的研究.

参考文献(References):

- [1] ZADEH L A. Fuzzy sets[J]. *Information and Control*, 1965, 8(3): 338 – 353.
- [2] TANAKA T, SUGENO M. Stability analysis and design of fuzzy control systems[J]. *Fuzzy Sets and Systems*, 1992, 45(2): 135 – 156.
- [3] WANG L X. Fuzzy systems are universal approximators[C] // *Proceedings of IEEE International Conference on Fuzzy Systems*. San Diego: IEEE Press, 1992: 1153 – 1162.
- [4] KOSKO B. Fuzzy systems as universal approximators[J]. *IEEE Transactions on Computers*, 1994, 43(11): 1329 – 1333.
- [5] CAO G S, REES N W, FENG G. Analysis and design for a class of complex control systems[J]. *Automatica*, 1997, 33(6): 1017 – 1028.
- [6] ZENG X J, SINGH M G. Approximation theory of fuzzy systems-MIMO case[J]. *IEEE Transactions on Fuzzy Systems*, 1995, 3(2): 219 – 235.
- [7] CASTRO J L. Fuzzy logic controllers are universal approximators[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1995, 25(4): 629 – 635.
- [8] YING H. General Takagi-Sugeno fuzzy systems are universal approximators[C] // *Proceeding of IEEE International Conference on Fuzzy Systems*. Anchorage, AK: IEEE Press, 1998, 1: 819 – 823.
- [9] TANIGUCHI T, TANAKA K, YANAFUJI K, et al. Fuzzy descriptor systems: stability analysis and design via LMIs[C] // *Proceedings of American Control Conference*. San Diego, USA: IEEE Press, 1999, 3: 1827 – 1831.
- [10] TANIGUCHI T, TANAKA K, WANG H O. Fuzzy descriptor systems and nonlinear model following control[J]. *IEEE Transactions on Fuzzy Systems*, 2000, 8(4): 442 – 452.
- [11] 刘晓东, 张庆灵, 王岩. T-S模糊广义系统的 H_∞ 控制[J]. *东北大学学报(自然科学版)*, 2002, 23(5): 428 – 430. (LIU Xiaodong, ZHANG Qingling, WANG Yan. H_∞ control of T-S fuzzy descriptor systems[J]. *Journal of Northeastern University(natural science)*, 2002, 23(5): 428 – 430.)
- [12] TANAKA K, IKEDA T, WANG H O. An LMI approach to fuzzy controller designs based on the relaxed stability conditions[C] // *Proceedings of IEEE Internet Conference on Fuzzy Systems*. Barcelona, Spain: IEEE Press, 1992: 171 – 176.

作者简介:

马俊峰 (1981—), 男, 东北大学系统科学研究所硕士, 目前研究方向为模糊控制、鲁棒控制, E-mail: junfeng_ma_china@yahoo.com.cn;

张庆灵 (1956—), 男, 东北大学控制理论与控制工程学科教授, 博士生导师, 主要研究方向为分散控制、鲁棒控制与广义系统理论等, E-mail: qlzhang@mail.neu.edu.cn.