

基于模糊子空间聚类的O阶L2型TSK模糊系统

邓赵红* 张江滨 蒋亦樟 史荧中 王士同

(江南大学数字媒体学院 无锡 214122)

摘要: 经典数据驱动型 TSK(Takagi-Sugeno-Kang)模糊系统在获取模糊规则时, 会考虑数据的所有特征空间, 其带来一个重要缺陷: 如果数据的特征空间维数过高, 则系统获取的模糊规则繁杂, 使系统复杂度增加而导致解释性下降。该文针对此缺陷, 探讨了一种基于模糊子空间聚类的O阶L2型 TSK 模糊系统(Fuzzy Subspace Clustering based zero-order L2- norm TSK Fuzzy System, FSC-0-L2-TSK-FS)构建新方法。新方法构建的模糊系统不仅能缩减模糊规则前件的特征空间, 而且获取的模糊规则可对应于不同的特征子空间, 从而具有更接近人类思维的推理机制。模拟和真实数据集上的建模结果表明, 新方法增强了面对高维数据所建模型的解释性, 同时所建模型得到了较之于一些经典方法更好或可比较的泛化性能。

关键词: Takagi-Sugeno-Kang(TSK)模糊系统; 医疗诊断; 解释性; 高维数据

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2015)09-2082-07

DOI: 10.11999/JEIT150074

Fuzzy Subspace Clustering Based Zero-order L2-norm TSK Fuzzy System

Deng Zhao-hong Zhang Jiang-bin Jiang Yi-zhang Shi Ying-zhong Wang Shi-tong

(School of Digital Media, Jiangnan University, Wuxi 214122, China)

Abstract: The classical data driven Takagi-Sugeno-Kang (TSK) fuzzy system considers all the features of trained data, and faces a challenge that the interpretation is degenerated and the obtained fuzzy rule is complex when trained by high dimensional data. In this paper, a new fuzzy model, *i.e.*, Fuzzy Subspace Clustering based zero-order L2-norm TSK Fuzzy System (FSC-0-L2-TSK-FS) is proposed to overcome this difficulty. The proposed fuzzy system not only reduces the feature spaces of the rule of antecedent, but also makes different rules implement the inference in different subspaces. The inference mechanism of the proposed fuzzy model training algorithm is very similar to the inference procedure of human. The experimental studies on the synthetic and real datasets prove that the interpretation of model constructed by the proposed method is enhanced when trained by high dimensional data and the generalization performance is better or comparative to several classical TSK fuzzy systems training methods.

Key words: Takagi-Sugeno-Kang (TSK) fuzzy system; Medical diagnosis; Interpretability; High-dimensional data

1 引言

模糊系统是以模糊集合^[1]和模糊推理理论为基础演变而来的一种人工智能系统, 其主要特点是能够将自然界的模糊语言转化成与人类推理机制相似的模糊规则。目前, 模糊系统已被广泛应用于各个领域, 如图像处理、智能决策等^[2,3]。与其他的人工智能技术相比, 模糊系统不仅具有强大的学习能力, 而且具有高度解释性^[4,5]。模糊系统凭借这一独特优势, 在医疗诊断中被广泛应用^[6,7]。当一种智能模型在医疗诊断中被应用时, 人们希望它具有强的解释

能力和良好的可靠性。模糊系统通过模糊推理机制, 对某种医学疾病建立相应的模糊规则, 模拟专家诊断疾病的过程, 帮助医生解决复杂的医学诊断问题。

在已有的几种经典模糊系统中, Takagi-Sugeno-Kang(TSK)^[8,9]模糊系统由于其输出的简洁性和良好的逼近性能被广泛探讨和运用^[10,11]。不同于早期基于专家经验的模糊系统构建方法, 数据驱动型模糊系统正日益成为最主要的模糊系统构建方法。经典数据驱动型 TSK 模糊系统在模型训练过程中, 采用训练样本的全部特征来构建模糊规则。而对于模糊逻辑系统来说, 其解释性主要来自于模糊规则中的模糊集可对应于人类的语言描述, 故经典 TSK 模糊系统在对高维数据进行训练时, 会因采用大量的特征而使其规则太复杂, 使得清晰度和解释性下降。

2015-01-13 收到, 2015-05-11 改回, 2015-06-29 网络优先出版
国家自然科学基金(61170122), 江苏省杰出青年基金(BK20140001)
和新世纪优秀人才支持计划(NCET120882)资助课题

*通信作者: 邓赵红 dzh666828@aliyun.com

在医疗诊断中，要求系统具有很强的解释性，而大量医学数据都具有较高的特征空间。若将经典TSK模糊系统用于医疗检测，势必导致系统的解释性骤降。本文所提出的FSC-0-L2-TSK-FS模型能够利用FSC聚类，将模糊规则映射到不同子空间。当从高维数据中获取模糊规则时，规则前件并不需要数据的全部特征，而是采用FSC聚类所抽取的子空间特征，从而有效地降低了所建模型的复杂性，增强了规则的清晰度和解释性。将FSC-0-L2-TSK-FS模型与医疗诊断相结合，每条模糊规则可从不同的视角(特征空间)对疾病进行医疗诊断，更符合专家诊断的推断过程。

2 高维数据驱动的模糊系统建模

2.1 经典TSK模糊系统简介

根据文献[12]，经典模糊系统包括Takagi-Sugeno-Kang模糊系统模型(TSK-FS)，Mamdani-Larsen模糊系统模型(ML-FS)^[13]和广义模糊系统模型(GFM)^[14]。由于TSK模型的广泛应用，本文以该模型为例探讨高维数据驱动的模糊系统建模以及面临的挑战。对于经典TSK模糊模型，其模糊规则定义如下。

第 k 条模糊规则：

$$\left. \begin{array}{l} \text{IF } x_1 \text{ is } A_1^k \wedge x_2 \text{ is } A_2^k \wedge \cdots \wedge x_d \text{ is } A_d^k \\ \text{THEN } f^k(\mathbf{x}) = p_0^k + p_1^k x_1 + \cdots + p_d^k x_d, \\ k = 1, 2, \dots, K \end{array} \right\} \quad (1)$$

其中 $\mathbf{x} = (x_1 \ x_2 \ \cdots \ x_d)^T$ 为输入向量； A_i^k 表示输入向量 \mathbf{x} 第 i 维特征所对应的第 k 条模糊规则的模糊子集； \wedge 为模糊合取操作； K 表示模糊规则数。每条规则都对应于输入向量 \mathbf{x} ，并把输入空间的模糊子集 $A^k \subset R^d$ 映射到输出空间的模糊集 $f_k(\mathbf{x})$ ， $A^k = (A_1^k, A_2^k, \dots, A_d^k)^T$ 。令 $\mu^k(\mathbf{x})$ 为模糊集 $f_k(\mathbf{x})$ 对应的模糊隶属函数，第 k 条规则的隶属函数 $\mu^k(\mathbf{x})$ 可由式(1)中IF部分中的各维度对应的隶属度值通过合取操作 \wedge 获取，即

$$\mu^k(\mathbf{x}) = \mu_1^k(x_1) \wedge \mu_2^k(x_2) \wedge \cdots \wedge \mu_d^k(x_d) \quad (2)$$

若将乘算子、加算子分别作为合取和析取操作算子，加法算子作为组合算子，同时采用重心反模糊操作，经过去模糊化之后，最终的输出可表示为

$$f(\mathbf{x}) = \frac{\sum_{k=1}^K \mu^k(\mathbf{x}) f^k(\mathbf{x})}{\sum_{k=1}^K \mu^k(\mathbf{x})} = \frac{\sum_{k=1}^K \prod_{j=1}^d \mu_j^k(x_j) f^k(\mathbf{x})}{\sum_{k=1}^K \prod_{j=1}^d \mu_j^k(x_j)} \quad (3)$$

2.2 经典TSK模糊系统高维数据之挑战

实践表明，经典TSK模糊系统在处理高维数据时，其解释性和简洁性下降严重，具体表现在以下

两方面。

(1)规则前件。经典TSK模型规则前件对应于输入向量 $\mathbf{x} = (x_1 \ x_2 \ \cdots \ x_d)^T$ 的所有特征。若 d 的数值很大，则第 k 条规则所对应的模糊子集 A^k 个数增多，模糊规则前件变得复杂，解释性下降。

(2)规则后件。经典TSK模型的输出 $f_k(\mathbf{x})$ 为自定义的模糊子集，最常见的是输入向量的线性组合形式，即 $f_k(\mathbf{x}) = p_0^k + p_1^k x_1 + \cdots + p_d^k x_d$ 。若 d 的数值很大，则第 k 条规则待求参数 p_i^k 个数增多，解释性和简洁性下降。

3 基于模糊子空间聚类的O阶L2型TSK模糊系统建模

3.1 模糊规则构建

根据2.2节所提出的经典TSK模糊系统之不足，本文探讨了一种基于模糊子空间聚类的O阶L2型TSK模糊系统(Fuzzy Subspace Clustering based zero-order L2-norm TSK Fuzzy System, FSC-0-L2-TSK-FS)构建方法，其所对应的模糊规则如下。

第 k 条模糊规则：

$$\left. \begin{array}{l} \text{IF } x_1^k \text{ is } A_1^k \wedge x_2^k \text{ is } A_2^k \wedge \cdots \wedge x_{m_k}^k \text{ is } A_{m_k}^k \\ \text{THEN } f^k(\mathbf{x}) = p_0^k, k = 1, 2, \dots, K \end{array} \right\} \quad (4)$$

与经典TSK模糊系统规则相比，该规则具有如下特色。

(1)规则前件。第 k 条规则对应于输入向量 \mathbf{x} 的子集，即 $\mathbf{x}^k = (x_1^k \ x_2^k \ \cdots \ x_{m_k}^k)^T$ ，并且每条规则对应于 \mathbf{x} 不同的特征子空间。这样既缩减了规则前件对应于输入向量的特征维数，又可使每条规则从不同的视角进行推理，增强了模糊系统的解释性。

(2)规则后件。每条规则的输出只有常数项，除去后面与 \mathbf{x} 冗长的线性组合形式，使得TSK模糊系统在处理高维数据时仍能保持高的简洁性。

3.2 基于模糊子空间聚类的规则前件学习

模糊子空间聚类^[15](FSC)能将高维数据空间转化到其相关的子空间中进行聚类。在模糊子空间聚类中，每一类给定权重向量 \mathbf{w} ，用于表示不同的特征(维度)对于该类的贡献程度。FSC-0-L2-TSK-FS可根据FSC中特征权重思想，为规则抽取重要特征。根据文献[16]，FSC的优化目标函数及其约束条件可写成

$$\left. \begin{array}{l} J_{\text{FSC}} = \sum_{i=1}^K \sum_{k=1}^d w_{ik}^\tau \left(\sum_{j=1}^N u_{ij}(x_{jk} - v_{ik})^2 + \varepsilon_0 \right) \\ \text{s.t. } u_{ij} \in \{0, 1\}, \sum_{i=1}^K u_{ij} = 1, 0 < \sum_{j=1}^N u_{ij} < N, \\ 0 \leq w_{ij} \leq 1, \sum_{k=1}^D w_{ik} = 1 \end{array} \right\} \quad (5)$$

式中, K 为聚类数, N 为样本数, d 为样本维数, ε_0 是个小的正则化常数, 如 $\varepsilon_0 = 1e-5$ 。 \mathbf{U} 是硬划分矩阵, u_{ij} 表示第 j 个样本是否属于第 i 类; \mathbf{V} 是聚类中心, v_{ik} 表示第 i 个聚类中心第 k 维的数值; \mathbf{W} 是模糊权值矩阵, w_{ik} 表示第 k 个特征(维度)在第 i 个聚类中的贡献程度。

对于给定的训练数据集 $D_{tr} = \{\mathbf{x}_i, y_i\}$, $\mathbf{x}_i \in R^d$, $y_i \in R$, $i = 1, 2, \dots, N$, 通过 FSC 算法进行聚类, 得到的每一类对应于 FSC-0-L2-TSK-FS 模型构建的一条模糊规则。将矩阵 \mathbf{W} 按照一定的准则进行处理, 可为每条模糊规则选取出不同的特征。矩阵 \mathbf{U} 和矩阵 \mathbf{V} 为构建模糊子集 \mathbf{A} 相应的模糊隶属函数提供了依据, 其中 $\mathbf{A} = (\mathbf{A}^1 \ \mathbf{A}^2 \ \dots \ \mathbf{A}^K)^T$, \mathbf{A}^k 表示第 k 条规则所对应的模糊子集。

矩阵 \mathbf{W} 的每一行向量 \mathbf{w}_k 与模糊规则 R^k 呈一一对应关系。根据矩阵 \mathbf{W} , 选取出的第 k 条模糊规则的特征表示为

$$x_1^k \ x_2^k \ \dots \ x_{m_k}^k, \quad k = 1, 2, \dots, K \quad (6)$$

其中 $\mathbf{x}^k = (x_1^k \ x_2^k \ \dots \ x_{m_k}^k)^T \in R^{m_k}$ 表示的是第 k 条模糊规则选取的特征所构成的向量。具体地, 给出两种选取策略。(1)固定特征选取数目。给定一个整型数值 m , $1 \leq m \leq D$, 第 k 条模糊规则根据向量 \mathbf{w}_k 的值, 选取前 m 个贡献程度最大的特征。此时, 每条模糊规则选取的特征数目相同。(2)设定阈值 γ 。给定 $\eta = \gamma/D$, 第 k 条模糊规则根据向量 \mathbf{w}_k 的值, 选取贡献程度不小于 η 的特征, 其中 $0 < \gamma \leq D$, 为人工设置的参数。此时, 每条规则选取的特征数目可不同。

FSC-0-L2-TSK-FS 模型的每条模糊规则对应输入向量 \mathbf{x} 的特征子空间, 特征子空间 \mathbf{x}^k 又与模糊子集 \mathbf{A}^k 对应。若采用高斯隶属函数作为模糊子集的隶属度函数, 第 k 条规则的模糊子集 A_i^k 的隶属函数可以表示为

$$\mu_{A_i^k}(x_i^k) = \exp\left[\frac{-(x_i^k - c_i^k)^2}{2\delta_i^k}\right] \quad (7)$$

其中 x_i^k 为第 k 条模糊规则根据 \mathbf{w}_k 选取的第 i 个特征, 参数 c_i^k 和 δ_i^k 可通过 FSC 的聚类结果来估计, 有

$$c_i^k = \frac{\sum_{j=1}^N u_{jk} x_{ji}^k}{\sum_{j=1}^N u_{jk}} \quad (8)$$

$$\delta_i^k = h \frac{\sum_{j=1}^N u_{jk} (x_{ji}^k - c_i^k)^2}{\sum_{j=1}^N u_{jk}} \quad (9)$$

其中, x_{ji}^k 表示的是第 j 个训练样本 \mathbf{x}_j 在第 k 条模糊规则中所选取的第 i 个特征维度, $\mathbf{x}_j = (x_{j1} \ x_{j2} \ \dots \ x_{jd})^T$; u_{jk} 是 FSC 聚类得到的硬划分矩阵 \mathbf{U} 中的元素, h 为人工可调的尺度参数。

3.3 基于 ε -不敏感损失函数的规则后件学习

根据 3.2 节规则前件的特征选取以及式(3), FSC-0-L2-TSK-FS 模型的输出可以表示成

$$f(\mathbf{x}) = \sum_{k=1}^K \frac{\mu^k(\mathbf{x}^k)}{\sum_{k'=1}^K \mu^{k'}(\mathbf{x}^{k'})} f^k(\mathbf{x}^k) = \sum_{k=1}^K \tilde{\mu}^k(\mathbf{x}^k) p_0^k \quad (10)$$

其中

$$\mu^k(\mathbf{x}^k) = \prod_{i=1}^{m_k} \mu_{A_i^k}(x_i^k) \quad (11)$$

$$\tilde{\mu}^k(\mathbf{x}^k) = \mu^k(\mathbf{x}^k) \left/ \sum_{k'=1}^K \mu^{k'}(\mathbf{x}^{k'}) \right. \quad (12)$$

根据文献[17,18]的策略, 可将 $f(\mathbf{x})$ 表示成线性回归问题。令

$$\mathbf{x}_g = (\tilde{\mu}^1(\mathbf{x}^1) \ \tilde{\mu}^2(\mathbf{x}^2) \ \dots \ \tilde{\mu}^K(\mathbf{x}^K))^T \quad (13)$$

$$\mathbf{p}_g = (p_0^1 \ p_0^2 \ \dots \ p_0^K)^T \quad (14)$$

则式(10)可表示成

$$f(\mathbf{x}) = \mathbf{p}_g^T \mathbf{x}_g \quad (15)$$

本文采用 ε -不敏感损失函数和 L2 范式惩罚技术对模糊规则后件参数 \mathbf{p}_g 进行优化。根据文献[18, 19]的基本原理, FSC-0-L2-TSK-FS 模型的优化目标函数构建为

$$\min_{\mathbf{p}_g} E = \sum_{i=1}^N \left| \mathbf{p}_g^T \mathbf{x}_{gi} - y_i \right|_{\varepsilon} \quad (16)$$

通过引入 L2 范式惩罚项 $\xi_i^+ \geq 0$ 和 $\xi_i^- \geq 0$, 目标函数式(16)等价于

$$\left. \begin{aligned} \min_{\mathbf{p}_g, \xi_i^+, \xi_i^-} E &= \sum_{i=1}^N \left((\xi_i^+)^2 + (\xi_i^-)^2 \right) \\ \text{s.t.} \quad &\begin{cases} y_i - \mathbf{p}_g^T \mathbf{x}_{gi} < \varepsilon + \xi_i^+ \\ \mathbf{p}_g^T \mathbf{x}_{gi} - y_i < \varepsilon + \xi_i^- \end{cases}, \quad \forall i \end{aligned} \right\} \quad (17)$$

类似于 L2-SVR^[20]的学习原理, 在引入结构风险项和正则化项之后, 最终的优化目标函数可表示成

$$\min_{\mathbf{p}_g, \xi_i^+, \xi_i^-, \varepsilon} L = \frac{1}{\tau} \frac{1}{N} \sum_{j=1}^N \left((\xi_j^+)^2 + (\xi_j^-)^2 \right) + \frac{1}{2} \mathbf{p}_g^T \mathbf{p}_g + \frac{2}{\tau} \varepsilon \quad (18)$$

通过最优化理论, 将式(18)转化为对偶问题:

$$\left. \begin{aligned} \max_{\lambda_i^+, \lambda_i^-} & - \sum_{i=1}^N \sum_{j=1}^N (\lambda_i^+ - \lambda_i^-) (\lambda_j^+ - \lambda_j^-) \cdot \mathbf{x}_{gi}^T \mathbf{x}_{gj} \\ & - \frac{N\tau}{2} \sum_{i=1}^N (\lambda_i^+)^2 - \frac{N\tau}{2} \sum_{i=1}^N (\lambda_i^-)^2 \\ & + \sum_{i=1}^N \lambda_i^+ \cdot y_i \cdot \tau - \sum_{i=1}^N \lambda_i^- \cdot y_i \cdot \tau \\ \text{s.t.} \quad & \sum_{i=1}^N (\lambda_i^+ + \lambda_i^-) = 1, \quad \lambda_i^+ \geq 0, \quad \lambda_i^- \geq 0 \quad \forall i \end{aligned} \right\} \quad (19)$$

通过求解式(19)的标准二次规划, 可以求得 λ_i^+ 和 λ_i^-

的最优解。根据对偶理论，可以得到原最优问题 p_g 的解。

$$p_g = \frac{2}{\tau} \sum_{i=1}^N (\lambda_i^+ - \lambda_i^-) \mathbf{x}_{g_i} \quad (20)$$

最终，根据 3.2 节获取的 FSC-0-L2-TSK-FS 规则前件，以及本小节运用二次规划理论求解得到的模糊规则后件参数，整合得到 FSC-0-L2-TSK-FS 模型和模糊规则库。

4 实验研究

4.1 实验设置

实验部分分别采用模拟数据集和 UCI 中的 Heart, Breast 数据集进行实验评估和验证。本文选取 3 种经典 TSK 模糊系统模型作为对比算法，分别为基于 ε -不敏感损失函数的 L2 型 TSK 模糊系统^[18] (L2-TSK-FS)，基于 IQP 优化算法的 ε -不敏感损失函数的 TSK 模糊系统^[17] (ε -TSK-FS(IQP))、基于 LSSLI 优化算法的 ε -不敏感损失函数的 TSK 模糊系统^[17] (ε -TSK-FS(LSSLI))。实验中，采用 5 倍交叉验证策略对参数进行寻优。其中，模糊规则数 k 的寻优集合为 {2,3,4,5,6,7,8,9,10,11,25,32}；高斯函数尺度参数 h 的寻优集合为 $\{10^{-2}, 10^{-1}, 10^0, 10^1, 10^2\}$ ；规则化参数 τ 的寻优集合为 $\{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2, 10^3\}$ ；特征选取参数 γ 的寻优集合为 {0.6, 0.62, 0.64, ..., 1.2}；FSC 模糊指数参数 r 的寻优集合为 {1.05, 1.5, 5, 10, 50, 100, 1000}。

为对各算法所得模型的泛化性能进行对比，采用分类正确率作为实验评价指标。正确率越接近 1，说明模型的性能越好。为对所构建的模糊系统的复杂度进行对比，采用指标 Ψ 来描述模型的复杂度。其中 K 表示模糊规则数目， NT_i 表示第 i 条模糊规则前件所抽取的特征数目， NY_i 表示第 i 条模糊规则后件参数个数。 Ψ 值越小，说明模型的规则复杂度越低，清晰度越高，语言解释性越强。

$$J_{\text{cla}} = \frac{\text{正确分类个数}}{\text{样本总个数}} \times 100\% \quad (21)$$

$$\Psi = \sum_{i=1}^K (2NT_i + NY_i) \quad (22)$$

4.2 模拟数据集实验

为体现模型适宜于高维数据和解释性强的特性，构造的模拟数据需具有如下特征：(1)具有较高的维数；(2)具有若干重要特征，并适当添加一些非重要特征和干扰项；(3)添加的重要特征需对应于不同的特征子空间。根据设置，生成 400 组 30 维 4 类分类数据。100 组数据为一类，每类所对应的重要特征分别为 $R_1\{3,7,12,18,25\}$, $R_2\{6,7,8,9,10,11,12\}$, $R_3\{15,17,19,21,23,24\}$, $R_4\{2,7,12,17,22\}$ 。

算法的性能对比如表 1 所示。从中可以得出结论如下：(1)由于受到非重要特征强烈的干扰作用，3 种经典 TSK 模型的系统性能下降严重，而 FSC-0-L2-TSK-FS 模型仍能保持良好的系统性能。(2)3 种经典 TSK 模型采用模拟数据的全部特征(30 个特征)来构建模糊规则。规则复杂度 Ψ 非常大，说明模糊规则太复杂，导致模型解释性下降严重。FSC-0-L2-TSK-FS 模型只为规则选取一些重要特征，规则复杂度 Ψ 降低至 105，增强了模型的清晰度和解释性。

FSC-0-L2-TSK-FS 模型达到最优时，为每条模糊规则抽取的子空间特征如表 2 所示。通过表 2 可以发现，规则所抽取的子空间特征与模拟数据的重要特征绝大部分是相吻合的。说明 FSC-0-L2-TSK-FS 模型具有准确抓取数据重要特征的能力。图 1 所示为 FSC-0-L2-TSK-FS 模型的每条模糊规则空间特征分布情况。由图 1 可以看出，FSC-0-L2-TSK-FS 模型通过 FSC 算法得到的聚类有着明显的子空间特征，其权值最高的几个特征恰好与表 2 所对应。

4.3 真实数据集仿真实验

本部分通过采用 UCI 机器学习库中的 Heart, Breast 数据集来进行实验。Heart 数据集共有 270 组实验样本，每组样本有 13 个特征(维度)。270 组实验样本被分成两类，分别为正常类和心脏病类。Breast 数据集共有 699 组数据，每组数据包含 10 个特征(维度)。699 组数据中，有 458 组良性乳腺癌病例和 241 组乳腺癌病例。

各算法的实验结果和评价指标如表 3 所示。通过表 3 可得出结论有：(1)与 3 种经典 TSK 模型相

表 1 模拟数据实验算法性能对比

算法	最优参数	J_{cla} (%)	Ψ	采用的特征数目
L2-TSK-FS	$k = 11, h = 10^{-2}, \tau = 10^{-1}$	25.50	1001	每条规则 30 个特征
ε -TSK-FS(IQP)	$k = 32, h = 10^{-1}, \tau = 10^{-2}$	62.75	2912	每条规则 30 个特征
ε -TSK-FS(LSSLI)	$k = 25, h = 10^{-1}, \tau = 1$	61.75	2275	每条规则 30 个特征
FSC-0-L2-TSK-FS	$k = 7, \gamma = 1, \tau = 10^{-2},$ $h = 10^{-3}, r = 10$	99.75	105	7 条规则的特征数目 分别为：5,9,7,5,6,10,7

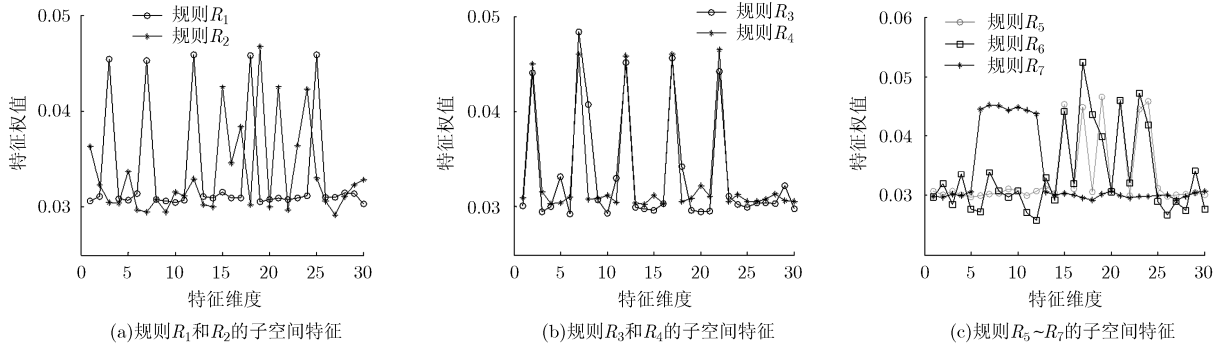


图1 模拟数据在 FSC-0-L2-TSK-FS 模型最优时每条规则的空间特征分布情况

表2 模拟数据实验 FSC-0-L2-TSK-FS 模型为每条规则抽取的特征

规则	规则特征提取
R_1	3,7,12,18,25
R_2	1,5,15,16,17,19,21,23,24
R_3	2,7,8,12,17,18,22
R_4	2,7,12,17,22
R_5	15,17,19,21,23,24
R_6	4,7,15,17,18,19,21,23,24,29
R_7	6,7,8,9,10,11,12

比, FSC-0-L2-TSK-FS 模型的性能与其相差无几, 在误差允许范围内。(2)3 种经典 TSK 模型无法为规则抽取重要的子空间特征, 导致规则复杂度 ψ 很高。FSC-0-L2-TSK-FS 可为规则抽取重要子空间特征, 规则复杂度 ψ 分别降低至 32 和 47, 增强了规则的简洁性和解释性。

为进一步验证 FSC-0-L2-TSK-FS 所构建的模型解释性增强之特性, 选取 L2-TSK-FS 作为对比算法, 选取 Breast 数据作为对比数据。表 4 为 L2-TSK-FS 达到最优时为每条规则抽取的特征及其高斯隶属函数表示和结论部分描述; 表 5 为 FSC-0-L2-

表3 UCI 数据实验算法性能对比

	Heart			Breast		
	最优参数	J_{cla} (%)	ψ	最优参数	J_{cla} (%)	ψ
L2-TSK	$k = 7, h = 1, \tau = 10^{-3}$	84.07	280	$k = 7, h = 1, \tau = 10^{-3}$	96.86	217
IQP	$k = 3, h = 10, \tau = 1$	83.70	120	$k = 3, h = 10, \tau = 1$	96.57	93
LSSLI	$k = 3, h = 10^{-2}, \tau = 100$	84.07	120	$k = 3, h = 10^{-2}, \tau = 100$	96.71	93
FSC-0-L2-TSK	$k = 2, \gamma = 0.988, \tau = 10^{-1},$ $h = 10, r = 50$	83.91	32	$k = 3, \gamma = 0.9, \tau = 10^{-1},$ $h = 10, r = 10$	96.60	47

表4 L2-TSK-FS 为每条规则抽取的特征及其高斯隶属函数表示(Breast)

规则	所抽取的特征及其高斯隶属函数表示以及结论部分的线性表达										
R_1											$y_1 = -0.0577 + 0.1754x_1 - 0.0852x_2 + 0.1103x_3 - 0.0770x_4 - 0.1400x_5 + 0.1316x_6 + 0.1357x_7 - 0.1232x_8 - 0.0633x_9 + 1.5035x_{10}$
R_2											$y_1 = -0.7086 + 0.0063x_1 + 4.5683x_2 - 0.0039x_3 + 0.0070x_4 - 0.0388x_5 - 0.0766x_6 + 0.0401x_7 + 0.0234x_8 + 0.0072x_9 + 1.3814x_{10}$
\vdots											\vdots
R_7											$y_1 = -0.0079 + 0.1348x_1 - 0.0692x_2 + 0.1793x_3 + 0.0707x_4 - 0.0384x_5 + 0.1628x_6 + 0.0389x_7 + 0.2163x_8 - 0.0861x_9 + 1.4422x_{10}$

表5 FSC-0-L2-TSK-FS 为每条规则抽取的特征及其高斯隶属函数表示(Breast)

规则	所抽取的特征及其高斯隶属函数表示以及结论部分的线性表达						
R_1							
	$y_1=1.8443$						
R_2							
	$y_1=0.4506$						
R_7							
	$y_1=1.4761$						

TSK-FS 达到最优时为每条规则抽取的特征及其高斯隶属函数表示和结论部分描述。通过对比可得出结论：(1)规则前件。L2-TSK-FS 为每条规则选用全部的特征，每一特征对应于一个模糊子集。对于模糊逻辑系统来说，其解释性主要来自于模糊规则中的模糊集可对应于人类的语言描述。L2-TSK-FS 模型因规则中含有的模糊子集太多，导致模型的解释性下降。FSC-0-L2-TSK-FS 为每条规则抽取不同的特征，并且不同规则对应不同的特征子空间。这样可使获取的模糊规则更加易于语言描述，也使得模糊规则从不同的视角对问题进行推理，增强了模型的简洁性和解释性。(2)规则后件。L2-TSK-FS 的后件为特征向量的线性组合形式，使得后件繁杂，简洁性下降；L2-TSK-FS 的规则后件为常数，增强了规则后件的简洁性。

5 结束语

本文通过利用 FSC 聚类中的特征权值矩阵，构造出一种 FSC-0-L2-TSK-FS 模型。新方法能运用较少的重要特征来构建模糊规则，每条模糊规则对应于数据的特征子空间；具有较强的去噪能力，能够去除掉对系统稳定性干扰强的特征，保证了系统的鲁棒性；能准确把握住数据中的重要特征，所构建的模糊规则对应于不同的特征子空间，推理的侧重点不相同，更加符合人类的推理机制，增强了模型的解释性。

目前，FSC-0-L2-TSK-FS 模型仍具有一定的缺陷。若给出的训练数据子空间特征不明显，会因丢失特征而导致系统泛化性能下降。如何从子空间特征不明显的高维数据中获取具有更强解释性的模糊系统是今后将进行的工作。

参考文献

- [1] Zadeh L A. Fuzzy sets[J]. *Information and Control*, 1965, 8(3): 338-353.
- [2] 李奕, 吴小俊. 基于监督学习的 Takagi Sugeno Kang 模糊系统图像融合方法研究[J]. *电子与信息学报*, 2014, 36(5): 1126-1132.
Li Yi and Wu Xiao-jun. A novel image fusion method using the Takagi Sugeno Kang fuzzy system based on supervised learning[J]. *Journal of Electronics & Information Technology*, 2014, 36(5): 1126-1132.
- [3] 宋恒, 王晨, 马时平, 等. 基于非单点模糊支持向量机的判决反馈均衡器[J]. *电子与信息学报*, 2008, 30(1): 117-120.
Song Heng, Wang Chen, Ma Shi-ping, et al.. A decision feedback equalizer based on non-singleton fuzzy support vector machine[J]. *Journal of Electronics & Information Technology*, 2008, 30(1): 117-120.
- [4] Lughofer E. On-line assurance of interpretability criteria in evolving fuzzy systems—achievements, new concepts and open issues[J]. *Information Sciences*, 2013(251): 22-46.
- [5] Riid A and Rüstern E. Adaptability, interpretability and rule weights in fuzzy rule-based systems[J]. *Information Sciences*, 2014(257): 301-312.
- [6] Thong N T and Son L H. HIFCF: an effective hybrid model between picture fuzzy clustering and intuitionistic fuzzy recommender systems for medical diagnosis[J]. *Expert System with Application*, 2015, 42(7): 3682-3701.
- [7] Sanz J A, Galar M, Jurio A, et al.. Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system[J]. *Applied Soft Computing*, 2014(20): 103-111.
- [8] Takagi T and Sugeno M. Fuzzy identification of systems and

- its applications to modeling and control[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1985(1): 116-132.
- [9] Sugeno M and Kang G T. Structure identification of fuzzy model[J]. *Fuzzy Sets and Systems*, 1988, 28(1): 15-33.
- [10] Jiang Yi-zhang, Chung Fu-lai, Ishibuchi H, *et al.*. Multitask TSK fuzzy system modeling by mining intertask common hidden structure[J]. *IEEE Transactions on Cybernetics*, 2015, 45(3): 548-561.
- [11] Fadali S and Jafarzadeh S. TSK observers for discrete type-1 and type-2 fuzzy systems[J]. *IEEE Transactions on Fuzzy Systems*, 2014, 22(2): 451-458.
- [12] Chung Fu-lai, Deng Zhao-hong, and Wang Shi-tong. From minimum enclosing ball to fast fuzzy inference system training on large datasets[J]. *IEEE Transactions on Fuzzy Systems*, 2009, 17(1): 173-184.
- [13] Mamdani E H. Application of fuzzy logic to approximate reasoning using linguistic synthesis[J]. *IEEE Transactions on Computers*, 1977, 100(12): 1182-1191.
- [14] Azeem M F, Hanmandlu M, and Ahmad N. Generalization of adaptive neuro-fuzzy inference systems[J]. *IEEE Transactions on Neural Networks*, 2000, 11(6): 1332-1346.
- [15] Gan Guo-jun and Wu Jian-hong. A convergence theorem for the fuzzy subspace clustering (FSC) algorithm[J]. *Pattern Recognition*, 2008, 41(6): 1939-1947.
- [16] Deng Zhao-hong, Choi Kup-sze, Chung Fu-lai, *et al.*. Enhanced soft subspace clustering integrating within-cluster and between-cluster information[J]. *Pattern Recognition*, 2010, 43(3): 767-781.
- [17] Leski J M. TSK-fuzzy modeling based on ϵ -insensitive learning[J]. *IEEE Transactions on Fuzzy Systems*, 2005, 13(2): 181-193.
- [18] Deng Zhao-hong, Choi Kup-sze, Chung Fu-lai, *et al.*. Scalable TSK fuzzy modeling for very large datasets using minimal-enclosing-ball approximation[J]. *IEEE Transactions on Fuzzy Systems*, 2011, 19(2): 210-226.
- [19] Juang Chia-feng and Chiang Loa. Zero-order TSK-type fuzzy system learning using a two-phase swarm intelligence algorithm[J]. *Fuzzy Sets and Systems*, 2008, 159(21): 2910-2926.
- [20] Tsang I W, Kwok J T Y, and Zurada J M. Generalized core vector machines[J]. *IEEE Transactions on Neural Networks*, 2006, 17(5): 1126-1140.
- 邓赵红: 男, 1981 年生, 副教授, 硕士生导师, 研究方向为计算智能和模糊计算.
- 张江滨: 男, 1990 年生, 硕士生, 研究方向为计算智能和模糊计算.
- 蒋亦樟: 男, 1988 年生, 博士生, 研究方向为人工智能和模式识别.
- 史奕中: 男, 1970 年生, 博士生, 研究方向为人工智能和模式识别.
- 王士同: 男, 1964 年生, 教授, 博士生导师, 研究方向为人工智能和模式识别.