



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

任务间共享和特有结构分解的多任务TSK模糊系统建模

赵壮壮, 王骏, 潘祥, 邓赵红, 施俊, 王士同

引用本文:

赵壮壮, 王骏, 潘祥, 等. 任务间共享和特有结构分解的多任务TSK模糊系统建模[J]. 智能系统学报, 2021, 16(4): 622–629.

ZHAO Zhuangzhuang, WANG Jun, PAN Xiang, et al. Multi-task TSK fuzzy system modeling based on inter-task common and special structure decomposition[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(4): 622–629.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202007009>

您可能感兴趣的其他文章

面向自闭症辅助诊断的无监督模糊特征学习新方法

A novel unsupervised fuzzy feature learning method for computer-aided diagnosis of autism

智能系统学报. 2019, 14(5): 882–888 <https://dx.doi.org/10.11992/tis.201808005>

引入外部词向量的文本信息网络表示学习

Representation learning using network embedding based on external word vectors

智能系统学报. 2019, 14(5): 1056–1063 <https://dx.doi.org/10.11992/tis.201809037>

基于车内外视觉信息的行人碰撞预警方法

Pedestrian collision warning system based on looking-in and looking-out visual information analysis

智能系统学报. 2019, 14(4): 752–760 <https://dx.doi.org/10.11992/tis.201801016>

基于核心向量机的多任务概念漂移数据快速分类

The core vector machine-based rapid classification of multi-task concept drift dataset

智能系统学报. 2018, 13(6): 935–945 <https://dx.doi.org/10.11992/tis.201712019>

多层递阶融合模糊特征映射的模糊C均值聚类算法

Fuzzy C-means clustering algorithm for multilayered hierarchical fusion fuzzy feature mapping

智能系统学报. 2018, 13(4): 594–601 <https://dx.doi.org/10.11992/tis.201703047>

符号网络的局部标注特征与预测方法

Local labeling features and a prediction method for a signed network

智能系统学报. 2018, 13(3): 437–444 <https://dx.doi.org/10.11992/tis.201710027>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202007009

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20210408.1021.002.html>

任务间共享和特有结构分解的 多任务 TSK 模糊系统建模

赵壮壮¹, 王骏², 潘祥¹, 邓赵红¹, 施俊², 王士同¹

(1. 江南大学人工智能与计算机学院, 江苏 无锡 214122; 2. 上海大学 通信与信息工程学院, 上海 200444)

摘要: 现有的多任务 Takagi-Sugeno-Kang (TSK) 模糊建模方法更注重利用任务间的相关性信息, 而忽略了单个任务的特殊性。针对此问题, 本文提出了一种考虑所有任务之间的共享结构和特有结构的 TSK 模糊系统多任务建模新方法。该方法将后件参数分解为共享参数和特有参数两个分量, 既充分利用了任务间共享信息, 又有效地保留了单个任务的特性。最后, 本文利用增广拉格朗日乘子法 (ALM) 求解该最优化问题。实验结果表明, 该方法比现有的模型获得了更好的表现。

关键词: TSK 模糊系统; 非线性; 多任务; 低秩; 稀疏; 参数分解; 泛化性能; 可解释性

中图分类号: TP181 **文献标志码:** A **文章编号:** 1673-4785(2021)04-0622-08

中文引用格式: 赵壮壮, 王骏, 潘祥, 等. 任务间共享和特有结构分解的多任务 TSK 模糊系统建模 [J]. 智能系统学报, 2021, 16(4): 622-629.

英文引用格式: ZHAO Zhuangzhuang, WANG Jun, PAN Xiang, et al. Multi-task TSK fuzzy system modeling based on inter-task common and special structure decomposition[J]. CAAI transactions on intelligent systems, 2021, 16(4): 622-629.

Multi-task TSK fuzzy system modeling based on inter-task common and special structure decomposition

ZHAO Zhuangzhuang¹, WANG Jun², PAN Xiang¹, DENG Zhaohong¹, SHI Jun², WANG Shitong¹

(1. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China; 2. School of Communication & Information Engineering, Shanghai University, Shanghai 200444, China)

Abstract: Existing Takagi-Sugeno-Kang (TSK) fuzzy system modeling methods pay more attention to the inter-task correlation but ignore the particularity of every single task. To address this issue, this paper proposes a novel multi-task modeling method for TSK fuzzy systems taking common and specific structures across all tasks (MTTSKFS-CS) into consideration. This method decomposes consequent parameters into shared and special ones, which not only takes advantage of the shared information among tasks but also effectively preserves the characteristics of individual tasks. Finally, the study uses the augmented Lagrange multiplier for optimization. The experimental results demonstrate the better performance of the proposed model compared with other existing methods.

Keywords: TSK fuzzy system; nonlinear; multitask; low-rank; sparse; parameter decomposition; generalization performance; interpretability

模糊逻辑和模糊推理被用来描述知识和表达的不确定性。而模糊系统就是由模糊逻辑和模糊推理发展而来的。相比于传统的机器学习模型,

收稿日期: 2020-07-06. 网络出版日期: 2021-04-08.

基金项目: 江苏省自然科学基金项目 (BK20181339); 国家自然科学基金项目 (61602007); 中央高校基础研究经费资助项目 (JUSRP11851).

通信作者: 王骏. E-mail: wangjun_shu@shu.edu.cn.

模糊系统能够更准确地描述和估计现实中不确定的复杂非线性系统模型^[1-4]。近年来, 学者们提出了众多的模糊系统建模方法, 其中 TSK 模糊系统因其能够将非线性系统转化为多个局部线性结构的逼近, 而成为最受欢迎的模型之一^[5-7]。

TSK 模糊系统由若干条模糊规则构成, 每条模糊规则由前件和后件组成。传统的模糊规则构

建依靠专家经验。近年来,数据驱动的模糊规则构建方法受到了充分的研究。通常可以划分为两个步骤:一是使用某种划分规则将训练数据分为若干子体来提取规则前件参数,在实际建模中,通常使用聚类来实现;二是学习优化后件参数,从机器学习的角度,可以视为一个线性回归问题^[8-11]。

TSK 模糊系统建模是重要的有监督学习的过程,因此需要充分的训练数据。然而,在很多实际应用中,样本数据经常是有限的高维数据,这就不可避免导致了模型的过拟合问题。而多任务学习可以从其他任务中获取相关信息,一定程度弥补训练数据不足的问题,进而提高模型的学习性能^[12-17]。在多任务建模中,任务之间往往具有明显的相关性,并存在着共享信息,因此充分利用多个任务间的共享信息进行多任务模糊系统建模,有助于提高每个任务的泛化性能。例如, Jiang 等^[18]提出了一种利用潜在任务间关系信息的多任务模糊系统,该方法为多个任务学习了一个共享的后件参数来表示任务间的共享信息。然而,这些方法都只着重于任务间后件参数的共享结构,而忽视了如何利用各任务的自身特点。

鉴于此,本文提出了一种新型多任务 TSK 模糊系统建模方法,在挖掘多任务间共享信息的同时,保留单个任务的特殊性。该方法将多任务的后件参数矩阵分解为共享参数矩阵和特有参数矩阵两个部分:共享参数矩阵表示了任务之间共享的结构信息,而特有参数矩阵保留了每个任务的不同于其他任务的差异信息。本文通过为共享参数矩阵和特有参数矩阵分别引入低秩和稀疏约束来实现这一目标。

1 多任务 TSK 模糊模型的基本原理

经典的单任务 TSK 模糊系统利用多个局部线性子模型来近似非线性模型。而多任务模糊系统就是多个单任务模糊系统的联合优化。在多任务设置中, $T, N = \sum_{t=1}^T N_t$ 和 D 分别表示任务、样本和特征的数量,其中 N_t 表示任务 t 的样本数量。对于任意输入向量 $\mathbf{x}_t = (x_t^1, x_t^2, \dots, x_t^D)^T \in \mathbf{R}^D$, 其表示任务 t 的一个样本的特征向量,其中 $x_t^d, d = 1, 2, \dots, D$ 是向量 \mathbf{x}_t 的第 d 变量。

因此,任务 t 的第 m 个规则可以表示为:

$$\begin{aligned} & \text{IF } x_t^1 \text{ is } A_{t,1}^{m,1} \wedge x_t^2 \text{ is } A_{t,2}^{m,2} \wedge \dots \wedge x_t^D \text{ is } A_{t,D}^{m,D}, \\ & \text{THEN } f_t^m(\mathbf{x}_t) = w_{t,1}^{m,0} + w_{t,1}^{m,1} x_t^1 + \\ & \quad w_{t,2}^{m,2} x_t^2 + \dots + w_{t,D}^{m,D} x_t^D \end{aligned} \quad (1)$$

式中: $A_{t,d}^{m,d}$ 表示输入变量 x_t^d 对应于任务 t 的第 m 个规则的模糊子集, $m = 1, 2, \dots, M, d = 1, 2, \dots, D$,

M 表示模糊规则数, \wedge 表示合取操作。本文选择式 (2) 的 Gaussian 函数作为模糊隶属度函数。

$$\mu_{A_{t,d}^{m,d}}(x_t^d) = \exp\left(-\frac{(x_t^d - c_{t,d}^{m,d})^2}{2\sigma_{t,d}^{m,d}}\right) \quad (2)$$

$$c_{t,d}^{m,d} = \frac{\sum_{i=1}^T \sum_{i=1}^{N_t} \mu_{t,i}^m x_{t,i}^d}{\sum_{i=1}^T \sum_{i=1}^{N_t} \mu_{t,i}^m} \quad (3)$$

$$\sigma_{t,d}^{m,d} = h \sum_{i=1}^T \sum_{i=1}^{N_t} \mu_{t,i}^m (x_{t,i}^d - c_{t,d}^{m,d}) \cdot \left(\sum_{i=1}^T \sum_{i=1}^{N_t} \mu_{t,i}^m \right)^{-1} \quad (4)$$

式中 $c_{t,d}^{m,d}$ 和 $\sigma_{t,d}^{m,d}$ 分别表示均值与方差。本文使用 FCM 聚类方法进行模糊划分,获得样本的模糊隶属度,其中 $\mu_{t,i}^m$ 表示任务 t 所属的第 i 个样本对于第 m 个聚类的模糊隶属度。式 (4) 中, h 是可调节参数。

定义 $(\mathbf{x}_{t,i}, y_{t,i})$ 表示任务 t 中的第 i 个样本和其对

应的标签, $i = 1, 2, \dots, N_t, t = 1, 2, \dots, T, \mathbf{X}_t = \begin{pmatrix} \mathbf{x}_{t,1}^T \\ \mathbf{x}_{t,2}^T \\ \vdots \\ \mathbf{x}_{t,N_t}^T \end{pmatrix} \in$

$\mathbf{R}^{N_t \times D}, \mathbf{y}_t = [y_{t,1}, y_{t,2}, \dots, y_{t,N_t}]^T \in \mathbf{R}^{N_t}$ 。对于任务 t , 其模糊系统的输出可以表示为

$$\hat{\mathbf{y}}_t = \Phi_t \mathbf{w}_t \quad (5)$$

式中: $\Phi_t = (\Phi_t^1, \Phi_t^2, \dots, \Phi_t^M)$ 称为模糊系统的字典, $\Phi_t^m = \text{diag}(\varphi^m(\mathbf{x}_{t,1}), \varphi^m(\mathbf{x}_{t,2}), \dots, \varphi^m(\mathbf{x}_{t,N_t})) \cdot (1, \mathbf{X}_t) \in \mathbf{R}^{N_t \times D}$, $\mathbf{y}_t = [y_{t,1}, y_{t,2}, \dots, y_{t,N_t}]^T \in \mathbf{R}^{N_t}$, $\varphi^m(\mathbf{x}_{t,i})$ 表示输入 $\mathbf{x}_{t,i}$ 对于第 m 条模糊规则的触发强度,可由式 (6) 获得。 $\mathbf{w}_t = ((\mathbf{w}_t^1})^T, (\mathbf{w}_t^2})^T, \dots, (\mathbf{w}_t^M})^T)^T \in \mathbf{R}^{(D+1)M}$ 是任务 t 的后件参数向量; \mathbf{w}_t^m 是对应于 Φ_t^m 的后件参数向量。

$$\varphi^m(\mathbf{x}_t) = \frac{\mu^m(\mathbf{x}_t)}{\sum_{k=1}^M \mu^k(\mathbf{x}_t)} \quad (6)$$

$$\mu^m(\mathbf{x}_t) = \prod_{d=1}^D \mu_{A_{t,d}^{m,d}}(x_t^d) \quad (7)$$

为方便计算,本文为多任务 TSK 模糊模型进一步定义 $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_T) \in \mathbf{R}^{(D+1)M \times T}$, 表示多任务模糊模型的后件参数联合矩阵。所以,基本多任务模糊系统的目标函数可以表示为

$$\min_{\mathbf{W}} \sum_{t=1}^T \|\mathbf{y}_t - \Phi_t \mathbf{w}_t\|_2^2 + R(\mathbf{W}) \quad (8)$$

2 任务间共享和特有结构分解的多任务 TSK 模糊系统建模新方法

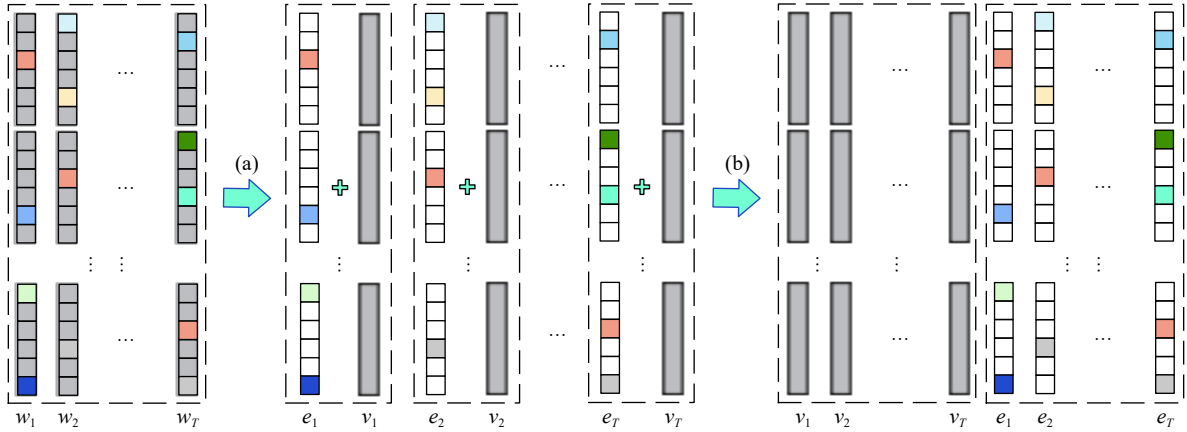
本节在基本多任务模糊系统的基础上,进一步提出新型多任务模糊系统建模方法。考虑到多任务之间是相互关联的,因此可以认为多个任务的模型参数包含着潜在的共享信息;另一方面,

各任务的模型参数中都包含了自身不同于其他任务的特有信息。如图1所示,本文将各后件参数联合矩阵 W 分解为共享参数矩阵 V 和特有参数矩阵 E , 即:

$$W = V + E \quad (9)$$

共享参数矩阵 V 包含了任务之间的共享参数信息,这种共享信息是指同一个特征在不同的任务中发挥相似的作用。具体表现在,如果某个特征在任务 i 中被赋予了一个较高的权重值,那么它

在相关任务 j 中也将被赋予较高的权重,反之亦然,即对应于任务 i 和任务 j 的共享参数 v_i 和 v_j 是相似的。因此,共享参数矩阵 V 具有低秩性,本文通过引入核范数 $\|V\|_*$ 来实现对共享参数矩阵 V 的低秩约束 $\|V\|_*$ 。特有参数矩阵 E 表示各项任务不同于其他任务的特有信息,这种特有信息体现在某一特征在不同任务中发挥不同的作用,即特有参数矩阵 E 是行稀疏的,本文通过引入 $\|E\|_{1,2}$ 正则化项实现行稀疏。



注: (a) 将后件参数矩阵 W 中的每个列向量 $w_t, t=1, 2, \dots, T$, 分解为共享参数 v_t 和特有参数 e_t 两个分量; (b) 调整各向量顺序

图1 后件参数矩阵 $W = (w_1, w_2, \dots, w_T)$ 可以拆分为低秩部分 $V = (v_1, v_2, \dots, v_T)$ 和稀疏部分 $E = (e_1, e_2, \dots, e_T)$

Fig. 1 The consequent parameter matrix $W = (w_1, w_2, \dots, w_T)$ can be decomposed into the low-rank component $V = (v_1, v_2, \dots, v_T)$ and the sparse component $E = (e_1, e_2, \dots, e_T)$

所以,通过对后件参数联合矩阵的分解,再分别施加低秩和行稀疏约束,在多任务建模中兼顾多任务之间共享信息和特有信息的作用,提出了任务间共享和特有结构划分的多任务 TSK 模糊系统的目标函数,表示如下:

$$\min_W \sum_{t=1}^T \|y_t - \Phi_t w_t\|_2^2 + \alpha \|V\|_* + \beta \|E\|_{1,2} \quad (10)$$

s.t. $W = V + E$

式中: α 和 β 是正则化参数,用于调节共享参数和特有参数在模型训练中发挥的作用,参数越大,惩罚力度越大。

3 目标函数优化

本文使用增广拉格朗日乘子法^[19-20]求解式(10)提出的最优化问题,其增广拉格朗日目标函数为

$$\min_{W, V, E, Y} \sum_{t=1}^T \|y_t - \Phi_t w_t\|_2^2 + \alpha \|V\|_* + \beta \|E\|_{1,2} + \langle Y, W - V - E \rangle + \frac{\mu}{2} \|W - V - E\|_F^2 \quad (11)$$

式中: $Y \in \mathbf{R}^{(D+1) \times T}$ 是拉格朗日乘子矩阵, $\langle \cdot, \cdot \rangle$ 表示两个矩阵的内积运算。为便于计算,使用 LADMAP 方法^[21]将目标函数重新表示为

$$\min_{W, V, E, Y} \sum_{t=1}^T \|y_t - \Phi_t w_t\|_2^2 + \alpha \|V\|_* + \beta \|E\|_{1,2} + \frac{\mu}{2} \|W - V - E + \frac{Y}{\mu}\|_F^2 \quad (12)$$

这是一个包含 W 、 V 、 E 、 Y 4 个优化变量的最优化问题,所以本文交替使用其他变量的固定值迭代优化每个变量,原问题被转化为下面若干个子问题:

1) 固定 V 、 E 、 Y , 式(12)化为如下目标函数:

$$\min_{W, V, E} \frac{1}{N} \sum_{t=1}^T \|y_t - \Phi_t w_t\|_2^2 + \frac{\mu}{2} \|W - V - E + \frac{Y}{\mu}\|_F^2 \quad (13)$$

对 w_t 求导,并令其等于 0, 整理之后,可以通过式(14)获得 w_t 的解:

$$w_t = (\mu I + 2(\Phi_t)^T \Phi_t)^{-1} [2(\Phi_t)^T y_t + \mu(v_t + e_t - y_t/\mu)] \quad (14)$$

因此,可以进一步得到 W 的解:

$$W = (w_1, w_2, \dots, w_T) \quad (15)$$

2) 固定 W 、 E 、 Y , 式(12)转化为目标函数:

$$\min_V \alpha \|V\|_* + \frac{\mu}{2} \|W - V - E + \frac{Y}{\mu}\|_F^2 \quad (16)$$

使用奇异值阈值算子^[22]求解式(16)的低秩问题, V 的最优解可以写成如下形式:

$$V = \mathcal{D}_{\frac{\alpha}{\mu}} \left(W - E + \left(\frac{1}{\mu} \right) Y \right) \quad (17)$$

3) 固定 W 、 V 、 Y , 式(11)变成如下目标函数:

$$\min_E \beta \|E\|_{1,2} + \frac{\mu}{2} \left\| W - V - E + \frac{Y}{\mu} \right\|_2^2 \quad (18)$$

式(18)等同于求解如下问题:

$$\min_E \frac{\beta}{\mu} \|E\|_{1,2} + \frac{1}{2} \|V - Q\|_2^2 \quad (19)$$

其中 $Q = W - V + Y/\mu$, 可以使用文献[23]中的方法求解式(19)的最优化问题, 得到 E 的最优解。

4) 更新拉格朗日乘子矩阵 Y 和正则化参数 μ :

$$Y = Y + \mu(W - V - E) \quad (20)$$

$$\mu = \min(\rho\mu, \mu_{\max}) \quad (21)$$

式中: ρ 是一个大于1的正数。

最终, 本文提出多任务 TSK 模糊系统建模方法具体描述如下:

算法 MTTSKFS-CS

输入 多任务数据集 X_1, X_2, \dots, X_T 和对应的标签 y_1, y_2, \dots, y_T ; 模糊规则数量 M ; 正则化参数 α, β, μ ; 可调节参数 h ; 正整数 $\rho > 1, \mu_{\max} = 10^3$;

输出 多任务的模糊前件和多任务后件参数联合矩阵。

训练过程

1) 生成模糊字典: 首先对全部任务的样本使用 FCM 聚类, 获得 M 个聚类中心。然后计算每个样本的模糊隶属度, 生成每个任务的模糊字典。

2) 联合学习多任务的后件参数: 求解式(12), 得到 W 、 V 、 E 最优解:

① 初始化: 设定 W 是一个随机矩阵, $V = W$, $E = W - V$ 。拉格朗日乘子矩阵 $Y = \frac{W}{\|W\|_2}$ 。

② 当式(12)不收敛

使用式(15)更新 W ;

使用式(17)更新 V ;

使用式(19)更新 E ;

使用式(20)更新 Y ;

使用式(21)更新 μ ;

end while

return W

3) 输出多任务的模糊前件和多任务后件参数联合矩阵。

4 结果与分析

4.1 实验设置

为了验证本文提出的多任务建模方法的有效

性, 在多个真实数据集上进行泛化性能实验。实验中用到的数据集包含2种类型: 一种是相同输入不同输出的 SIDO 数据集, 将数据集的每一个输出作为一个回归任务构成一个多任务数据集, 数据集的多个任务共享同一个输入数据, 但每个任务拥有不同的输入空间到输出空间的映射函数; 另一种是不同输入不同输出的 DIDO 数据集, 同样将每一个输出作为一个回归任务, 每个任务拥有不同的输入数据和不同的输入空间到输出空间的映射函数。

本节选择的 Slump、Parkinsons Telemonitoring、Winequality、House 数据集来自 UCI Machine Learning Repository 数据集网站。Slump 数据集用来模拟混凝土的坍落度, 涉及到坍落度、流动性和抗压强度3个输出, 是一个 SIDO 数据集。Parkinsons Telemonitoring 数据集由来自42名早期帕金森氏症患者的生物医学声音测量组成, 用来测量 motor 和 total UPDRS scores, 也是一个 SIDO 数据集。Winequality 数据集是一个利用基于理化测试数据来划分葡萄酒质量等级的数据集, 包括红葡萄酒和白葡萄酒两个子集, 本节分别将每个子集视为一个任务, 这是一个 DIDO 数据集。House 数据集被用于波士顿房价预测, 在本节根据特征变量“RAD”的值将数据集划分为 (Task 1: RAD<5; Task 2: 5<=RAD<7.5; Task 3: RAD>=7.5)3个子集, 作为一个 DIDO 数据集。Multivalued (MV) Data Modeling 数据集可以从 KEEL Datasets Repository 获得, 是一个具有特征间依赖关系的人工数据集, 本节中根据第8个特征变量可以将数据集划分为两个任务, 作为 DIDO 数据集。表1中列出来了上述所选用的数据集的特征维数、样本数量等具体细节。

表1 多任务数据集的详细信息
Table 1 Details of the multitask datasets

数据集	类型	任务	特征数	样本数
Slump	SIDO	Task 1	7	103
		Task 2	7	103
		Task 3	7	103
Parkinsons Telemonitoring	SIDO	Task 1	26	5875
		Task 2	26	5875
Winequality	DIDO	Task 1	11	1599
		Task 2	11	4898
Multivalued (MV) Data Modeling	DIDO	Task 1	10	1000
		Task 2	10	1000
Housing	DIDO	Task 1	13	191
		Task 2	13	182
		Task 3	13	132

我们在实验中比较了几种经典的回归算法, 参数的设置通过 5 折交叉验证来进行寻优, 这些包括多任务和单任务回归算法。算法中涉及到的 算法的详细介绍以及参数的寻优范围如表 2 所示。

表 2 实验中各算法参数的详细设置

Table 2 Detailed settings of all algorithm's parameters

算法	说明	参数寻优范围
MTTSKFS-CS	本文提出的多任务模糊系统建模方法	调节参数 $h \in \{10^{-4}, 10^{-3}, \dots, 10^4\}$ 模糊规则数 $M \in \{5, 10, \dots, 50\}$ 正则化参数 $\alpha \in \{2^{-4}, 2^{-3}, \dots, 2^5\}$ 正则化参数 $\beta \in \{10^{-4}, 10^{-3}, \dots, 10^5\}$
MT-TSK-FS	一种利用潜在任务间关系信息的多任务模糊系统 ^[18]	模糊规则数 $M \in \{5, 10, \dots, 100\}$ 正则化参数 $\tau_k \in \{2^{-4}, 2^{-3}, \dots, 2^5\}$ 正则化参数 $\lambda \in \{10^{-4}, 10^{-3}, \dots, 10^5\}$
CSL	一种 correlation-aware sparse and low-rank constrained 多任务建模方法 ^[24]	正则化参数 $\lambda_1 \in \{10^{-4}, 10^{-3}, \dots, 10^4\}$ 正则化参数 $\lambda_2 \in \{10^{-4}, 10^{-3}, \dots, 10^4\}$ 正则化参数 $\lambda_3 \in \{10^{-4}, 10^{-3}, \dots, 10^4\}$
Least_L21	基于 L _{2,1} 范数正则化的最小二乘多任务建模方法 ^[25]	正则化参数 $\gamma \in \{10^{-5}, 10^{-4}, \dots, 10^5\}$ 正则化参数 $\lambda \in \{10^{-4}, 10^{-3}, \dots, 10^5\}$
TSKFS-SVR	基于 SVR 的单任务 TSK 模糊系统建模方法 ^[26]	调节参数 $h \in \{10^{-4}, 10^{-3}, \dots, 10^4\}$ 模糊规则数 $M \in \{5, 10, \dots, 50\}$ SVM 代价函数 $C \in \{2^{-4}, 2^{-3}, \dots, 2^5\}$ 高斯核, 核宽参数 $\sigma^2 \in \{10^{-4}, 10^{-3}, \dots, 10^4\}$

本文选用 RRSE 来评价各对比算法的泛化性能, 定义如下:

$$J_{RRSE} = \sqrt{\sum_{i=1}^N (\hat{y}_i - y_i)^2 / \sum_{i=1}^N (\bar{y}_i - y_i)^2} \quad (22)$$

式中: \hat{y}_i 是模型的预测输出; \bar{y}_i 是真实标签的均值。 J_{RRSE} 越小表示模型的泛化能力越好。

4.2 泛化性能实验

我们分别在每个数据集上验证了 MTTSKFS-

CS 及对比算法。

分别计算各模型在每个数据集的每个子任务上的泛化性能, 其中“Average”表示算法在每个数据集的所有任务中的平均表现。若算法为单任务算法时, 分别对每个子任务进行建模, 来评价算法性能。本文提出的 MTTSKFS-CS 建模方法与对比方法在真实数据集上的实验结果如表 3 所示。

表 3 所有算法在各数据集上的泛化性能比较

Table 3 Comparison of generalization performance of all algorithms on datasets

数据集	任务	MTTSKFS-CS		MT-TSK-FS		CSL		Least_L21		TSKFS-SVR	
		RRSE	Std	RRSE	Std	RRSE	Std	RRSE	Std	RRSE	Std
Slump	Task 1	0.8971	0.0391	0.9405	0.0361	0.8978	0.0160	1.5586	0.9526	1.6684	0.0184
	Task 2	0.7842	0.0191	0.8987	0.0462	0.7759	0.0127	1.0839	0.0161	1.1458	0.0165
	Task 3	0.3468	0.0154	0.3592	0.0251	0.3608	0.0051	0.3955	0.0090	0.3312	0.0140
	Average	0.6760	0.0157	0.7328	0.0192	0.6781	0.0065	0.9040	0.0078	0.9342	0.0157
Parkinsons Telemonitoring	Task 1	0.9071	0.0023	0.9126	0.0099	0.9806	0.0014	1.0129	0.0015	0.9260	0.0011
	Task 2	0.9059	0.0004	0.9213	0.0183	0.9847	0.0005	1.0153	0.0022	0.9328	0.0011
	Average	0.9065	0.0010	0.9170	0.0125	0.9827	0.0008	1.0141	0.0017	0.9249	0.0009

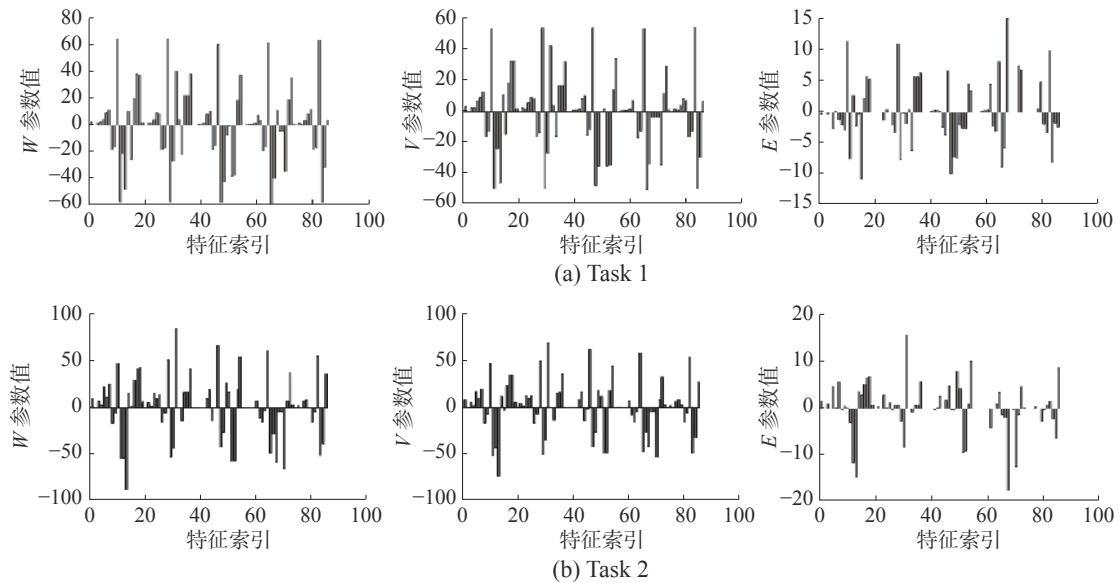
续表3

数据集	任务	MTTSKFS-CS		MT-TSK-FS		CSL		Least_L21		TSKFS-SVR	
		RRSE	Std	RRSE	Std	RRSE	Std	RRSE	Std	RRSE	Std
Winequality	Task 1	0.803 2	0.003 8	0.799 3	0.002 4	0.808 7	0.002 1	0.970 7	0.001 6	0.841 6	0.005 7
	Task 2	0.821 6	0.004 4	0.839 6	0.001 8	0.855 3	0.000 4	0.913 6	0.000 6	0.848 3	0.003 2
	Average	0.812 4	0.002 0	0.819 4	0.002 1	0.832 0	0.000 9	0.942 1	0.000 9	0.845 0	0.002 2
MV	Task 1	0.080 4	0.000 8	0.090 8	0.004 9	0.136 6	0.001 5	0.247 2	0.010 4	0.139 0	0.001 2
	Task 2	0.086 6	0.003 8	0.097 5	0.003 9	0.093 3	0.002 4	0.955 3	0.008 1	0.115 4	0.000 9
	Average	0.083 5	0.001 8	0.094 2	0.003 4	0.115 0	0.001 8	0.061 2	0.009 2	0.127 2	0.001 1
Housing	Task 1	0.400 3	0.011 7	0.410 5	0.019 9	0.405 7	0.009 0	0.675 8	0.017 7	0.541 3	0.017 9
	Task 2	0.378 5	0.008 2	0.380 8	0.008 6	0.382 7	0.007 3	0.681 0	0.013 3	0.677 9	0.023 4
	Task 3	0.626 9	0.012 5	0.641 1	0.007 7	0.634 1	0.013 9	0.718 9	0.020 1	0.757 6	0.039 5
	Average	0.468 5	0.005 1	0.477 5	0.008 5	0.474 2	0.005 7	0.691 9	0.009 9	0.658 9	0.013 4

4.3 后件参数划分的可视化分析

为了进一步说明本文提出的对于后件参数矩阵的低秩和稀疏划分的重要作用,我们将模型在 Parkinsons Telemonitoring 数据集上训练出的后件

参数 W 、 V 、 E 进行可视化。图2展示了后件参数的可视化结果,从中可以看到,训练出的参数结果,基本符合图1中的假设, W 被划分为了低秩部分 V 和稀疏部分 E 。

图2 W 、 V 、 E 的可视化结果Fig. 2 Visualization of W , V , E

4.4 算法收敛性分析

为了进一步研究本文提出的 MTTSKFS-CS 建模方法的收敛性,我们选取了 Multivalued (MV) Data Modeling、Parkinsons Telemonitoring 和 House 3 个数据集,通过交叉验证寻优得到最优参数,并在最优参数的基础上进行收敛性实验。算法在 3 个数据集上的收敛曲线如图3所示。从收敛曲线中可以看到,算法在前期可以快速收敛,并迅速进入稳定状态。实验结果说明,本文第

3 节提出的优化方法具有良好的收敛性能,能够真正达到模型最优化的目的,从而使模型获得较高的实用性。

4.5 实验结果分析

表3的实验结果证明 MTTSKFS-CS 方法在大多数任务上获得了比对比算法更好的性能表现。与单任务方法相比,多任务建模方法明显提升了每个任务的预测表现,拥有更好的泛化性能。与多任务算法的对比结果说明,本文提出的方法在

利用了多任务之间的共享信息的同时有效利用了单个任务自身的特有信息,从而获得了更好的表现。对于图2的可视化结果,我们可以看到,后

件参数联合矩阵 \mathbf{W} 被划分为了低秩部分 \mathbf{V} 和稀疏部分 \mathbf{E} , 实现了本文提出的模型设想,也间接验证了 MTTSKFS-CS 模型的有效性。

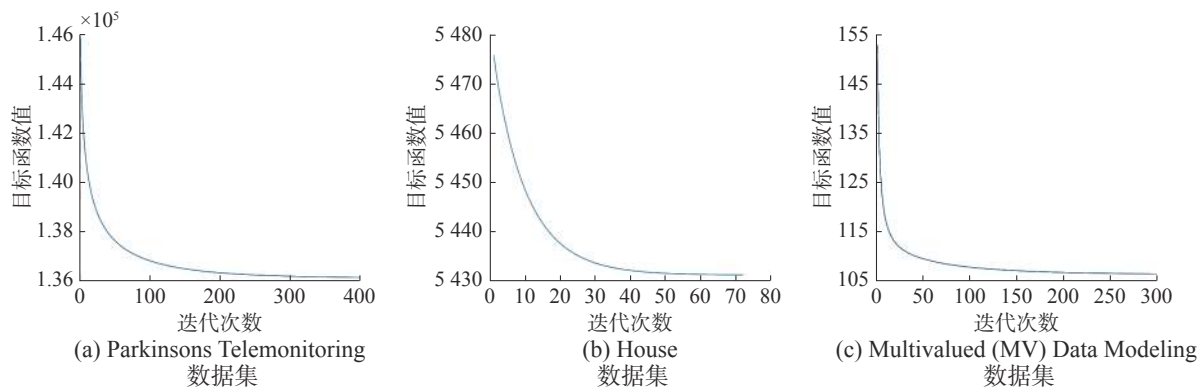


图3 MTTSKFS-CS 方法的收敛曲线

Fig. 3 Convergence curve of MTTSKFS-CS algorithm

5 结束语

本文提出了一种新型多任务模糊系统建模方法,首先使用模糊聚类方法获得多任务的模糊前件,然后通过合理划分后件参数联合矩阵为共享参数矩阵和特有参数矩阵,同时兼顾多任务之间的共享信息和各任务的特有信息。最后通过 ALM 方法求解最优化问题,获得模型的最优解。在多个真实多任务数据集上的实验结果说明了,本文提出的 MTTSKFS-CS 建模方法能够有效解决传统多任务模型只着重于多任务共享信息的问题。在今后的工作中,如何更好地在建模中平衡共享信息和特有信息将是我们研究的重点。

参考文献:

- [1] LI Chaoshun, ZHOU Jianzhong, CHANG Li, et al. T-S fuzzy model identification based on a novel hyperplane-shaped membership function[J]. *IEEE transactions on fuzzy systems*, 2017, 25(5): 1364–1370.
- [2] XU Peng, DENG Zhaohong, CUI Chen, et al. Concise fuzzy system modeling integrating soft subspace clustering and sparse learning[J]. *IEEE transactions on fuzzy systems*, 2019, 27(11): 2176–2189.
- [3] CHANG P C, LIU Chenhao. A TSK type fuzzy rule based system for stock price prediction[J]. *Expert systems with applications*, 2008, 34(1): 135–144.
- [4] ZHOU Shangming, GAN J Q. Extracting Takagi-Sugeno fuzzy rules with interpretable submodels via regularization of linguistic modifiers[J]. *IEEE transactions on knowledge and data engineering*, 2009, 21(8): 1191–1204.
- [5] DENG Zhaohong, CHOI K S, CHUNG F L, 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.
- [6] JUANG C F, HSIEH C D. TS-fuzzy system-based support vector regression[J]. *Fuzzy sets and systems*, 2009, 160(17): 2486–2504.
- [7] 林得富, 王骏, 张嘉旭, 等. Takagi-Sugeno 模糊系统双正则联合稀疏建模新方法 [J]. *计算机科学与探索*, 2019, 13(6): 1016–1026.
- LIN Defu, WANG Jun, ZHANG Jiaxu, et al. Novel Takagi-Sugeno fuzzy system modeling method via joint sparse learning using two regularizations[J]. *Journal of frontiers of computer science and technology*, 2019, 13(6): 1016–1026.
- [8] 张春香, 王骏, 张嘉旭, 等. 面向自闭症辅助诊断的联合组稀疏 TSK 建模方法 [J]. *计算机科学与探索*, 2020, 14(2): 2083–2093.
- ZHANG Chunxiang, WANG Jun, ZHANG Jiaxu, et al. Novel TSK modeling method with joint group sparse learning for autism aided diagnosis[J]. *Journal of frontiers of computer science and technology*, 2020, 14(2): 2083–2093.
- [9] ZHU Yuanguo. Fuzzy optimal control for multistage fuzzy systems[J]. *IEEE transactions on systems, man, and cybernetics, part B (cybernetics)*, 2011, 41(4): 964–975.
- [10] JANG J S R. ANFIS: adaptive-network-based fuzzy inference system[J]. *IEEE transactions on systems, man, and cybernetics*, 1993, 23(3): 665–685.
- [11] REZAEI B, ZARANDI M H F. Data-driven fuzzy modeling for Takagi-Sugeno-Kang fuzzy system[J]. *Information sciences*, 2010, 180(2): 241–255.
- [12] CAVALLANTI G, CESA-BIANCHI N, GENTILE C. Linear algorithms for online multitask classification[J].

- Journal of machine learning research, 2010, 11(5): 2901–2934.
- [13] EVGENIOU T, MICCHELLI C A, PONTIL M. Learning multiple tasks with kernel methods[J]. Journal of machine learning research, 2005, 6(4): 615–637.
- [14] ZHANG Jiangmei, YU Binfeng, JI Haibo, et al. Multi-task feature learning by using trace norm regularization[J]. *Open physics*, 2017, 15(1): 674–681.
- [15] ZHAO Qian, RUI Xiangyu, HAN Zhi, et al. Multilinear multitask learning by rank-product regularization[J]. *IEEE transactions on neural networks and learning systems*, 2020, 31(4): 1336–1350.
- [16] ZHONG Shi, PU Jian, JIANG Yugang, et al. Flexible multi-task learning with latent task grouping[J]. *Neuro-computing*, 2016, 189: 179–188.
- [17] XUE Ya, LIAO Xuejun, CARIN L, et al. Multi-task learning for classification with dirichlet process priors[J]. Journal of machine learning research, 2007, 8: 35–63.
- [18] JIANG Yizhang, DENG Zhaohong, CHUNG F L, et al. Multi-task TSK fuzzy system modeling using inter-task correlation information[J]. *Information sciences*, 2015, 298: 512–533.
- [19] WANG Jun, LIN Defu, DENG Zhaohong, et al. Multi-task TSK fuzzy system modeling by jointly reducing rules and consequent parameters[J]. *IEEE transactions on systems, man, and cybernetics: systems*, 2019.
- [20] MENG Fan, YANG Xiaomei, ZHOU Chenghu, et al. The augmented lagrange multipliers method for matrix completion from corrupted samplings with application to mixed gaussian-impulse noise removal[J]. *PLoS one*, 2014, 9(9): e108125.
- [21] LIN Zhouchen, LIU Risheng, SU Zhixun. Linearized alternating direction method with adaptive penalty for low-rank representation[C]//Proceedings of 25th Annual Conference on Neural Information Processing Systems. Granada, Spain, 2011: 612–620.
- [22] CAI Jianfeng, CANDÈS E J, SHEN Zuwei. A singular value thresholding algorithm for matrix completion[J]. *SIAM journal on optimization*, 2010, 20(4): 1956–1982.
- [23] LIU Guangcan, LIU Zhouchen, YAN Shuicheng. Robust recovery of subspace structures by low-rank representation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(1): 171–184.
- [24] JIANG Pengbo, WANG Xuetong, LI Qionglin, et al. Correlation-aware sparse and low-rank constrained multi-task learning for longitudinal analysis of alzheimer’s disease[J]. *IEEE journal of biomedical and health informatics*, 2019, 23(4): 1450–1456.
- [25] ZHOU Jiayu, CHEN Jianhui, YE Jieping. Malsar: multi-task learning via structural regularization—user’s manual version 1.1[EB/OL]. (2019-12-12) [2020-07-08] <https://github.com/jiayuzhou/MALSAR>.
- [26] JUANG C F, CHIU S H, SHIU S J. Fuzzy system learned through fuzzy clustering and support vector machine for human skin color segmentation[J]. *IEEE transactions on systems, man, and cybernetics-part A: systems and humans*, 2007, 37(6): 1077–1087.

作者简介:



赵壮壮, 硕士研究生, 主要研究方向为模式识别与人工智能、模糊系统。



王骏, 副教授, 主要研究方向为机器学习、模糊系统、医学影像分析。



潘祥, 副教授, 主要研究方向为医学图像诊断、计算机视觉、AI 医疗诊断。主持国家自然科学基金项目 1 项, 安徽省自然科学基金项目 1 项。获得授权发明专利 6 项, 受理发明专利 2 项。发表学术论文 20 余篇。