

## 基于核映射的高阶Takagi-Sugeno模糊模型

蔡前凤<sup>1</sup>, 郝志峰<sup>2</sup>, 杨晓伟<sup>3</sup>

(1. 广东工业大学 应用数学学院, 广东 广州 510090; 2. 广东工业大学 计算机学院, 广东 广州 510090;

3. 华南理工大学 理学院, 广东 广州 510061)

**摘要:** 本文研究规则后件为非线性函数的高阶Takagi-Sugeno(TS)模糊系统. 为求解规则后件的函数表达式, 首先通过一个核映射将原输入空间映射到高维特征空间, 使原空间的非线性子模型转化为高维特征空间的线性子模型, 获得了规则后件的非线性函数的计算公式. 然后, 给出了用核模糊聚类和最小二乘支持向量机设计模糊系统的一种新算法. 最后通过4个公开数据集上的仿真实验验证了所提算法的逼近能力、推广能力和鲁棒性能.

**关键词:** 模糊系统; 模糊聚类; 支持向量机; 核函数

**中图分类号:** TP18      **文献标识码:** A

## Higher-order Takagi-Sugeno fuzzy model based on kernel mapping

CAI Qian-feng<sup>1</sup>, HAO Zhi-feng<sup>2</sup>, YANG Xiao-wei<sup>3</sup>

(1. Faculty of Applied Mathematics, Guangdong University of Technology, Guangzhou Guangdong 510090, China;

2. Faculty of Computer Science, Guangdong University of Technology, Guangzhou Guangdong 510090, China;

3. School of Sciences, South China University of Technology, Guangzhou Guangdong 510061, China)

**Abstract:** This paper is concerned with higher-order Takagi-Sugeno(TS) fuzzy systems, where the consequent of a fuzzy rule is a nonlinear combination of input variables. To solve this problem, an implicit nonlinear kernel-mapping is introduced to map the original input space to some higher dimensional feature space, where locally nonlinear submodels of TS fuzzy systems are transformed into locally linear submodels; and then, the expressions of the consequent functions are presented. Furthermore, a novel algorithm of designing higher-order TS fuzzy systems is developed by combining the kernel-based fuzzy clustering with least squares support-vector-machines(LSSVM). Finally, the approximation accuracy, the generalization ability and robustness of the proposed algorithm have been demonstrated by simulation experiments on four well-known data sets.

**Key words:** fuzzy systems; fuzzy clustering; support-vector-machine; kernel function

### 1 引言(Introduction)

模糊系统是基于模糊规则的推理系统, 与传统的建模方法的相比, 其优越性在于能将数值信息和专家经验结合在一起用于复杂系统的建模, 并能有效处理带有模糊性的不确定性信息. 目前, 模糊系统已被广泛地用于许多工程领域: 控制和系统辨识、信号和图像处理、通信、模式识别和数据挖掘等. 然而, 模糊系统存在易陷入局部极值点、推广能力弱和维数灾难等问题.

支持向量机(support-vector-machine, SVM)是基于结构风险最小化(structural risk minimization, SRM)原则的学习算法, 具有全局优化、适应性强、推广能力好等优点, 较好的解决了小样本、过学习、高维数、局部极值点等实际问题. 标准SVM需要求解一个复杂的二次优化问题, 最小二乘支持向量机(least squares support vector machine, LSSVM)是标

准SVM的改版形式, 它将标准SVM的 $\epsilon$ -不敏感损失函数用最小二乘损失函数代替, 不等式约束改为等式约束, 只需要求解一个线性方程组就可得到全局最优解. 与标准SVM相比, LSSVM算法更简单, 参数也更少.

近年来, 用SVM设计模糊系统成为一个新的研究热点. 文献[1~4]研究了规则后件为常数的模糊系统, 即零阶Takagi-Sugeno(TS)模糊系统. 这些研究工作的基本思想都是用支持向量机提取模糊规则, 用支持向量来构造模糊规则的前件和后件; 支持向量机的核函数均由模糊基函数的乘积构成, 模糊规则数等于支持向量的个数. 采用这种方法能够提高模糊系统的推广能力, 确定模糊规则数. 但是由于模糊规则数等于支持向量的个数, 当支持向量的个数很多时, 模糊规则数就会远远超过实际所需的规则个数, 增加了模糊模型的复杂性. 文献[5~7]研究了规

则后件为线性函数的模糊系统,即一阶TS模糊系统,主要思想是用模糊聚类算法构造规则前件,用支持向量机器学习后件的参数,支持向量机的核函数是用前件隶属函数构造的.因为规则数等于聚类数,从而文献[5~7]中所提算法避免了由于支持向量过多而导致太多规则的问题.

本文研究的是规则后件为非线性函数的TS模糊系统,简称为高阶TS模糊系统(higher-order TS fuzzy systems, HFS).与零阶TS模糊系统和一阶TS模糊系统相比,高阶TS模糊系统的规则具有更强的非线性表达能力,因此不仅能减少模糊系统的复杂性,而且具有更好的逼近性能[8].

在非线性函数中,多项式函数的一般表达形式是明确的.因此,目前高阶模糊系统的研究主要集中在多项式模糊规则系统[8~12].文献[8,9]提出用聚类算法来设计规则前件,规则后件中多项式参数用最小二乘法估计.在多项式的阶数为2时,后件中参数有 $M(1+n+\frac{1}{2}(n^2+n))$ 个,其中 $M$ 是规则数, $n$ 是输入维数.这就存在一个问题:随着规则数,输入向量的维数和多项式阶数的增加,模型参数将急剧增加.为了避免参数的快速增加,文献[10,11]中都是事先给出多项式的两种不同形式(阶数小于等于2),然后用遗传算法选择规则后件的表达形式.文献[10]中后件参数用最小二乘法求解,而文献[11]中后件参数用误差反向传播算法求解.文献[12]利用泰勒公式对多项式模糊规则进行解释,这说明高阶模糊系统是具有可解释性的.

对于其它的非线性函数,由于函数表达式难以确定,鲜有文献讨论.针对这个问题,本文通过核函数,构造了一个从原输入空间到高维特征空间的非线性映射,将原空间的非线性子模型转化为高维特征空间的线性子模型,获得了高阶TS模糊系统的规则后件的求解方法.此时,规则前件由基于核的模糊C-均值聚类算法(kernelized fuzzy C-means, KFCM)获得,规则数等于聚类数,从而避免了支持向量过多导致的冗余规则问题;规则后件参数用LSSVM算法求解,因此参数个数不会随规则数和输入维数的增加而快速增长,同时也提高了模糊系统的推广性能.本文所用数学符号说明见附录1.

## 2 高阶TS模糊系统的结构(Description of higher-order TS fuzzy systems)

高阶TS模糊系统的第 $k$ 个模糊规则的表达形式如下:

$$\text{Rule } k: \text{ If } x_1 \text{ is } A_1^{(k)} \text{ and } x_2 \text{ is } A_2^{(k)} \text{ and } \dots \text{ and } x_n \text{ is } A_n^{(k)}, \text{ Then } y \text{ is } f_k(\mathbf{x}). \quad (1)$$

这里 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 是实值输入向量, $A_j^{(k)}$ 为

模糊集, $f_k(\mathbf{x})$ 是非线性函数, $k = 1, \dots, M, j = 1, \dots, n$ . $n$ 是输入向量的维数, $M$ 是规则个数.高阶TS模糊系统的实值输出为

$$f(\mathbf{x}) = \frac{\sum_{k=1}^M f_k(\mathbf{x}) \mu_{A^{(k)}}(\mathbf{x})}{\sum_{k=1}^M \mu_{A^{(k)}}(\mathbf{x})}, \quad (2)$$

其中

$$\mu_{A^{(k)}}(\mathbf{x}) = t(\mu_{A_1^{(k)}}(x_1), \mu_{A_2^{(k)}}(x_2), \dots, \mu_{A_n^{(k)}}(x_n)),$$

$t(\cdot)$ 是 $t$ -范数, $\mu_{A_j^{(k)}}(x_j)$ 表示 $A_j^{(k)}$ 的隶属函数.令

$$S_k(\mathbf{x}) = \frac{\mu_{A^{(k)}}(\mathbf{x})}{\sum_{k=1}^M \mu_{A^{(k)}}(\mathbf{x})},$$

有

$$f(\mathbf{x}) = \sum_{k=1}^M f_k(\mathbf{x}) S_k(\mathbf{x}), \quad (3)$$

其中 $\sum_{k=1}^M \mu_{A^{(k)}}(\mathbf{x}) \neq 0$ .

高阶TS模糊系统实际上是将一个非线性问题分解为 $M$ 个局部非线性子模型的组合.为了获得非线性子模型的后件函数 $f_k(\mathbf{x})$ 的具体形式,本文引入从输入空间到高维空间 $H_1$ 的核映射 $\varphi$ , $\varphi$ 满足 $\varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$ ,其中 $K(\cdot, \cdot)$ 是核函数,如线性核、高斯核或多项式核等等.于是,本文将原空间中的非线性子模型映射为高维特征空间 $H_1$ 中的线性子模型.此时高阶TS模糊系统的第 $k$ 个模糊规则的表达形式(1)变为

$$\text{Rule } k: \text{ If } \mathbf{x} \text{ is } A^{(k)}, \text{ Then } y \text{ is } f_k(\mathbf{x}) = \mathbf{p}^{(k)T} \varphi(\mathbf{x}) + p_0^{(k)}. \quad (4)$$

从而高阶TS模糊系统的实值输出为

$$f(\mathbf{x}) = \sum_{k=1}^M S_k(\mathbf{x}) (\mathbf{p}^{(k)T} \varphi(\mathbf{x}) + p_0^{(k)}). \quad (5)$$

根据式(5),原空间的非线性问题分解成高维特征空间 $H_1$ 中的 $M$ 个局部线性子模型的组合.在此基础上,本文引入从高维特征空间 $H_1$ 到另一个高维特征空间 $H$ 的非线性映射 $\psi$ .记 $\mathbf{x}' = \varphi(\mathbf{x})$ ,映射 $\psi$ 的形式如下:

$$\begin{aligned} \psi^k(\mathbf{x}') &= (S_k(\mathbf{x}) \mathbf{x}'^T, S_k(\mathbf{x}))^T, \\ \psi(\mathbf{x}') &= (\psi^1(\mathbf{x}')^T, \psi^2(\mathbf{x}')^T, \dots, \psi^M(\mathbf{x}')^T)^T. \end{aligned}$$

令 $\Phi(\mathbf{x}) = \psi(\varphi(\mathbf{x}))$ , $\mathbf{p}^{(k)'} = (\mathbf{p}^{(k)T}, p_0^{(k)})^T$ , $\mathbf{P} = (\mathbf{p}^{(1)'}^T, \dots, \mathbf{p}^{(M)'}^T)^T$ ,式(5)可写为

$$f(\mathbf{x}) = \mathbf{P}^T \Phi(\mathbf{x}). \quad (6)$$

通过从原空间到高维特征空间 $H$ 的非线性映射 $\Phi$ ,本

文将原空间中的非线性输入输出模型转化为了一个高维特征空间 $H$ 上的线性回归模型. 因此, 可采用支持向量机算法来构造模糊系统.

不妨记 $\phi^k(\mathbf{x}) = \psi^k(\mathbf{x}') = \psi^k(\varphi(\mathbf{x}))$ , 于是映射 $\Phi$ 的形式为

$$\Phi(\mathbf{x}) = (\Phi^1(\mathbf{x})^T, \Phi^2(\mathbf{x})^T, \dots, \Phi^M(\mathbf{x})^T)^T. \quad (7)$$

### 3 高阶TS模糊系统的回归模型(Regression model for higher-order fuzzy systems)

设训练样本集为 $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l)\}$ , 其中:  $\mathbf{x}_i \in \mathbb{R}^n, y_i \in \mathbb{R}, i = 1, 2, \dots, l$ . 根据结构风险最小化原则, 将高阶TS模糊系统的目标函数设定为:

$$\min_{\mathbf{P}} \frac{1}{2} \sum_{i=1}^l (y_i - \mathbf{P}^T \Phi(\mathbf{x}_i))^2 + \frac{\tau}{2} \mathbf{P}^T \mathbf{P}, \quad (8)$$

其中 $\tau > 0$ . 令 $y_i - \mathbf{P}^T \Phi(\mathbf{x}_i) = e_i$ , 式(8)可转化为:

$$\begin{aligned} \min Q(\mathbf{P}, \mathbf{e}) &= \frac{1}{2} \mathbf{P}^T \mathbf{P} + \frac{1}{2\tau} \sum_{i=1}^l e_i^2 \\ \text{s.t. } &y_i - \Phi(\mathbf{x}_i)^T \mathbf{P} = e_i, i = 1, \dots, l, \end{aligned} \quad (9)$$

其中 $\mathbf{e} = (e_1, e_2, \dots, e_l)^T$ .

比较优化问题(9)与LSSVM, 本文发现在下列条件下优化问题(9)与LSSVM是等价的: i) LSSVM的核函数为 $\Phi(\mathbf{x})^T \Phi(\mathbf{z})$ ; ii)  $b = 0$ ; iii)  $\gamma = \tau^{-1}$ .

因此, 本文提出一种新的模糊核函数 $K^*$ , 形式如下:  $K^*(\mathbf{x}, \mathbf{z}) = \Phi(\mathbf{x})^T \Phi(\mathbf{z})$ . 其中:  $\mathbf{x} = (x_1, \dots, x_n)^T, \mathbf{z} = (z_1, \dots, z_n)^T \in \mathbb{R}^n$ 是任意两个样本. 将 $\Phi(\mathbf{x})$ 的形式(式(7))代入上式有

$$\begin{aligned} K^*(\mathbf{x}, \mathbf{z}) &= \sum_{k=1}^M S_k(\mathbf{x}) S_k(\mathbf{z}) (\Phi(\mathbf{x})^T \varphi(\mathbf{z}) + 1) = \\ &\sum_{k=1}^M S_k(\mathbf{x}) S_k(\mathbf{z}) (K(\mathbf{x}, \mathbf{z}) + 1). \end{aligned} \quad (10)$$

根据文献[13]中的定理, 其具体形式如下:

**定理 1** 令 $K_1$ 和 $K_2$ 是在 $X \times X$ 上的核函数,  $a \in \mathbb{R}^+$ ,  $f$ 是 $X$ 上的实值函数, 那么下面的函数均是核函数:

- 1)  $K(\mathbf{x}, \mathbf{z}) = K_1(\mathbf{x}, \mathbf{z}) + K_2(\mathbf{x}, \mathbf{z})$ ,
- 2)  $K(\mathbf{x}, \mathbf{z}) = K_1(\mathbf{x}, \mathbf{z}) K_2(\mathbf{x}, \mathbf{z})$ ,
- 3)  $K(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) f(\mathbf{z})$ .

在式(10)中, 由于 $S_i(\mathbf{x})$ 是实值函数,  $K(\mathbf{x}, \mathbf{z})$ 是核函数, 显然有 $K^*(\mathbf{x}, \mathbf{z})$ 是核函数. 因此, 优化问题(9)是基于 $K^*(\mathbf{x}, \mathbf{z})$ 的LSSVM, 可用LSSVM算法求解.

**定理 2** 如果 $\sum_{k=1}^M \mu_{A^{(k)}}(\mathbf{x}) \neq 0$ , 令 $\gamma = \tau^{-1}$ , 则原始优化问题(9)可转化为下面的线性方程组

$$(\Omega + \gamma^{-1} \mathbf{e}) \boldsymbol{\alpha} = \mathbf{Y}, \quad (11)$$

其中 $\Omega_{ij} = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j)$ . 并且 $\mathbf{P}$ 可表示为 $\Phi(\mathbf{x}_i)$  ( $i = 1, 2, \dots, l$ )的线性组合, 即

$$\mathbf{P} = \sum_{i=1}^l \alpha_i \Phi(\mathbf{x}_i). \quad (12)$$

证 如果 $\sum_{k=1}^M \mu_{A^{(k)}}(\mathbf{x}) \neq 0$ , 令 $\gamma = \tau^{-1}$ , 优化问题(9)的拉格朗日函数为

$$\begin{aligned} L(\mathbf{P}, \mathbf{e}, \boldsymbol{\alpha}) &= \frac{1}{2} \mathbf{P}^T \mathbf{P} + \frac{1}{2} \gamma \sum_{i=1}^l e_i^2 - \\ &\sum_{i=1}^l \alpha_i (e_i - y_i + \Phi(\mathbf{x}_i)^T \mathbf{P}), \end{aligned} \quad (13)$$

其中 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_l)^T$ 是拉格朗日乘子. 由式(13)的极值条件知:

$$\frac{\partial L}{\partial \mathbf{P}} = 0 \rightarrow \mathbf{P} = \sum_{i=1}^l \alpha_i \Phi(\mathbf{x}_i), \quad (14)$$

$$\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i - \gamma e_i = 0, \quad (15)$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow e_i - y_i + \Phi(\mathbf{x}_i)^T \mathbf{P} = 0. \quad (16)$$

由式(14)直接可推出式(12). 将式(14)和(15)代入式(16)就将原优化问题(9)转化为线性方程组(11).

证毕.

根据定理2, 对式(6)中参数的求解可转化为解线性方程组(11). 记 $H = \Omega + \gamma^{-1} \mathbf{e}$ , 有 $H \boldsymbol{\alpha} = \mathbf{Y}$ , 由于 $K^*(\mathbf{x}, \mathbf{z})$ 是核函数, 所以 $\Omega$ 是半正定阵, 而 $\gamma > 0$ , 因此可用共轭梯度法求解线性方程组(11).

根据定理2, 本文还可推导出规则后件的函数 $f_k(\mathbf{x})$ 的表达形式.

**定理 3** 设 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_l)^T$ 是线性方程组(11)的解, 则高阶TS模糊系统的第 $k$ 个规则后件 $f_k(\mathbf{x})$ 的函数表达形式为:

$$f_k(\mathbf{x}) = \sum_{i=1}^l \alpha_i S_k(\mathbf{x}_i) (K(\mathbf{x}_i, \mathbf{x}) + 1). \quad (17)$$

证 因为 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_l)^T$ 是线性方程组(11)的解, 由定理2, 有 $\mathbf{P} = \sum_{i=1}^l \alpha_i \Phi(\mathbf{x}_i)$ , 从而有

$$\begin{bmatrix} \mathbf{p}^{(1)'} \\ \vdots \\ \mathbf{p}^{(M)'} \end{bmatrix} = \sum_{i=1}^l \alpha_i \begin{bmatrix} \Phi^1(\mathbf{x}_i) \\ \vdots \\ \Phi^M(\mathbf{x}_i) \end{bmatrix}.$$

所以

$$\mathbf{p}^{(k)'} = \sum_{i=1}^l \alpha_i \Phi^k(\mathbf{x}_i), 1 \leq k \leq M.$$

因此对第 $k$ 个规则后件函数, 有

$$f_k(\mathbf{x}) = \mathbf{p}^{(k)'}{}^T \begin{bmatrix} \varphi(\mathbf{x}) \\ 1 \end{bmatrix} =$$

$$\begin{aligned} \sum_{i=1}^l \alpha_i (\phi^k(\mathbf{x}_i))^T \begin{bmatrix} \varphi(\mathbf{x}) \\ 1 \end{bmatrix} &= \\ \sum_{i=1}^l \alpha_i (S_k(\mathbf{x}_i) \varphi(\mathbf{x}_i)^T, S_k(\mathbf{x}_i)) \begin{bmatrix} \varphi(\mathbf{x}) \\ 1 \end{bmatrix} &= \\ \sum_{i=1}^l \alpha_i S_k(\mathbf{x}_i) (K(\mathbf{x}_i, \mathbf{x}) + 1). \end{aligned}$$

证毕.

当核函数 $K(\mathbf{x}_i, \mathbf{x})$ 是线性核时, 所构造的模糊系统就是一阶TS模糊系统. 如果 $K(\mathbf{x}_i, \mathbf{x})$ 是多项式核, 所构造的模糊系统就是多项式型模糊系统. 因此, 根据核函数 $K(\mathbf{x}_i, \mathbf{x})$ 的不同形式, 本文可以得到不同形式的模糊系统, 极大丰富了模糊系统的表达形式.

#### 4 基于核映射的模糊系统 (Fuzzy systems based on kernel mapping)

将式(12)代入式(6), 得到的高阶TS模糊系统的实值输出为

$$f(\mathbf{x}) = \sum_{i=1}^l \alpha_i K^*(\mathbf{x}_i, \mathbf{x}). \quad (18)$$

将KFCM与最小二乘支持向量机算法结合起来可得一种设计高阶TS模糊系统的新方法(KLSSVFS). 高阶TS模糊系统的规则前件用KFCM获得, 规则后件的参数由LSSVM得到, 记

$$\mathbf{z}_i = (\mathbf{x}_i^T, y_i)^T.$$

具体算法如下:

##### 算法1 KLSSVFS.

- 1) 选取规则数, 核函数 $K$ 与核参数.
- 2) 用FCM初始化聚类中心 $\mathbf{v}_k, k = 1, \dots, M$ .
- 3) 采用KFCM对输入输出乘积空间进行划分, 获得新的聚类中心和划分矩阵 $U = (u_{ki})$ ; 根据文献[14], 新的聚类中心的计算公式如下:

如果核函数是高斯核,

$$\mathbf{v}_k = \frac{\sum_{i=1}^l u_{ki} K(\mathbf{z}_i, \mathbf{v}_k) \mathbf{z}_i}{\sum_{j=1}^l u_{kj} K(\mathbf{z}_j, \mathbf{v}_k)}.$$

如果核函数是多项式核,

$$\mathbf{v}_k = \frac{\sum_{i=1}^l u_{ki} K_{d-1}(\mathbf{z}_i, \mathbf{v}_k) \mathbf{z}_i}{(\sum_{j=1}^l u_{kj}) K_{d-1}(\mathbf{v}_k, \mathbf{v}_k)}.$$

其中 $d$ 是多项式核的阶数.

- 4) 利用新的聚类中心和划分矩阵 $U$ 构造规则前件的隶属函数 $\mu_{A_j^{(k)}}(x)$ , 可选用高斯、柯西或三角函数的形式. 前件隶属函数 $\mu_{A_j^{(k)}}(x)$ 的中心为聚类中心

的第 $j$ 个分量, 隶属函数的宽度计算公式如下:

$$\sigma_{kj}^2 = \frac{\sum_{i=1}^l u_{ki} (x_{ij} - v_{kj})^2}{\sum_{i=1}^l u_{ki}}.$$

- 5) 求解线性方程组(11), 得到优化问题(9)的最优解 $\alpha$ .

- 6) 构造决策函数 $f(\mathbf{x}) = \sum_{i=1}^l \alpha_i K^*(\mathbf{x}_i, \mathbf{x})$ .

需要说明的是, 通过第3节的分析, 本文发现KLSSVFS本质上是基于模糊核函数 $K^*$ 的LSSVM.

#### 5 仿真结果 (Simulation results)

本节首先通过Mackey-Glass混沌系统上的仿真实验分析不同核函数对所提模型的影响, 然后通过3个公开数据集上的实验来说明基于核映射的高阶TS模糊系统在逼近精度、推广能力和鲁棒性能方面的改进. 所有的实验均在MATLAB的环境下完成. FMID(fuzzy modeling and identification)前件采用投影的方法建立, 后件采用GLS算法, 使用了FMID工具箱, 标准SVM算法采用LIBSVM工具箱, LSSVM使用了LSSVM工具箱. 模糊系统中 $t$ -范数用乘积算子, 自适应神经模糊系统(adaptive-network-based fuzzy inference system, ANFIS)和KLSSVFS的前件隶属函数为高斯隶属函数. 评价指标采用RMSE(root-mean-square-error).

**例1** 预测混沌时间序列Mackey-Glass(MG)系统<sup>[15]</sup>

$$\frac{dx}{dt} = \frac{a \cdot x(t - \tau)}{1 + x^{10}(t - \tau)} - b \cdot x(t),$$

其中回归时嵌入维数为4. 数据集由1000个输入输出对组成, 500个用于学习, 500个用于测试. 本文在此数据集上分析了不同核函数对KLSSVFS性能的影响. 对于不同的核函数, 下面参数值选取相同:  $\gamma = 100, M = 8$ .

表1给出了KLSSVFS分别使用线性核、多项式核和高斯核时在MG系统上的训练精度和测试精度, 这两种精度均采用RMSE作为评价指标. KLSSVFS有两组核参数: KFCM中的核参数和模糊核函数的核参数. 如果使用高斯核, 核参数记为 $(\sigma_1, \sigma_2)$ , 其中:  $\sigma_1$ 是KFCM中的核参数,  $\sigma_2$ 是模糊核函数中的核参数.  $\sigma_1$ 的选择范围为 $\{2, 2^2, \dots, 2^9, 2^{10}\}$ . 通过实验发现, 对于MG系统,  $\sigma_2$ 的取值不宜太大, 因此 $\sigma_2$ 的选择范围为 $[0.01, 10]$ . 如果使用多项式核, 核参数记为 $(d_1, d_2)$ , 选择范围均为 $[1, 10]$ . 表1中给出了用10-折交叉验证法得到的最优参数值. 从表1可以看出多项式核的训练精度和测试精度要高于线性核, 而高斯核的训练结果和测试结果都是最优的.

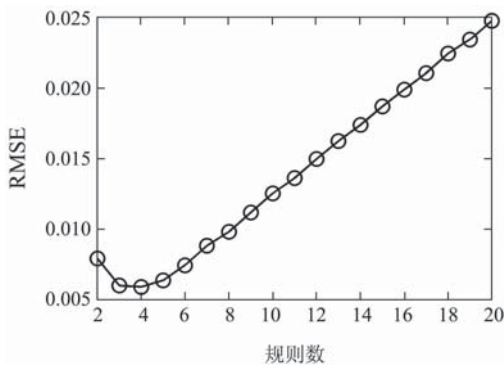
表 1 MG系统上KLSSVFS的实验结果  
Table 1 Experimental results of KLSSVFS on MG systems

类型	核参数	训练精度	测试精度
线性核	—	0.0308	0.0303
多项式核	(2, 2)	0.0206	0.0206
高斯核	(128, 0.1)	0.0022	0.0022

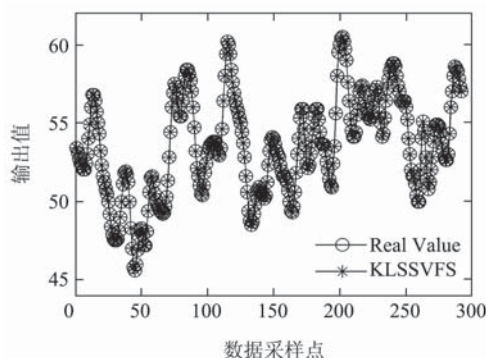
通过例1的实验结果, 本文发现KLSSVFS使用高斯核时的性能要远远优于多项式核和线性核. 因此, 在下面的实验中核函数 $K$ 均使用高斯核.

**例 2** Gas-furnace数据集在模糊系统识别中被广泛使用. 如Xu和Lu<sup>[16]</sup>, Wang和Langari<sup>[17]</sup>, Sugeno<sup>[18]</sup>及Li和Yang<sup>[19]</sup>都使用过. 该数据集一共有296个样本. 输入:  $x(t) = (y(t-1), u(t-4))$ , 输出:  $y(t)$ . 这里使用Gas-furnace数据集来考察所提算法的逼近能力, 因此所有数据都用于训练.

KLSSVFS的参数为:  $\gamma = 200, \sigma_1 = 16, \sigma_2 = 0.01$ . 图1(a)给出了Gas-furnace数据集上RMSE随规则数的递增发生的改变. 从图1(a)可发现, KLSSVFS在规则数为4时, 逼近精度是最高的; 当规则数大于4时, 逼近精度随着规则数的增加而减少. 图1(b)给出了规则数为2时, KLSSVFS在Gas-furnace数据集上的逼近结果.



(a) 不同规则数时KLSSVFS的逼近精度



(b)  $M = 2$ 时KLSSVFS的模拟结果

图 1 KLSSVFS Gas-furnace 数据集上的实验结果

Fig. 1 Experimental result of KLSSVFS on Gas-furnace data set

表2给出了KLSSVFS与其它6种模糊系统学习算法及标准SVM, LSSVM的比较结果, 其中SVs表示支持向量数.

表 2 不同算法逼近精度比较  
Table 2 Training errors of different methods

算法	输入数	规则数/SVs	RMSE
Xu和Lu	2	25	0.573
Wang和Langari	2	5	0.397
Sugeno	3	6	0.435
Li和Yang	2	2	0.401
FMID	2	4	0.337
ANFIS	2	4	0.351
标准SVM	2	289	0.363
LSSVM	2	—	0.352
KLSSVFS	2	2	0.008

从表2中发现: KLSSVFS的逼近精度要比这8种算法都要高, 并且所需规则数也是最少的. 在传统的模糊系统算法中, FMID的逼近精度是最高的,  $RMSE = 0.337$ , 所用规则数是4. 标准SVM的10-折交叉验证最优参数为 $\epsilon = 0.001, C = 64, \sigma = 4$ . 此时,  $RMSE = 0.363$ , 支持向量数是289. LSSVM的10-折交叉验证的最优参数为 $\gamma = 5000, \sigma = 7.37$ ,  $RMSE = 0.352$ . 而KLSSVFS在规则数等于2时,  $RMSE = 0.008$ . KLSSVFS的逼近精度与其它算法相比提高了40多倍.

**例 3** Stock数据集和Sunspots数据集. Stock数据集是一个股票价格数据集, 由Sugeno和Yasukawa<sup>[15]</sup>给出, 是模糊系统用于非线性回归的经典数据集之一. 该数据集由100个样本构成, 每个样本含10个输入属性和一个输出属性. Sunspots数据集是一个时间序列数据集, 主要是对太阳黑子数进行预测, 数据来源于Box和Jenkins<sup>[16]</sup>. 数据集含从1700到1979每年的太阳黑子数, 一共280个样本. 模型输入为 $x_n = (x(n-1), x(n-2), \dots, x(n-12))^T$ , 模型输出为 $y(n) = x(n)$ . 本文在这两个公开数据集上考察KLSSVFS算法的推广性能和稳定性.

为了使得算法性能和所选样本的顺序无关, 本文采用如下10-折交叉验证: 1) 对于某个数据集, 采用10-折交叉法随机生成10个子集, 并使每个子集的样本数尽量相等. 同时, 将其中的9个子集组合成训练集, 剩下的1个子集作为测试集. 这样, 本文得到10组训练集和测试集; 2) 在10组训练集上采用10-折交叉法搜索最优参数, 分别计算在相应测试集上的RMSE; 3) 将10个测试集上的RMSE的平均值作为该数据集的RMSE, 用来评估该算法的性能.

在10-折交叉验证的第2步, 对于支持向量机算法, 为了获得最好的测试精度, 必须为每一组数据集

选择最优参数,这里采用网格搜索方法. LSSVM中主要参数为 $\gamma$ 和核参数 $\sigma$ ,使用LSSVM工具箱在区间上进行网格搜索. KLSSVFS中 $\sigma_1$ 的选择范围与例1相同, $\gamma$ 和 $\sigma_2$ 的搜索方法与LSSVM相同.

规则数的选取方法:规则数从2到10以步长1进行变化,在10组训练集上运行FMID和ANFIS,计算在相应测试集上的RMSE. 10个测试集上的RMSE的平均值的最小值对应的规则数作为这两种算法的最优规则数. KLSSVFS的规则数小于或等于这两者中最小规则数.

图2和图3给出了KLSSVFS的中间结果. 这些结果与10-折交叉验证的第2步相对应. 图中,横轴有10个刻度,每一个刻度表示一组训练/测试集,纵轴表示最高测试集精度. 此外,本文把同一个算法获得的最小测试RMSE用线段连接起来. 在图中给出了FMID, ANFIS, LSSVM和KLSSVFS获得的RMSE.

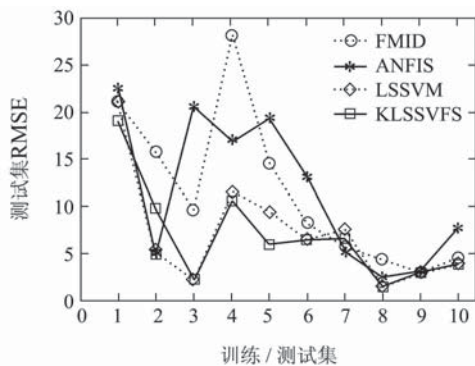


图2 Stock数据集KLSSVFS的10-折交叉验证的测试结果  
Fig. 2 Result of stock price data set for ten-fold cross validation

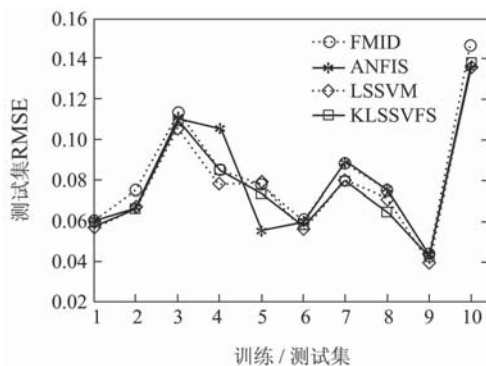


图3 sunspots数据集KLSSVFS的10-折交叉验证测试结果  
Fig. 3 Result of sunspots data set for ten-fold cross validation

图2表明:对于Stock数据集,与FMID相比,除了第7组外, KLSSVFS的测试精度始终高于FMID;与ANFIS相比,除了第2, 7组外, KLSSVFS的测试精度高于ANFIS;与LSSVM相比,除了第2组外, KLSSVFS的测试精度是最高的. 即除了第2, 7组外, KLSSVFS的测试精度都优于其他算法. 图3表明:对于Sunspot数据集,与FMID比, KLSSVFS的连接线段始终位于下方. 与ANFIS相比,除了第5组之外,

KLSSVFS的测试精度高于ANFIS. 与LSSVM相比,除了第4组之外,其余数据组KLSSVFS的测试精度高于或接近于LSSVM.

通过以上分析,本文得到结论: KLSSVFS的性能几乎在横轴的每个坐标处都高于FMID和ANFIS; KLSSVFS的推广性能几乎高于或接近于LSSVM的推广性能.

表3给出了不同算法在10组训练/测试集上的平均测试RMSE. 对于Stock数据集, KLSSVFS的测试精度是最高的. 对于Sunspots数据集, KLSSVFS高于FMID和ANFIS,与LSSVM接近. 从而KLSSVFS的推广性能优于FMID和ANFIS,不逊于LSSVM.

表3还给出了不同算法之间的方差,其表示算法的稳定性(对某些数据集出现不好的预测结果). 对于Stock数据集, KLSSVFS的方差是最低的;对于Sunspots数据集, KLSSVFS的方差比FMID和ANFIS要低,比LSSVM要高. 主要原因在于在第10组测试集上KLSSVFS的RMSE太高,造成太大的偏差. 因此, KLSSVFS的稳定性要优于FMID和ANFIS.

表3 实际数据集上的测试精度±标准方差  
Table 3 Test errors ± standard deviation on real data sets

模型	Stock		Sunspots	
	M	RMSE	M	RMSE
FMID	3	11.557 ± 8.331	3	0.0824 ± 0.0292
ANFIS	4	11.611 ± 7.7999	2	0.0796 ± 0.0288
LSSVM	-	7.198 ± 5.8228	-	0.0768 ± 0.0273
KLSSVFS	2	6.908 ± 5.2622	2	0.0772 ± 0.0280

例4 鲁棒性能比较. 对于Stock数据集,选取第9组训练集和测试集,因为在第9组数据上,4种算法的测试精度相近;对于Sunspots数据集,选取第1组训练集和测试集,因为在第1组数据上,4种算法的测试精度相近. 在训练集中加入自由度为2的 $\chi^2$ 噪声,测试集保持不变.

表4给出了这4种算法在加入了噪声的数据集上的测试精度.

表4 噪声数据集上的测试精度  
Table 4 Test errors on noisy data sets

数据集	FMID	ANFIS	LSSVM	KLSSVFS
Stock	6.7229	8.2786	5.2885	4.8801
Sunspot	3.9281	2.4318	2.1614	2.0260

表5给出了LSSVM和KLSSVFS的测试结果对应的最优参数,这些参数均用10-折交叉验证法网格搜索得到. 从表4看出,在两个数据集上KLSSVFS的测试精度都是最高的,其中Stock数据集上RMSE

= 4.8801, Sunspot数据集上RMSE = 2.026, 这个结果比LSSVM要好, 远远优于FMID和ANFIS. 这一结果显示, KLSSVFS的鲁棒性能要优于FMID, ANFIS和LSSVM.

表5 实验结果所对应的参数

Table 5 Parameters of experimental results

参数	LSSVM		KLSSVFS	
	Stock	Sunspot	Stock	Sunspot
$\gamma$	2.6	0.95	2	0.62
核参数	19	18.5	(2,16)	(8,16.8)

## 6 结论(Conclusions)

高阶TS模糊系统是非线性子模型的组合, 因此具有比一阶TS模糊系统更好的逼近能力. 本文给出高阶TS模糊系统的规则后件的求解方法, 并提出了将KFCM聚类算法和最小二乘支持向量机结合起来构造高阶TS模糊系统的新算法. 通过在4个公开数据上的实验, 分析了所设计的高阶TS模糊系统的逼近精度、推广能力和鲁棒性能. 实验结果表明了与传统的模糊模型相比, 本文所提出新算法在逼近精度、推广能力和鲁棒性能3方面都得到了改进. 与LSSVM相比, 所设计的新算法只需极少的规则数, 就能获得相同的推广能力, 而且具有更好的鲁棒性能和显著的逼近能力.

## 参考文献(References):

- [1] CHEN Y X, WANG J Z. Support vector learning for fuzzy rule-based classification systems[J]. *IEEE Transactions on Fuzzy Systems*, 2003, 11(6): 716 – 728.
- [2] CHEN Y X, WANG J Z. Kernel machines and additive fuzzy systems: classification and function approximation[C] // *Proceedings of the 12th IEEE International Conference on Fuzzy Systems*. St. Louis, Missouri: IEEE, 2003, 2: 789 – 795.
- [3] CHIANG J H, HAO P Y. Support vector learning mechanism for fuzzy rule-based modeling: a new approach[J]. *IEEE Transactions on Fuzzy Systems*, 2004, 12(1): 1 – 12.
- [4] 字正华, 赵爽, 王光昶. 模糊逻辑系统与支持向量机的关系探索[J]. *计算机工程*, 2004, 30(1): 117 – 119. (ZI Zhenghua, ZHAO Shuang, WANG Yongchang. Study of relationship between fuzzy logic system and support vector machine[J]. *Computer Engineering*, 2004, 30(1): 117 – 119.)
- [5] LESKI J K. On support vector regression machines with linguistic interpretation of the kernel matrix[J]. *Fuzzy Sets and Systems*, 2006, 157(8): 1092 – 1113.
- [6] 蔡前凤, 郝志峰, 刘伟. 基于模糊划分和支持向量机的TS模糊系统[J]. *模式识别与人工智能*, 2009, 22(3): 411 – 416. (CAI Qianfeng, HAO Zhifeng, LIU Wei. TS fuzzy systems based on fuzzy partition and support vector machines[J]. *Pattern Recognition and Artificial Intelligence*, 2009, 22(3): 411 – 416.)
- [7] 蔡前凤, 郝志峰, 杨晓伟. 基于最小二乘支持向量机的TSK模糊模型[J]. *华南理工大学学报(自然科学版)*, 2009, 37(5): 130 – 134. (CAI Qianfeng, HAO Zhifeng, YANG Xiaowei. TSK fuzzy model base on least-squares support vector machines[J]. *Journal of South China University of Technology(Natural Science Edition)*, 2009, 37(5): 130 – 134.)

- [8] DEMIRLI K, MUTHUKUMARAN P. Higher order fuzzy system identification using subtractive clustering[J]. *Journal of Intelligent and Fuzzy Systems*, 2000, 9(3): 129 – 158.
- [9] COCCIONI M, LAZZERINI B, MARCELLONI F. Estimating the concentration of optically active constituents of seawater by Takagi-Sugeno models with quadratic rule consequents[J]. *Pattern Recognition*, 2007, 40(10): 2846 – 2860.
- [10] OH S K, PEDRYCZ W, ROH S B. Genetically optimized fuzzy polynomial neural networks with fuzzy set-based polynomial neurons[J]. *Information Sciences*, 2006, 176(23): 3490 – 3519.
- [11] PARK B J, PEDRYCZ W, OH S K. A design of genetically oriented fuzzy relation neural networks (FrNNs) based on the fuzzy polynomial inference scheme[J]. *IEEE Transactions on Fuzzy Systems*, 2009, 17(6): 1310 – 1323.
- [12] HERRERA L J, POMARES H, ROJAS I, et al. TaSe, a Taylor series-based fuzzy system model that combines interpretability and accuracy[J]. *Fuzzy Sets and Systems*, 2005, 153(3): 403 – 427.
- [13] 邓乃杨, 田英杰. 数据挖掘中的新方法: 支持向量机[M]. 北京: 科学出版社, 2004. (DENG Naiyang, TIAN Yingjie. *New Method in Data Mining: Support Vector Machines*[M]. Beijing: Science Press, 2004.)
- [14] CAI Q F, LUI W. TS fuzzy model using kernel-based fuzzy c-means clustering[C] // *Proceedings of the 18th International Conference on Fuzzy Systems*. Jeju Island, Korea: IEEE, 2009, (1): 308 – 312.
- [15] BOX G E P, JENKINS G M. *Time Series Analysis, Forecasting and Control*[M]. San Francisco: Holden Day, 1970.
- [16] XU C W, LU Y Z. Fuzzy model identification and self-learning for dynamic systems[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1987, 17(4): 683 – 689.
- [17] WANG L, LANGARI R. Complex systems modeling via fuzzy logic[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1996, 26(1): 100 – 106.
- [18] SUGENO M, YASUKAWA T. A fuzzy-logic-based approach to qualitative modeling[J]. *IEEE Transactions on Fuzzy Systems*, 1993, 1(1): 7 – 31.
- [19] LI W, YANG Y. A new approach to TS fuzzy modeling using dual kernel-based learning machines[J]. *Neurocomputing*, 2008, 71(16-18): 3660 – 3665.

## 附录 数学符号说明(Appendix for mathematical notations)

$x_i$ : 第*i*个样本的输入向量;  $y_i$ : 第*i*个样本的输出值;  $n$ : 输入向量维数;  $l$ : 样本个数;  $M$ : 规则个数;  $A^{(k)}$ : 第*k*个规则的前件模糊集;  $\mu(\cdot)$ : 隶属函数;  $f_k(x)$ : 第*k*个规则的后件函数;  $U$ : 模糊聚类的划分矩阵;  $p^{(k)}, p_0^{(k)}$ : 在特征空间 $H_1$ 中第*k*个规则的后件参数;  $K(x, z)$ : 核函数;  $\gamma$ : LSSVM 的控制参数;  $C$ : 标准SVM 的控制参数.

### 作者简介:

蔡前凤 (1973—), 女, 博士, 副教授, 目前研究方向为模糊系统、支持向量机与智能计算, E-mail: caiqianfeng@163.com;

郝志峰 (1968—), 男, 教授, 博士生导师, 目前研究方向为进化计算、数据挖掘与模式识别, E-mail: mazhfao@scut.edu.cn;

杨晓伟 (1969—), 男, 教授, 硕士生导师, 目前研究方向为神经网络、支持向量机与仿真算法, E-mail: xwyang@scut.edu.cn.