

具备视角协同学习能力的多视角 TSK 型模糊系统

程 旻 顾晓清 蒋亦樟* 杭文龙 钱鹏江 王士同
(江南大学数字媒体学院 无锡 214122)

摘要: 传统模糊系统建模方法本质上是一种单视角学习模式, 面向适合多视角处理的场景时, 它们通常只能将每一视角割裂开来独立建模, 这导致其所得系统泛化性能往往不令人满意。针对此缺陷, 该文探讨具备多视角学习能力的模糊系统建模方法。为此, 基于经典的 L2 型 TSK 模糊系统, 通过引入具备多视角学习能力的协同学习项, 该文提出了核心的多视角 TSK 型模糊系统(MV-TSK-FS)建模方法。MV-TSK-FS 不仅能有效地利用各视角不同特征构成的独立样本信息, 还能充分地利用各视角间由于相互关联而存在内在信息, 以最终达到提高系统泛化性能的效果。在模拟数据集与真实数据集上的实验结果验证了较之于传统单视角模糊建模方法该多视角模糊系统有着更好的泛化性和适用性。

关键词: 多视角学习; 协同学习; 模糊建模; TSK 型模糊系统

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2016)08-2054-08

DOI: 10.11999/JEIT151209

Multi-view TSK Fuzzy System via Collaborative Learning

CHENG Yang GU Xiaoping JIANG Yizhang HANG Wenlong
QIAN Pengjiang WANG Shitong
(School of Digital Media, Jiangnan University, Wuxi 214122, China)

Abstract: Conventional fuzzy system modeling methods essentially belong to the single-view learning modality. In multi-view-oriented data scenarios, they can only cope with each view separately, which is prone to incurring their unsatisfactory generalization performance. In response to such problem, the fuzzy system modeling method with the ability of multi-view learning is pursued. To this end, based on the classic L2 norm Takagi-Sugeno-Kang (TSK) fuzzy system, by means of the collaborative learning items qualified for multi-view learning, the core Multi-View TSK Fuzzy System (MV-TSK-FS) modeling method is presented. MV-TSK-FS can not only effectively utilize the independent components composed of the characteristics affiliated to each view, but also take full advantage of the potential information occurred by the interrelated effects among views, which eventually facilitates its relatively strong generalization ability. The experimental results performed on both synthetic and real-life datasets indicate that, compared with some traditional single-view methods, this propounded multi-view fuzzy modeling system owns preferable applicability as well as generalization.

Key words: Multi-view learning; Collaborative learning; Fuzzy modeling; Takagi-Sugeno-Kang (TSK) type fuzzy system

1 引言

经典的哲学理论告诉我们多视角学习方法较之

传统的单视角学习策略, 在看待问题方面更为全面与客观, 并且大量的实际应用, 如多视角文本分类及多视角行为识别等均说明了该方法的有效性。近年来多视角学习在机器学习领域得到了迅猛发展, 具体工作可归纳为: (1)分类领域的多视角学习^[1-3]; (2)无监督学习, 特别是聚类领域的多视角学习^[4-6]; (3)回归领域的多视角学习^[7]。其中, 在分类和聚类领域的多视角学习成果已较为丰富, 但相对而言, 针对回归领域的多视角学习理论研究则较少。然而, 在回归领域多视角学习方式同样具有很好的应用前景及研究价值。以食品的发酵过程为例^[8], 在对同一食品的发酵过程进行检测时, 我们可从不同的角度,

收稿日期: 2015-10-29; 改回日期: 2016-03-15; 网络出版: 2016-05-05

*通信作者: 蒋亦樟 jyz0512@163.com

基金项目: 国家自然科学基金(61300151), 江苏省自然科学基金(BK20130155), 江苏省产学研前瞻性联合研究项目(BY2013015-02), 中央高校基本科研业务费专项资金资助重点项目(JUSRP51614A)

Foundation Items: The National Natural Science Foundation of China (61300151), The Natural Science Foundation of Jiangsu Province (BK20130155), The R&D Frontier Grant of Jiangsu Province (BY2013015-02), The Fundamental Research Funds for the Central Universities (JUSRP51614A)

如风速或搅拌速度, 来记录食品的发酵指标参数值(即输出值)。因角度不同从而导致输入空间构成不同(独立性), 但又因观测对象相同, 这使得各视角的最终输出是一致的(关联性), 根据上述性质即可得到一个多视角样本集。同时, 由于发酵过程是一个随着时间变化的连续性函数, 因此对其的预测可视为回归问题。上述需求, 促使我们应当在回归领域开展多视角学习研究。

模糊系统作为智能计算领域重要的研究分支, 其独特的可解释性以及强大的学习能力, 得到了较广泛的应用^[9,10]。传统的模糊系统建模方法均采用单视角学习方式, 其在面对多视角问题时, 一般的处理方式是人为地将多视角样本集根据不同的视角划分成若干个独立的样本, 然后根据每一视角的样本构建出专属于该视角的模糊系统。这种模糊系统存在各视角模糊系统泛化能力低且各视角间的系统泛化性能良莠不齐的现象。造成上述影响的主要原因在于单视角的建模方法只是简单地根据每一视角的样本的独立性特征构建出了适应于该视角的模糊系统而忽视了各视角样本之间的关联关系, 使得多视角样本集的特征无法得到充分的利用, 从而导致了上述现象。

针对上述挑战, 本文提出了一种具有多视角学习能力的模糊系统建模方法。鉴于 Takagi-Sugeno-Kang(TSK)^[11]型模糊系统在面对中小规模数据建模时优越的性能, 本文将其作为具体研究对象来构建多视角模糊系统。在此模型的基础上, 通过引入多视角学习机制, 构造出了新的目标函数, 使得该模糊系统不仅能够考虑各视角的独立信息还能够考虑各视角间的关联信息, 进而达到提升系统泛化能力及稳定各视角间系统性能的目的。各行各业科学技术的发展正在日益变化, 看待问题的方法及角度也越来越丰富, 这使得本文所提出的多视角 TSK 型模糊系统具有较好的应用前景。

2 传统的 TSK 型模糊系统

经典模糊系统模型可分为以下 3 类: TSK 型模型^[11], Mamdani-Larsen(ML)型模型^[12]及广义模糊模型(Generalized Fuzzy Model, GFM)^[13]。其中, TSK 型模型应用最为广泛^[11]。在本节中我们将对传统 TSK 型模糊系统及其对应的学习策略进行简要的回顾。

2.1 TSK 型模糊系统的基本概念

TSK 型模糊系统的模糊推理规则可表达为

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

其中, 每条规则都有与之对应的输入向量 $\mathbf{x} = [x_1, x_2, \dots, x_d]^T$, 并把输入空间的模糊集 $A^k \subset R^d$ 映射到输出空间的模糊集 $f^k(x) \subset R$ 。此处 A_i^k 表示第 i 维输入在第 k 条模糊规则中对应的模糊子集, \wedge 表示模糊合取操作。

TSK 型模糊系统在去模糊化后其最终输出可表示为

$$y^0 = \sum_{k=1}^K \frac{\mu^k(x)}{\sum_{k'=1}^K \mu^{k'}(x)} f^k(x) = \sum_{k=1}^K \tilde{\mu}^k(x) f^k(x) \quad (2)$$

对于式(2)中的 $\mu^k(x)$ 与 $\tilde{\mu}^k(x)$ 之间的关系通过模糊集合 A^k 联系, 其两者之间的关系为

$$\mu^k(x) = \prod_{i=1}^d \mu_{A_i^k}(x_i) \quad (2a)$$

$$\tilde{\mu}^k(x) = \mu^k(x) / \sum_{k'=1}^K \mu^{k'}(x) \quad (2b)$$

通过引入高斯函数后, 式(2a)中的 $\mu_{A_i^k}(x_i)$ 部分可进一步表示为

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

其中, 参数 c_i^k 以及 δ_i^k 都可通过经典的聚类算法获取。本文以采用模糊 C 均值聚类(Fuzzy C-Means clustering, FCM)^[14]算法为例, c_i^k 以及 δ_i^k 可通过式(2d), 式(2e)得到

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

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

式中, u_{jk} 即通过 FCM 算法得到的输入向量 $\mathbf{x}_j = [x_{j1}, x_{j2}, \dots, x_{jd}]^T$ 隶属于第 k 类的隶属度。 h 为可人工调整的尺度参数。

根据文献[8]描述, 可令

$$\mathbf{x}_e = (\mathbf{1}, \mathbf{x}^T)^T \quad (3a)$$

$$\tilde{\mathbf{x}}^k = \tilde{\mu}^k(\mathbf{x}) \mathbf{x}_e \quad (3b)$$

$$\mathbf{x}_g = \left((\tilde{\mathbf{x}}^1)^T, (\tilde{\mathbf{x}}^2)^T, \dots, (\tilde{\mathbf{x}}^K)^T \right)^T \quad (3c)$$

$$\mathbf{p}^k = (p_0^k, p_1^k, \dots, p_d^k)^T \quad (3d)$$

$$\mathbf{p}_g = \left((\mathbf{p}^1)^T, (\mathbf{p}^2)^T, \dots, (\mathbf{p}^K)^T \right)^T \quad (3e)$$

根据式(3a)~式(3e)可将式(2)改写为

$$y^0 = \mathbf{p}_g^T \mathbf{x}_g \quad (3f)$$

通过上述描述, 可知 TSK 型模糊系统的参数学习问题最终转化为了线性回归模型^[8]。

2.2 基于 ε -不敏感损失函数的 L2 型 TSK 型模糊系统

TSK 模型基于 ε -不敏感损失函数的目标函数式一般被定义为

$$\min_{\mathbf{p}_g} L = \sum_{i=1}^N |y_i^0 - y_i|_{\varepsilon} = \sum_{i=1}^N |\mathbf{p}_g^T \mathbf{x}_{g_i} - y_i|_{\varepsilon} \quad (4)$$

$$\text{s.t.} \begin{cases} y_i - \mathbf{p}_g^T \mathbf{x}_{g_i} < \varepsilon + \xi_i^+ \\ \mathbf{p}_g^T \mathbf{x}_{g_i} - y_i < \varepsilon + \xi_i^- \end{cases}, \quad \forall i$$

通过引入松弛变量 ξ_i^+ 及 ξ_i^- , 并根据二次规划理论, 式(4)可转化为经典的二次规划问题, 其具体形式为

$$\min_{\mathbf{p}_g, \xi^+, \xi^-, \varepsilon} L(\mathbf{p}_g, \xi^+, \xi^-, \varepsilon) = \frac{1}{\tau} \cdot \frac{1}{N} \sum_{i=1}^N \left((\xi_i^+)^2 + (\xi_i^-)^2 \right) + \frac{1}{2} (\mathbf{p}_g^T \mathbf{p}_g) + \frac{2}{\tau} \cdot \varepsilon \quad (5)$$

$$\text{s.t.} \begin{cases} y_i - \mathbf{p}_g^T \mathbf{x}_{g_i} < \varepsilon + \xi_i^+ \\ \mathbf{p}_g^T \mathbf{x}_{g_i} - y_i < \varepsilon + \xi_i^- \end{cases}, \quad \forall i$$

式(5)引入拉格朗日乘数项, 可将式(4)化为式(6)的QP问题:

$$\arg \max_{\tilde{\alpha}} -\tilde{\alpha}^T \tilde{\mathbf{K}} \tilde{\alpha} + \tilde{\alpha}^T \tilde{\beta} \quad (6)$$

$$\text{s.t.} \quad \tilde{\alpha}^T \mathbf{1} = 1, \quad \alpha_i \geq 0, \quad \forall i$$

其中, $\tilde{\mathbf{K}} = [\tilde{k}_{ij}]_{2N \times 2N}$, $\tilde{k}_{ij} = \mathbf{z}_i^T \mathbf{z}_j + ((N\tau)/2)\delta_{ij}$, 且 $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$.

通过对式(5)的求解, 最终可得到参数 $(\lambda^+)^*$, $(\lambda^-)^*$ 的最优解, 并利用式(7a)~式(7d)最终求得 TSK 型模糊系统的泛化项 \mathbf{p}_g 的最优解.

$$\mathbf{p}_g = \frac{2}{\tau} \sum_{i=1}^N (\lambda_i^+ - \lambda_i^-) \mathbf{x}_{g_i} \quad (7a)$$

$$\xi_i^+ = N\lambda_i^+ \quad (7b)$$

$$\xi_i^- = N\lambda_i^- \quad (7c)$$

$$\varepsilon = \sum_{i=1}^N (\lambda_i^+ - \lambda_i^-) y_i - \frac{N}{2} \sum_{i=1}^N \left((\lambda_i^+)^2 + (\lambda_i^-)^2 \right) - \frac{1}{\tau} \sum_{i=1}^N \sum_{j=1}^N (\lambda_i^+ - \lambda_i^-) (\lambda_j^+ - \lambda_j^-) \cdot \mathbf{x}_{g_i}^T \mathbf{x}_{g_j} \quad (7d)$$

根据式(2d)及式(2e)所获之模糊系统前件参数并结合式(7a)所获得的模糊系统后件参数即可生成对应的 TSK 型模糊系统.

3 多视角 TSK 型模糊系统学习框架

根据第2节所提之经典 TSK 型模糊系统建模方法可知传统的系统建模策略均是针对单视角场景设计的, 当该种经典建模方法在面对多视角建模场景时, 较为可行的建模策略则是针对不同的视角利用

该视角下的样本构建出与之对应的模糊系统, 具体方法如图1所示. 虽然上述策略为单视角建模方法在多视角场景下的应用提供了一种可行的方案, 但是其针对每一视角较为独立的建模方式在一定程度上破坏了多视角数据集之间的关联关系, 这通常会造造成不同视角所获得之模糊系统泛化能力存在良莠不齐的现象.

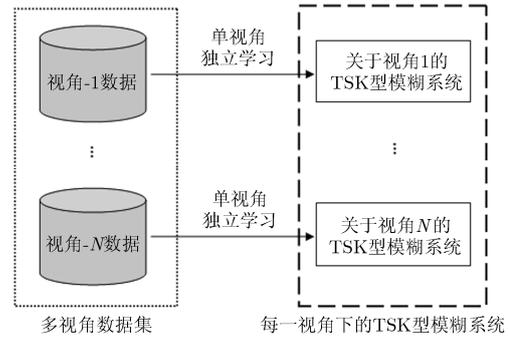


图1 面对多视角学习场景时传统的单视角 TSK 型模糊系统学习框架展示

为了使得传统的单视角 TSK 型模糊系统建模方法具备多视角学习的能力进而得到多视角 TSK 型模糊系统(MV-TSK-FS), 本文给出了如图2所示的建模策略. 观察图2可知, 该建模方法较之传统单视角的建模方法而言, 其不再将各视角下的数据样本孤立开来独立的模糊系统构建, 而是通过多视角学习机制来训练不同视角下的模糊系统使得最终所获之各视角模糊系统泛化能力得到提升并性能趋于一致. 具体的多视角 TSK 型模糊系统的模型构建方法将于第4节进行详细阐述.

4 多视角 TSK 型模糊系统建模方法

为改进单视角 TSK 型模糊系统, 使其具备多视角学习能力, 本节将在第2节所提及的 TSK 型模糊系统建模方式上应用多视角的学习原理构建多视角 TSK 型模糊系统(Multi-View TSK type Fuzzy

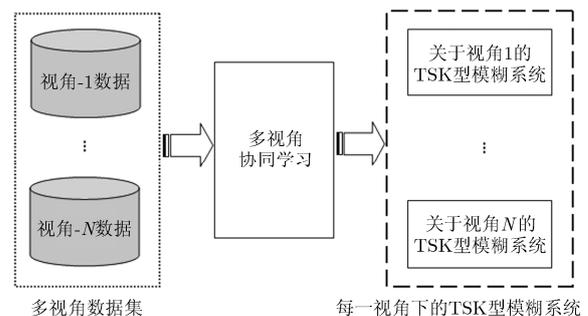


图2 面对多视角学习场景时新型的单视角 TSK 模糊系统学习框架展示

System, MV-TSK-FS)。为了公式简洁性, 我们仅以 2 视角为例进行多视角模型构建。

4.1 具备多视角协同学习能力的目标函数构造

为了使得经典的 L2 型惩罚函数构造的 TSK 型模糊系统具备多视角学习的能力, 本小节以双视角作为具体的研究场景探讨了多视角 TSK 型模糊系统的目标函数构造。在构建多视角 TSK 型模糊系统时, 为了能够充分利用不同视角数据样本的特有特征(独立性信息)及各视角间样本内在的关联关系(关联性信息), 本文给出了如下的具备多视角学习能力的目标函数式:

$$\min_{\mathbf{p}_{g,A}, \xi_A^+, \xi_A^-, \varepsilon_A} g_A(\mathbf{p}_{g,A}, \xi_A^+, \xi_A^-, \varepsilon_A) + g_B(\mathbf{p}_{g,B}, \xi_B^+, \xi_B^-, \varepsilon_B) + \Psi_S(\boldsymbol{\eta}^+, \boldsymbol{\eta}^-, \varepsilon_S) \quad (8)$$

$$\text{s.t.} \left\{ \begin{array}{l} y_{i,A} - \mathbf{p}_{g,A}^T \mathbf{x}_{g,i,A} < \varepsilon_A + \xi_{i,A}^+ \\ \mathbf{p}_{g,A}^T \mathbf{x}_{g,i,A} - y_{i,A} < \varepsilon_A + \xi_{i,A}^- \\ y_{i,B} - \mathbf{p}_{g,B}^T \mathbf{x}_{g,i,B} < \varepsilon_B + \xi_{i,B}^+ \\ \mathbf{p}_{g,B}^T \mathbf{x}_{g,i,B} - y_{i,B} < \varepsilon_B + \xi_{i,B}^- \\ \mathbf{p}_{g,B}^T \mathbf{x}_{g,i,B} - \mathbf{p}_{g,A}^T \mathbf{x}_{g,i,A} < \varepsilon_S + \eta_i^+ \\ \mathbf{p}_{g,A}^T \mathbf{x}_{g,i,A} - \mathbf{p}_{g,B}^T \mathbf{x}_{g,i,B} < \varepsilon_S + \eta_i^- \end{array} \right., \quad \forall i$$

其中,

$$g_A(\mathbf{p}_{g,A}, \xi_A^+, \xi_A^-, \varepsilon_A) = \frac{1}{N\tau_A} \sum_{i=1}^N [(\xi_{i,A}^+)^2 + (\xi_{i,A}^-)^2] + \frac{1}{2} (\mathbf{p}_{g,A}^T \mathbf{p}_{g,A}) + \frac{2}{\tau_A} \cdot \varepsilon_A \quad (8a)$$

$$g_B(\mathbf{p}_{g,B}, \xi_B^+, \xi_B^-, \varepsilon_B) = \frac{1}{N\tau_B} \sum_{i=1}^N [(\xi_{i,B}^+)^2 + (\xi_{i,B}^-)^2] + \frac{1}{2} (\mathbf{p}_{g,B}^T \mathbf{p}_{g,B}) + \frac{2}{\tau_B} \cdot \varepsilon_B \quad (8b)$$

$$\Psi_S(\boldsymbol{\eta}^+, \boldsymbol{\eta}^-, \varepsilon_S) = \frac{1}{N\tau_S} \sum_{i=1}^N [(\eta_i^+)^2 + (\eta_i^-)^2] + \frac{2}{\tau_S} \cdot \varepsilon_S \quad (8c)$$

观察式(8)可发现该式由式(8a)~式(8c) 3 项构成, 其中式(8a)及式(8b)分别为根据 A 和 B 视角样本特征构建的 TSK 型模糊系统, 上述两项保证了所提方法能够充分利用不同视角下数据样本特征的差异性构建出属于该视角特征的专有模糊系统。但仅考虑视角间的样本特征差异性(独立性信息)对于整个多视角 TSK 型模糊系统的构建是不够的, 为增强多视角学习的能力, 所提方法进一步考虑了视角间的关联关系, 即各视角决策(输出)的一致性。根据上述特性本文构造了具备协同学习能力的式(8c), 该式可使得视角 A 与视角 B 进行协同学习, 最终达到各视角所对应的模糊系统之输出趋于一致从而避免因视角特征的变化而造成各视角模糊系统的泛化

性能的差异。需要说明的是正则化参数 $\tau_A > 0$, $\tau_B > 0$, $\tau_S > 0$ 控制了回归模型的复杂度及误差程度, 其数值可人工设定, 亦可通过交叉验证策略^[15] 选择一个合适的值。

4.2 MV-TSK-FS 参数学习规则

为求解式(8), 根据拉格朗日优化理论, 引入相关拉格朗日乘子, 可将式(8)转换为式(9):

$$\arg \max_{\mathbf{v}} \left. \begin{array}{l} -\frac{1}{2} \mathbf{v}^T \mathbf{K} \mathbf{v} + \mathbf{v}^T \boldsymbol{\beta} \\ \text{s.t.} \quad \mathbf{v}^T \mathbf{L}_A = 2/\tau_A, \mathbf{L}_A = \underbrace{[1, \dots, 1]}_{2N}, \underbrace{[0, \dots, 0]}_{2N}, \underbrace{[0, \dots, 0]}_{2N} \\ \mathbf{v}^T \mathbf{L}_B = 2/\tau_B, \mathbf{L}_B = \underbrace{[0, \dots, 0]}_{2N}, \underbrace{[1, \dots, 1]}_{2N}, \underbrace{[0, \dots, 0]}_{2N} \\ \mathbf{v}^T \mathbf{L}_S = 2/\tau_S, \mathbf{L}_S = \underbrace{[0, \dots, 0]}_{2N}, \underbrace{[0, \dots, 0]}_{2N}, \underbrace{[1, \dots, 1]}_{2N} \\ v_i \geq 0, \quad \forall i \end{array} \right\} \quad (9)$$

其中,

$$\tilde{\mathbf{K}}_A = [\tilde{k}_{ij,A}]_{2N \times 2N}, \tilde{k}_{ij,A} = \mathbf{z}_{i,A}^T \mathbf{z}_{j,A} + 0.5 \cdot N\tau_A \delta_{ij,A}, \quad (9a)$$

$$\delta_{ij,A} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$\tilde{\mathbf{K}}_B = [\tilde{k}_{ij,B}]_{2N \times 2N}, \tilde{k}_{ij,B} = \mathbf{z}_{i,B}^T \mathbf{z}_{j,B} + 0.5 \cdot N\tau_B \delta_{ij,B}, \quad (9b)$$

$$\delta_{ij,B} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$\tilde{\mathbf{K}}_S = [\tilde{k}_{ij,S}]_{2N \times 2N}, \tilde{k}_{ij,S} = -3\mathbf{z}_{i,B}^T \mathbf{z}_{j,B} + \mathbf{z}_{i,A}^T \mathbf{z}_{j,A} + 0.5 \cdot N\tau_S \delta_{ij,S}, \quad (9c)$$

$$\delta_{ij,S} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$\tilde{\mathbf{K}}_{AS} = [\tilde{k}_{ij,AS}]_{2N \times 2N}, \tilde{k}_{ij,AS} = \mathbf{z}_{i,A}^T \mathbf{z}_{j,A} \quad (9d)$$

$$\tilde{\mathbf{K}}_{BS} = [\tilde{k}_{ij,BS}]_{2N \times 2N}, \tilde{k}_{ij,BS} = \mathbf{z}_{i,B}^T \mathbf{z}_{j,B} \quad (9e)$$

$$\mathbf{K} = \begin{pmatrix} \tilde{\mathbf{K}}_A & 0 & \tilde{\mathbf{K}}_{AS} \\ 0 & \tilde{\mathbf{K}}_B & -\tilde{\mathbf{K}}_{BS} \\ \tilde{\mathbf{K}}_{AS} & -\tilde{\mathbf{K}}_{BS} & \tilde{\mathbf{K}}_S \end{pmatrix} \quad (9f)$$

至此, 式(8)被化为了如式(9)的经典 QP 问题, 本文将采用文献[8]中的 QP 求解方法来求解式(9)。

通过 QP 优化后, 即可得到式(9)对应的优化后的求解参数 $(\boldsymbol{\lambda}_A^+)^*$, $(\boldsymbol{\lambda}_A^-)^*$, $(\boldsymbol{\lambda}_B^+)^*$, $(\boldsymbol{\lambda}_B^-)^*$, $(\boldsymbol{\beta}_S^+)^*$, $(\boldsymbol{\beta}_S^-)^*$, 利用这些优化后的参数, 我们可进一步地获取对应于各视角下的已优化 TSK 型模糊系统后件参数为

$$(\mathbf{p}_{g,A})^* = \sum_{i=1}^N \left((\lambda_{i,A}^+)^* - (\lambda_{i,A}^-)^* \right) \mathbf{x}_{g,i,A} + \sum_{i=1}^N \left((\beta_i^+)^* - (\beta_i^-)^* \right) \mathbf{x}_{g,i,A} \quad (10a)$$

$$\begin{aligned}
(\mathbf{p}_{g,B})^* &= \sum_{i=1}^N \left((\lambda_{i,B}^+)^* - (\lambda_{i,B}^-)^* \right) \mathbf{x}_{g_i,B} \\
&+ \sum_{i=1}^N \left((\beta_i^+)^* - (\beta_i^-)^* \right) \mathbf{x}_{g_i,B} \quad (10b)
\end{aligned}$$

式(10a)为视角 A 对应的 TSK 型模糊系统的后件参数, 而式(10b)则为视角 B 对应的 TSK 型模糊系统的后件参数。根据所获之后件参数结合式(2d)及式(2e)所得到的模糊系统前件参数即可生成一个多视角 TSK 型模糊系统。具体的多视角 TSK 型模糊系统(MV-TSK-FS)的学习算法如表 1 所示。

表 1 MV-TSK-FS 算法

算法 1 MV-TSK-FS 算法	
步骤 1	设置正则化参数 τ_A, τ_B, τ_S 以及为每一视角设定模糊规则数 K_A, K_B (一般, 取 $K_A = K_B$);
步骤 2	利用 FCM 算法或其他具备样本空间划分能力的算法得到各视角样本所对应于回归模型的数据集 $D_A = \{\mathbf{x}_{g_i,A}, y_{i,A}\}, D_B = \{\mathbf{x}_{g_i,B}, y_{i,B}\}, i = 1, 2, \dots, N$;
步骤 3	利用式(9)求解得到拉格朗日乘子 $(\lambda_A^+)^*, (\lambda_A^-)^*, (\lambda_B^+)^*, (\lambda_B^-)^*, (\beta_S^+)^*, (\beta_S^-)^*$;
步骤 4	根据步骤 3 得到对偶问题式(9)的解, 利用式(10a)及式(10b)即可求得 $(\mathbf{p}_{g,A})^*$ 及 $(\mathbf{p}_{g,B})^*$;
步骤 5	根据步骤 4 获取的各视角模型参数即可对应生成满足各视角特征的 TSK 型模糊系统。

由算法 1 可知, MV-TSK-FS 模糊系统的时间复杂度主要由两部分构成: 前件参数的学习(对应步骤 2)和后件参数的学习(对应步骤 3 与步骤 4)。在多视角学习中, MV-TSK-FS 的每个视角对应的模糊系统的前件参数由 FCM 聚类算法获得, 其时间复杂度是 $O(MTNK + MTK)$, 其中 M 为视角的个数, T 为算法迭代次数, N 为样本数, K 为模糊系统规则条数。MV-TSK-FS 的每个视角的后件参数的学习关键在二次规划求极值问题上, 时间复杂度为 $O(N^3)$, 通过使用 SMO 等分解方法处理二次规划问题, 其时间复杂度可降为 $O(N^2)$, 因此, 多视角学习中, MV-TSK-FS 的后件参数的学习的时间复杂度为 $O(MN^2) \sim O(MN^3)$ 之间。此外, 因本文采用的对比方法, 即经典 TSK 模糊系统建模方法^[8]为一种快速建模方法, 其时间复杂度仅为 $O(TNK + TK) + O(N^2)$, 因此所提出的多视角 MV-TSK-FS 模糊系统在时间性能上要劣于单视角 TSK 模糊系统建模方法。究其主要原因在于, MV-TSK-FS 模糊系统建模策略融入了多视角协同学习策略, 而单视角 TSK 模糊系统并不具备多视角学习能力, 因此 MV-TSK-FS 模糊系统更为耗时, 但其泛化性能较之经典单视角 TSK 模糊系统有着较大的提升。

5 实验研究

5.1 实验设置

为对本文所提之多视角 TSK 型模糊系统(MV-TSK-FS)的性能进行验证与评估, 本节将分别采用人工合成的多视角数据集以及真实的谷氨酸发酵多视角数据集进行实验研究。其中合成数据以及谷氨酸真实数据将分别于 5.2.1 节和 5.2.2 节中给出具体描述。此外, 实验中所涉及的一些符号定义如下: D1 和 D2 分别表示为多视角场景下视角 1 和视角 2 的训练数据集, D1_test 和 D2_test 则为测试数据集; [D1;D2]表示为视角 1 及视角 2 训练数据集以特征扩展的方式合并得到的单视角训练数据集, [D1; D2]_test 为测试数据集; TSK-FS(View-D1), TSK-FS(View-D2)及 TSK-FS(View-[D1;D2])分别为单视角 TSK-FS 基于 D1, D2 及[D1;D2]数据集所构之模糊系统; MV-TSK-FS(View-D1)和 MV-TSK-FS(View-D2)则为本文所提之多视角 MV-TSK-FS 建模方法基于 D1 和 D2 所构之模糊系统。

为了对本文的实验结果给出合理的评价, 采用式(11)作为性能评价指标^[8]。

$$J = \sqrt{\frac{\sum_{i=1}^N (y'_i - y_i)^2}{\sum_{i=1}^N (y'_i - \bar{y})^2}} \quad (11)$$

其中, N 为测试样本集尺寸; y_i 为第 n 个测试样本的真实输出; y'_i 表示第 n 个测试样本对应的模糊系统输出; $\bar{y} = 1/N \cdot \sum_{i=1}^N y_i$ 。 J 越小则表示泛化能力越好。

在实验中, 所涉人工设定参数均采用如表 2 所述之优化策略。实验环境: 实验硬件平台为 Intel Pentium×2 CPU, 其主频为 1.6 GHz, 内存为 1 GB。编程环境为 MATLAB 7.0。

5.2 本文方法与经典 L2-TSK-FS 性能比较

5.2.1 合成数据集 为了能够体现多视角数据集各视角特征构成的独立性以及视角之间的关联关系, 本文根据上述的多视角数据的特点分别选用 $\sin(x_1)$ 函数以及 $\cos(x_2)$ 函数来表征两个独立的视角, 并利用 $\sin(x)$ 函数与 $\cos(\pi/2 - x)$ 函数具备相同的函数值的特点来形成两视角之间的关联关系(令 $x_1 = x$ 而 $x_2 = \pi/2 - x$), 这样构成的多视角样本集在各视角下, 虽然输入特征不同(视角间的独立性), 但最终各视角的输出却是一致的(视角间的关联性)。具体的函数定义及相关采样区间的描述可见表 3 所述。具体人工数据实验结果如表 4 和图 3 所示。

根据表 4 和图 3 所示的合成数据集上的实验结果, 我们归纳出如下观察结果:

(1)从表 4 以及图 3(a)和图 3(b)的实验结果看,

表 2 相关实验设置

设置	数据集	传统单视角 TSK-FS 模型	新型多视角 MV-TSK-FS 模型
模糊系统规则数：本实验将在{5,10,20,30,50,100}的区间范围内进行对比分析			
可调参数设置	模拟数据集	正则化参数 τ 的最佳值将在 $\{2^{-6}, 2^{-5}, \dots, 2^5, 2^6\}$ 区间内通过 5 倍交叉验证策略寻优取得。	正则化参数 τ_A, τ_B 及 τ_S 的最佳值将在 $\{2^{-6}, 2^{-5}, \dots, 2^5, 2^6\}$ 区间内通过 5 倍交叉验证策略寻优取得。
	与模拟数据集相似，本实验将在{5,10,20,30}的区间范围内进行对比分析		
	真实数据集	正则化参数 τ 的最佳值将在 $\{2^{-6}, 2^{-5}, \dots, 2^5, 2^6\}$ 区间内通过 5 倍交叉验证策略寻优取得。	正则化参数 τ_A, τ_B 及 τ_S 的最佳值将在 $\{2^{-6}, 2^{-5}, \dots, 2^5, 2^6\}$ 区间内通过 5 倍交叉验证策略寻优取得。
评价指标	(1) J 符号表示为所获模型在测试集上的泛化性能。 (2) Δ 符号表示为不同视角间泛化性能差的绝对值。		

表 3 相关数据源生成函数定义

	定义函数	功能
视角 1 数据集	$Y_1 = f^{\text{View-1}}(x_1) = \sin(x_1); x_1 = x$	用于产生 128 个视角 1 训练数据集(D1)及 64 个测试数据集(D1_test)
视角 2 数据集	$Y_2 = f^{\text{View-2}}(x_2) = \cos(x_2); x_2 = \pi/2 - x$	用于产生 128 个视角 2 训练数据集(D2)及 64 个测试数据集(D2_test)
x 取值范围:	$x \in [-4, 4]$	

表 4 多视角 MV-TSK-FS 建模方法及传统 TSK-FS 建模方法在模拟数据集上的泛化性能比较

规则数(M)	TSK-FS 模型				MV-TSK-FS 模型		
	View(D1)	View(D2)	View(D1;D2)	Δ	View(D1)	View(D2)	Δ
5	0.1179	0.1060	0.1002	0.0119	0.0751	0.0757	0.0006
10	0.0839	0.0878	0.0830	0.0039	0.0176	0.0180	0.0004
20	0.0384	0.0411	0.0521	0.0027	0.0085	0.0091	0.0006
30	0.0332	0.0367	0.0370	0.0035	0.0083	0.0090	0.0007
50	0.0387	0.0360	0.0381	0.0027	0.0052	0.0047	0.0005
100	0.0463	0.0487	0.0454	0.0024	0.0050	0.0049	0.0001

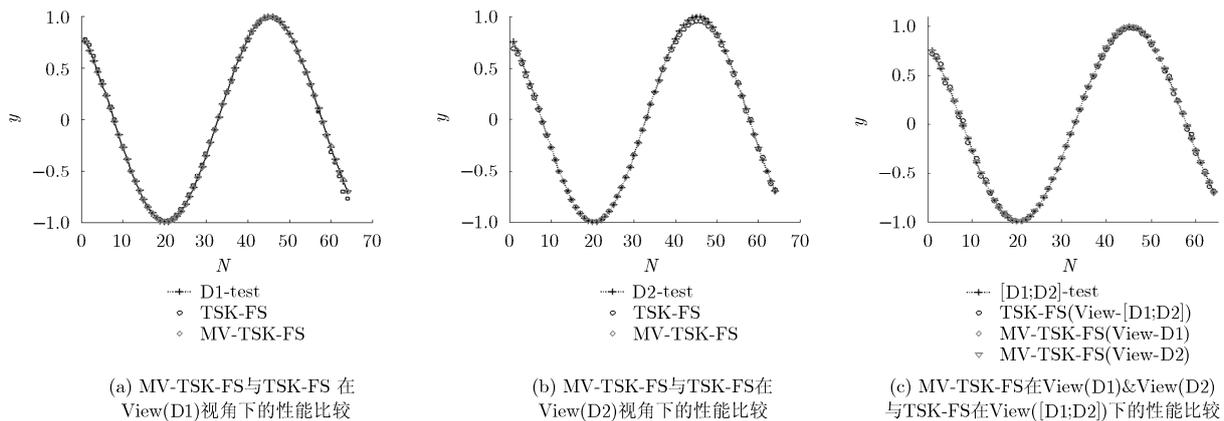


图 3 MV-TSK-FS 建模方法及传统的 TSK-FS 建模方法在模糊规则数 $K = 30$ 的情况下各自的性能比较

无论是 View(D1)视角还是 View(D2)视角的结果，均说明了由本文方法所构之模糊系统性能要明显优于经典单视角 TSK 型模糊系统的性能。这说明本文方法因具备多视角学习的能力较之经典的单视角方法能够更加有效地利用多视角数据集中各视角较为

独立的特征空间信息以及视角与视角之间的关联关系，从而使得本文方法能够更好地提升在各视角上所建之模糊系统的泛化能力。

(2)从图 3(c)的实验结果看，各视角虽然存在一定的关联关系如一致性的输出成分，但视角间的差

异性亦是明显的,如输入成分(输入特征空间)。这种差异性导致经典单视角 TSK 型模糊系统在使用人为构造的单视角样本[D1;D2]进行训练后所构建模型的泛化性能明显劣于任意视角下基于本文方法构建的模型的性能。根据上述的结果说明了仅考虑视角间的关联性而忽视了视角间独立性(存在差异的特征空间)也是不可行的。

(3)进一步观察图 3(a)-图 3(c)以及表 4 的 TSK-FS 性能的第 4 列与 MV-TSK-FS 性能的第 3 列,可以发现经典单视角方法得到的各视角模糊系统的性能指标差明显大于本文方法得到的各视角模糊系统的性能指标差,这说明单视角方法得到的模糊系统性能并不稳定,各视角模糊系统的性能存在良莠不齐的现象,而本文方法由于采用了多视角协同学习的策略,使得其最终生成的各视角方法系统的性能在提升的同时还具备相互接近的稳定状态,这一点从图 3(c)中的“ \diamond ”线和“ ∇ ”线的重合程度可以得到进一步的证实。

5.2.2 真实数据集 本小节利用具备多视角特征的发酵过程数据集^[8]来进行建模仿真实验以进一步验证本文方法与经典 TSK 型模糊系统的性能比较,通过对采集到的工厂发酵数据进行筛选和分析后,本文选择其中具有代表性并且覆盖较为广泛的 21 批共 294 组数据样本作为研究对象,该数据集包含两个视角样本分别为:(1)根据搅拌速度为主特征的视角样本 D1;(2)根据通风指数为主特征的视角样本 D2。最终实验结果如表 5 所示。

从表 5 的实验结果可知,其实验结果与上节中在合成数据的观察结果基本是一致的。具备多视角

学习能力的 TSK 型模糊系统在模型训练时,其能够通过充分利用多视角数据集的独立性成分和关联性成分训练得到有关各视角的模糊系统。对于这些得到的各视角模糊系统,我们无论是从独立的单一视角看,还是从不同视角之间系统的泛化性能关联关系角度分析,所得结论均显示出了本文方法比传统单视角方法在处理多视角问题具有更好的泛化性能,且各视角下的系统性能更为稳定。因此利用本文方法构建的模糊系统可以更加准确地监控发酵过程,并进行相应的预测。

6 结束语

本文针对传统的模糊系统建模方法在处理多视角场景仅利用了视角间的独立性信息而忽视了视角间样本的关联性信息,从而导致各视角所构之模糊系统泛化能力差且性能存在良莠不齐的现象,提出了多视角协同学习的思想并将该方法融入具体的 TSK 型模糊系统中,得到了具备多视角学习能力的多视角 TSK 型模糊系统。该方法能够有效且充分地利用多视角数据集中的独立性成分和视角之间的关联性成分,进而提高各视角下模糊系统的泛化性能并通过协同学习机制进一步稳定各视角模糊系统的性能。通过在模拟数据集以及真实的发酵过程数据集上的仿真实验结果,显示本文方法比传统的单视角建模方法具备更好的适应性。但由于多视角算法的复杂性,在本文的研究工作中仅给出了双视角下的 MV-TSK-FS 的具体建模方法,在今后的工作中将致力于研究出一种推导及计算更为简洁且更适于推广的多视角 TSK 型模糊系统的建模方法,同时也将考虑其他经典模糊系统模型的多视角化构建问题。

表 5 多视角 MV-TSK-FS 建模方法及传统 TSK-FS 建模方法在真实发酵数据集上的泛化性能比较

输出	规则数(M)	TSK-FS 模型				MV-TSK-FS 模型		
		View(D1)	View(D2)	View([D1;D2])	Δ	View(D1)	View(D2)	Δ
S($h+2$)	5	0.1850	0.2136	0.1248	0.0286	0.0937	0.0971	0.0034
	10	0.2759	0.3253	0.1230	0.0494	0.0854	0.0871	0.0017
	20	0.1636	0.1591	0.1064	0.0045	0.0804	0.0815	0.0011
	30	0.0899	0.0988	0.0910	0.0089	0.0788	0.0790	0.0002
X($h+2$)	5	0.4747	0.4966	0.4972	0.0219	0.2721	0.2810	0.0089
	10	0.4904	0.5126	0.4964	0.0222	0.3088	0.3132	0.0044
	20	0.3855	0.5533	0.4432	0.1678	0.2397	0.2301	0.0096
	30	0.4763	0.5945	0.5012	0.1182	0.2890	0.2943	0.0053
P($h+2$)	5	0.0994	0.1893	0.1213	0.0899	0.0555	0.0692	0.0137
	10	0.0858	0.1097	0.0969	0.0239	0.0629	0.0631	0.0002
	20	0.0733	0.1104	0.1088	0.0371	0.0575	0.0608	0.0033
	30	0.0736	0.1002	0.0769	0.0266	0.0476	0.0502	0.0026

参 考 文 献

- [1] LI Guangxia, CHANG Kuiyu, and HOI S C H. Multiview semi-supervised learning with consensus[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2012, 24(11): 2040–2051. doi: 10.1109/TKDE.2011.160.
- [2] HONG Chaoqun, YU Jun, YOU Jane, *et al.* Multi-view ensemble manifold regularization for 3D object recognition [J]. *Information Sciences*, 2015, 320: 395–405. doi: 10.1016/j.ins.2015.03.032.
- [3] ZHANG Shunli, YU Xin, SUI Yao, *et al.* Object tracking with multi-view support vector machines[J]. *IEEE Transactions on Multimedia*, 2015, 17(3): 265–278. doi: 10.1109/TMM.2015.2390044.
- [4] 蒋亦樟, 邓赵红, 王骏, 等. 熵加权多视角协同划分模糊聚类算法[J]. *软件学报*, 2014, 25(10): 2293–2311. doi: 10.13328/j.cnki.jos.004510.
JIANG Yizhang, DENG Zhaohong, WANG Jun, *et al.* Collaborative partition multi-view fuzzy clustering algorithm using entropy weighting[J]. *Journal of Software*, 2014, 25(10): 2293–2311. doi: 10.13328/j.cnki.jos.004510.
- [5] TZORTZIS G F and LIKAS A C. Multiple view clustering using a weighted combination of exemplar-based mixture models[J]. *IEEE Transactions on Neural Network*, 2010, 21(12): 1925–1938. doi: 10.1109/TNN.2010.2081999.
- [6] JIANG Yizhang, CHUNG Fulai, WANG Shitong, *et al.* Collaborative fuzzy clustering from multiple weighted views[J]. *IEEE Transactions on Cybernetics*, 2015, 45(4): 688–701. doi: 10.1109/TCYB.2014.2334595.
- [7] MERUGU S, ROSSET S, and PERLICH C. A new multi-view regression approach with an application to customer wallet estimation[C]. *Proceedings of 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Philadelphia, USA, 2006: 656–661.
- [8] DENG Zhaohong, JIANG Yizhang, CHOI Kupsze, *et al.* Knowledge-leverage-based TSK fuzzy system modeling[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, 24(8): 1200–1212. doi: 10.1109/FUZZ-IEEE.2014.6891544.
- [9] ÅSTRÖM K J and MCAVOY T J. Intelligent control[J]. *Journal of Process Control*, 1992, 2(3): 115–127.
- [10] 邓赵红, 张江滨, 蒋亦樟, 等. 基于模糊子空间聚类的 α 阶 L2 型 TSK 模糊系统[J]. *电子与信息学报*, 2015, 37(9): 2082–2088. doi: 10.11999/JEIT150074.
DENG Zhaohong, ZHANG Jiangbin, JIANG Yizhang, *et al.* Fuzzy subspace clustering based zero-order L2-norm TSK fuzzy system[J]. *Journal of Electronics & Information Technology*, 2015, 37(9): 2082–2088.
- [11] TAKAGI T and SUGENO M. Fuzzy identification of systems and its application to modeling and control[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1985, 15(1): 116–132. doi: 10.1109/TSMC.1985.6313399.
- [12] MAMDANI E H. Application of fuzzy logic to approximate reasoning using linguistic synthesis[J]. *IEEE Transactions on Computers*, 1977, 26(12): 1182–1191. doi: 10.1109/TC.1977.1674779.
- [13] AZEEM M F, HANMANDLU M, and AHMAD N. Generalization of adaptive neural-fuzzy inference systems[J]. *IEEE Transactions on Neural Networks*, 2000, 11(6): 1332–1346. doi: 10.1109/72.883438.
- [14] JIANG Yizhang, CHUNG Fulai, and WANG Shitong. Enhanced fuzzy partitions vs data randomness in FCM[J]. *Journal of Intelligent and Fuzzy Systems*, 2014, 27(4): 1639–1648. doi: 10.3233/IFS-141130.
- [15] ITO K and NAKANO R. Optimizing support vector regression hyperparameters based on cross-validation[C]. *Proceedings of the International Joint Conference on Neural Networks*, Jantzen Beach, Portland, Oregon, 2003: 2077–2082.
- 程 旻: 男, 1991 年生, 博士生, 研究方向为计算智能和模糊计算.
- 顾晓清: 女, 1981 年生, 讲师, 博士生, 研究方向为计算智能和模糊计算.
- 蒋亦樟: 男, 1988 年生, 讲师, 博士, 研究方向为计算智能和模糊计算.
- 杭文龙: 男, 1988 年生, 博士生, 研究方向为人工智能和模式识别.
- 钱鹏江: 男, 1979 年生, 副教授, 硕士生导师, 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.