

多尺度核随机配置网络的多目标回归算法

代伟, 吴尚, 南静, 刘鑫[†]

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221116)

摘要: 多目标回归通过将多个回归任务的相关性纳入建模中以提高模型表现, 但是当前方法对于多目标内在更加普遍且复杂的非线性结构关系建模效果并不理想. 本文提出多尺度核随机配置网络(MSK-SCN)构建高维输入与多目标之间的映射以及多目标内在的非线性关系. 通过设计隔离建模机制, 使得每个子模型根据其负责的目标误差状态完成隐含层参数配置, 进而避免多个回归任务共享模型参数导致建模质量下降; 建立多尺度核空间, 利用不同尺度参数的核函数将低维空间数据映射到高维特征空间以增强数据的表达能力, 进而提高模型挖掘多目标间非线性关系的能力. 基于虚拟数据集和真实数据集进行实验验证, 结果表明MSK-SCN的表现优于当前的多目标算法, 证实了其在多目标回归问题上的有效性.

关键词: 随机配置网络; 多目标回归; 多核学习; 多尺度核

引用格式: 代伟, 吴尚, 南静, 等. 多尺度核随机配置网络的多目标回归算法. 控制理论与应用, 2026, 43(4): 915 – 926

DOI: 10.7641/CTA.2024.40038

Multi-target regression by multi-scale kernel stochastic configuration network

DAI Wei, WU Shang, NAN Jing, LIU Xin[†]

(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou Jiangsu 221116, China)

Abstract: Multi-objective regression improves model performance by incorporating the correlation of multiple regression tasks into modelling, but the current method is not ideal for more general and complex nonlinear structural relationships inherent in multi-target modeling. In this paper, a multi-scale kernel stochastic configuration network (MSK-SCN) is proposed to construct the correlations between high-dimensional inputs and multi-targets as well as the intrinsic nonlinear relationships of multi-targets. The isolation modeling mechanism is designed to make each submodel complete the hidden layer parameter configuration according to its responsible target error state, so as to avoid the modeling quality degradation caused by multiple regression tasks sharing model parameters. The multi-scale kernel space is established, and the kernel function of different scale parameters is used to map the low-dimensional data to the high-dimensional feature space to enhance the expression ability of the data, improving the ability of the model to mine the nonlinear relationship between multiple targets. The simulation results show that MSK-SCN performs better than the current multi-target algorithm through experiments on virtual data sets and real data sets, which proves its effectiveness in multi-target regression problems.

Key words: stochastic configuration network; multi-target regression; multi-kernel learning; multi-scale kernel

Citation: DAI Wei, WU Shang, NAN Jing, et al. Multi-target regression by multi-scale kernel stochastic configuration network. *Control Theory & Applications*, 2026, 43(4): 915 – 926

1 引言

数据回归是指寻找使得预测误差达到最小的一种变量间最佳拟合关系^[1]. 神经网络模型由于强大的非线性映射的能力而被广泛用于回归任务中, 然而利用

梯度迭代法训练网络参数的传统神经网络存在收敛速度慢、易于陷入局部极小以及网络尺寸缺乏理论指导等问题^[2]. 随机向量函数链^[3]采用随机学习算法配置网络参数, 避免了梯度迭代的缺点, 具有快速的学

收稿日期: 2024-01-17; 录用日期: 2024-11-20.

[†]通信作者. E-mail: 15B904027@hit.edu.cn.

本文责任编辑: 王凌.

国家重点研发项目(2022YFB3304700), 江苏省杰出青年基金项目(BK20240102), 国家自然科学基金项目(62373361)资助.

Supported by the National Key Research and Development Program of China (2022YFB3304700), the Science Fund for Distinguished Young Scholars of Jiangsu Province (BK20240102) and the National Natural Science Foundation of China (62373361).

习能力和良好的泛化能力^[4]. 然而, 现有研究表明, 如果随机参数范围设置不合理, 那么随机向量函数连接网络的无限逼近能力则无法保证^[5].

2017年, 一种带有约束机制的随机学习算法被提出, 即随机配置网络(stochastic configuration network, SCN)^[6], 其隐含层节点在监督机制的约束下从候选节点池中选择最佳的参数对, 并且以点增量的方式构建模型, 使得模型结构紧致且具有良好的泛化能力. 同时通过自适应地调整随机参数范围, 保证了SCN的无限逼近能力, 成为解决多输出回归问题的有效工具^[7]. 但是其在多目标回归(multi-target regression, MTR)建模过程中仅关注高维输入与多输出(目标)的映射关系, 而忽略了多目标之间的相关性^[8], 这导致原始SCN模型在MTR上表现不太理想. 为提高SCN学习多目标内在相关性的能力, 在之前的工作中提出了多目标随机配置网络(multi-target SCN, MTSCN)^[9], 通过引入结构矩阵在无需知道先验知识的条件下, 以数据驱动的方式自行学习多输出内在的相关性, 同时利用参数交替优化的方式更新模型.

当前MTR方法可分为两类: 问题转化法与算法自适应法^[10], 前者将多目标的预测输出扩展到输入空间并基于循环迭代更新模型参数, 而后者利用一个模型同时拟合多个目标任务. 在问题转化法中, 文献[11]利用多个基模型拟合多个目标, 并将模型输出增广到原始输入空间形成元输入, 再使用元输入建模以修正模型参数. 文献[12]引入深度网络思想, 将文献[11]的两次建模推广至多次建模以更新学习信息. 但是, 机械地将多目标的预测堆叠至输入空间易产生冗余特征, 导致影响模型性能. 为此, 文献[13]应用聚类算法揭示多目标间的相似性并为每个目标生成一个独立相似度矩阵, 再通过查询相似度矩阵学习目标特定特征. 文献[14]通过探索目标间的相关性, 为每个目标构建其独特的标签特定特征, 从而降低输入空间中特征的冗余性. 但是, 这类方法在模型构建时缺乏一定的约束机制, 导致模型的收敛性无法得到保证. 为此, 文献[15]将多目标变量以模块化形式堆叠到输入空间, 并在模块之间以及模型内部基于残差逼近的学习方式更新参数, 保证了所建立模型的收敛性. 在问题转化法中, 早期方法如文献[16]往往基于线性回归模型建模目标相关性, 这限制了对高维输入与多输出之间复杂的非线性关系的学习能力. 为此, 文献[17–18]基于核技巧完成输入与输出间的非线性关系建模, 提高模型学习复杂MTR任务的能力. 在文献[17]的基础上, 文献[19]引入流行正则化保证预测值与真实值在局部与全局结构的一致性, 进而提高模型泛化能力. 文献[20]从特征选择角度使用结构矩阵编码目标相关性, 并结合 $L_{2,1}$ 范数最小化优化模型参数. 但是, 这几类方法都是以线性加权的方式学习多目标之间的相

关联系, 难以挖掘到更加复杂的非线性结构信息, 导致模型的应用能力受限. 为此, 文献[21]则通过学习一个半正定相似度矩阵, 以编码多目标输出间的关系, 然而这难以充分描述目标相关性. 为此, 文献[22]基于文献[21], 利用一个非线性核函数增强模型挖掘数据信息的能力, 进而提高多目标间非线性结构知识的表达能力. 但是, 单个核函数通常只能捕捉到特定类型的数据关系, 而难以学习到数据中更加复杂或多样化的非线性结构信息.

当将现有MTR方法用于SCN时, 不仅难以避免现存的问题, 还会存在由多目标误差难以保持一致所导致的参数折中选择质量不高的问题. 受上述问题驱动, 本文提出多尺度核随机配置网络(multi-scale kernel SCN, MSK-SCN), 同时对高维输入到多输出之间的联系以及多目标内部复杂的非线性结构信息进行建模. 通过为每个目标分配专属的SCN模型以基于目标的误差状态完成参数配置, 进而避免因模型共享隐层参数导致的建模质量下降. 为学习多目标之间的非线性结构信息, 通过引入尺度参数选择灵活的多尺度核函数构建多尺度核空间^[23]. 本文的主要创新之处在于:

1) 设计隔离建模机制, 以根据每个目标的误差状态完成节点参数配置, 解决了多目标共享隐层参数问题进而提高建模质量, 并且完成了高维输入与多输出之间联系的建模;

2) 建立基于不同尺度参数的核函数组成的多尺度核空间, 增强模型的数据信息表达效果, 进而提高了非线性结构信息的挖掘能力;

3) 构建隐层输出全局更新和结构矩阵梯度下降相结合的参数优化方法, 保证辅助输出和最终输出的误差单调降性.

2 相关工作

本部分主要介绍多目标回归(MTR), 随机配置网络(SCN)与多尺度核随机配置网络(MTSCN).

2.1 多目标回归

$D = \{(X, Y) | X \in \mathbb{R}^{n \times d}, Y \in \mathbb{R}^{n \times m}\}$ 表示训练数据集, $X \in \mathbb{R}^{n \times d}$ 和 $Y \in \mathbb{R}^{n \times m}$ 分别为输入与输出, 其中 n , d 和 m 分别表示样本数量、输入维度和目标个数. MTR的目的在于学习一个模型 $f: X \rightarrow Y$ 使其对于一个新的输入向量 x 的预测值 $\hat{y} = f(x)$ 能够最佳地逼近真实目标 y .

2.2 随机配置网络

SCN每个节点的输入权值和偏置都是通过监督机制从参数候选池中选择的最佳参数, 其构建的模型隐含层输出矩阵可以表示为 $H = [h_1 \ h_2 \ \cdots \ h_L]$,

h_L 表示第 L 个隐含层节点的输出,具体如下:

$$h_L = [g_L(w_L^T x_1 + b_L) \quad g_L(w_L^T x_2 + b_L) \quad \cdots \quad g_L(w_L^T x_n + b_L)]^T, \quad (1)$$

其中: g_L 是激活函数; w_L 与 b_L 是输入权值和偏置; $x_i \in \mathbb{R}^d, i = 1, 2, \dots, n$ 是输入样本; d 表示特征数. 由此, SCN的目标函数可以表示为

$$F = \|Y - H\beta\|_F^2, \quad (2)$$

其中: Y 是真实输出; $\beta = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_L]^T \in \mathbb{R}^{L \times m}$ 为输出权值, 其包括两种计算方式: 局部更新和全局更新. 前者只针对新增节点采用构造方法更新参数并保持之前节点的权值不变, 第 L 个节点的权值公式是

$$\beta_{L,q} = \frac{e_{L-1,q}^T \cdot h_L}{h_L^T \cdot h_L}, \quad q = 1, 2, \dots, m, \quad (3)$$

其中: m 是输出数, $e_{L-1,q}$ 表示第 q 个目标的误差. 全局更新在新节点增加后调整所有节点输出权值, 因此提高了收敛速度, 其计算公式为

$$\beta^* = \arg \min_{\beta} \|Y - H\beta\|_F^2 = H^+ Y, \quad (4)$$

其中 H^+ 是Moore-Ponrse广义逆.

2.3 多目标随机配置网络

为解决SCN无法学习多目标的内在相关性, MT-SCN引入结构矩阵, 利用其对辅助输出线性编码, 探索多输出变量的潜藏信息, 目标函数表示为

$$F(\beta, U) = \min_{\beta, U} \|Y - H\beta U\|_F^2 + \lambda_1 \|\beta\|_F^2 + \lambda_2 \|U\|_{2,1}, \quad (5)$$

其中: U 为结构矩阵, λ_1 和 λ_2 是两个正则化参数. 由于 β 与 U 之间存在耦合性且目标函数是非联合凸, 因此采用交替优化进行求解.

β 和 U 的计算公式如下所示:

$$\beta = \text{Sylvester}(A, B, C), \quad (6)$$

$$U = ((H\beta)^T (H\beta) + \lambda_2 D)^{-1} (H\beta)^T T, \quad (7)$$

其中: $A = \lambda_1 (H^T H)^{-1}$, $B = U U^T$, Sylvester是西尔维斯特方程, $C = (H^T H)^{-1} H^T T U^T$, D 是一个与结构矩阵 U 相关的对角矩阵. 随着节点增加, U 与 β 利用彼此信息交替更新, 进而学习多目标的相关性知识以提高SCN解决MTR问题的能力.

3 多尺度核随机配置网络

3.1 隔离建模机制

在MTSCN构造过程中, 新增节点 L 的隐含层节点参数配置需要满足如下的监督机制约束:

$$\xi_{L,q} = \langle h_L(\beta_L U)_{*,q}, h_L(\beta_L U)_{*,q} \rangle + 2\lambda_1 \langle \beta_{L,q}, \beta_{L,q} \rangle - (1 - r - \mu_L) \langle e_{L-1,q}, e_{L-1,q} \rangle \geq 0, \quad (8)$$

其中: h_L 如式(1)所示, $e_{L-1,q}$ 是第 q 个目标的误差. 新节点从候选节点池中选择一组参数, 使得 $\sum_{q=1}^m \xi_{L,q}$ 最大, 因而作为该节点的输入权值及其偏置. 这里可以看出, 多个目标共用了一组参数 (w_L, b_L) ; 然而, 多个目标的误差状态在建模过程中往往有所差异, 因此基于相同隐层参数的模型在建模多目标时难以满足不同目标的构建需求, 导致多目标模型难以有效减小每个目标的残差.

为了选择符合每个目标特性的最优参数, MSK-SCN建立了隔离建模机制, 为每个目标分配一个专属SCN模型. 根据多目标在建模过程中的不同误差状态, 针对性地为每个目标配置个性化参数以提高节点质量, 进而加快模型拟合多目标的速度. 该策略将多目标分解为多个单目标并对每个目标单独建模. 根据每个目标的自身误差状态选择隐层参数, 使得不同目标的构建过程完全隔绝而没有联系, 避免了多目标共享的隐层参数无法满足每个目标的当前误差状态而降低模型质量的问题. 以第 i 个目标 y_i 为例, 隐含层第 L 个节点的参数配置如下:

$$\xi_{L,i} = \frac{(e_{L-1,i}^T(X) \cdot h_L(X))^2}{(h_L^T(X) \cdot h_L(X) + \frac{1}{\lambda_1})^2 / (h_L^T(X) \cdot h_L(X) + \frac{2}{\lambda_1})} (1 - r - \mu_L) e_{L-1,i}^T(X) e_{L-1,i}(X), \quad (9)$$

其中: $\lambda_1 > 0$ 是正则化参数; $X \in \mathbb{R}^{n \times d}$ 是原始输入; $(w_{L,i}^*, b_{L,i}^*)$ 则是在隔离建模机制下由监督机制筛选的个性化参数, 此时模型仅根据 $q = i$ 的误差状态完成参数配置. 因此, 每个专属模型的新增节点只根据其负责的目标误差状态完成参数配置, 进而加快该目标误差的下降速度, 而不会对其他目标的建模产生影响. 从而当前模型的隐含层输出为

$$H_i = [h_{1,i}(X) \quad \cdots \quad h_{l,i}(X) \quad \cdots \quad h_{L,i}(X)], \quad (10)$$

其中在隔离建模下第 l 个节点对输入 X 的映射输出表示为

$$h_{l,i}(X) = [g_{l,i}(w_{l,i}^{*T} x_1 + b_{l,i}^*) \quad \cdots \quad g_{l,i}(w_{l,i}^{*T} x_n + b_{l,i}^*)]^T.$$

此时对于目标 y_i 的辅助输出为

$$\tilde{y}_{L,i} = \sum_{j=1}^L \beta_{j,i} g_{j,i}(w_{j,i}^{*T} x + b_{j,i}^*), \quad (11)$$

其中: $\beta_{j,i}$ 表示第 j 个节点的输出权值, $w_{j,i}^*$ 和 $b_{j,i}^*$ 是在隔离建模机制下针对目标 y_i 筛选出的个性化输入权值和偏置. 当最终输出误差未达到预先设定的标准, 则继续增加节点以更新所学相关信息进而减小误差. 关于式(10)中输出权值的求解, 采用全局更新策略以加

快模型的收敛速度.

$$\beta_i^* = \arg \min_{\beta_i} \|y_i - \tilde{y}_{L,i}\| + \lambda_1 \left\| \sum_{j=1}^L \beta_{j,i} \right\|, \quad (12)$$

其中 λ_1 是正则化参数. 因此目标 y_i 的辅助输出又可以表达成 $\tilde{y}_i = H_i \beta_i^*$. 在隔离机制下, 其余目标与 y_i 的个性化参数选择过程一样. 在未达到停止条件时, 通过自身专属模型确定新增节点的最优参数. 因此, 当 m 个目标完成个性化参数配置后, 基于全局更新策略得到新的辅助输出, 并组成辅助输出层, 即

$$\tilde{Y} = [\tilde{y}_1 \ \tilde{y}_2 \ \cdots \ \tilde{y}_m] = [H_1 \beta_1^* \ H_2 \beta_2^* \ \cdots \ H_m \beta_m^*]. \quad (13)$$

3.2 多尺度核空间

隔离建模机制为每个目标筛选个性化参数, 其本质是在构建共享输入与每个目标的专一映射, 进而建模输入到输出之间的关系. 然而每个目标的建模过程完全处于封闭状态而没有任何联系, 导致模型无法学习多目标之间的相关性. 为建模多目标之间的非线性结构知识, MSK-SCN基于多个尺度参数的同类核函数组成了多尺度核空间, 利用核函数对应的非线性映射将原始的低维空间投射到高维甚至无限维的特征空间, 以提高信息的表达能力, 进而保证多目标变量之间的复杂非线性关系的挖掘能力. 并且不同尺度参数的核对应着一个特定的样本关系学习任务, 因此可以根据目标的数量确定多尺度核的数量上界. MSK-SCN的目标函数可以定义如下:

$$F(U) = \arg \min_U \|Y - KU\|_F^2 + \lambda_2 \|U\|_F^2, \quad (14)$$

其中: Y 表示真实输出, U 是结构矩阵, $\lambda_2 > 0$ 为正则化参数, K 是多尺度核组成的增广核矩阵, 即

$$K = \sum_{i=1}^m s_i \cdot K_i(\tilde{Y}, Y) = \sum_{i=1}^m s_i \cdot \langle \varphi_i(\tilde{Y}), \varphi_i(Y) \rangle_F, \quad (15)$$

其中: K_i 对应着第 i 个尺度参数的核矩阵, $\varphi_i(\cdot)$ 表示将变量从原始空间转换到泛函希尔伯特空间 F 的非线性映射函数, s_i 为 K_i 在建模中的对应权值. 最终输出是经过结构矩阵对增广核矩阵进行信息编码而获得的, 即

$$\hat{Y} = KU = [\hat{y}_1 \ \hat{y}_2 \ \cdots \ \hat{y}_m]. \quad (16)$$

为充分挖掘多目标之间的非线性结构知识, 设计了基于不同尺度参数的核矩阵增广形成多尺度核空间. 为度量不同尺度参数的核在不同任务场景下的作用, 假定目标空间的Gram矩阵为基准核, 如 $K_i(Y, Y)$ 其作为对不同参数核拟合效果的参考基准. 当每个目标在隔离机制下完成模型更新后, 多尺度核空间分别

学习辅助输出和目标空间的结构知识, 并且基于余弦相似度刻画在不同尺度参数下辅助输出的核矩阵对于基准核的接近程度, 进而表征不同尺度参数核在建模过程中的贡献度. 当余弦相似度值越大, 则表明该尺度参数对应的核更加适合该场景的建模任务, 反之该核对任务贡献较小. 尺度参数 i 对应的核矩阵 K_i 与其基准核之间的余弦相似度计算公式为

$$\cos \theta_i = \frac{\langle \text{vec}(K_i(\tilde{Y}, Y)), \text{vec}(K_i(Y, Y)) \rangle}{\|K_i(\tilde{Y}, Y)\| \cdot \|K_i(Y, Y)\|} = \frac{\text{tr}(K_i(\tilde{Y}, Y)^T K_i(Y, Y))}{\|K_i(\tilde{Y}, Y)\| \cdot \|K_i(Y, Y)\|}, \quad (17)$$

其相应的权重比值是

$$s_i = \begin{cases} 1, & \cos \theta_i \geq \sigma, \\ 0, & \cos \theta_i < \sigma, \end{cases} \quad (18)$$

其中 σ 是黄金比例系数0.618. 在建模过程中, 当某尺度核与基准核间的近似度未达到预设标准, 则表明该尺度对应的核空间无法有效提取到多输出之间的相关信息, 对其权值赋0, 在后续建模中不再使用.

辅助输出层经过不同尺度的非线性核映射到高维特征空间后, 可以有效构建多目标内在的非线性结构知识的提取通道. 式(14)关于 U 的导数为

$$\nabla F(U) = -K^T(Y - KU) + \lambda_2 U. \quad (19)$$

利用梯度寻优策略, U 的迭代公式为

$$U^{t+1} = U^t - \eta \nabla F(U^t), \quad (20)$$

其中: U^t 是第 t 次迭代的结构矩阵值, η 是学习率. MSK-SCN的增量构造过程如图1所示, 其中 \oplus 是求和符号, 通过的不断调整 β 和 U 实现对高维输入到多输出以及多目标内在的关系学习.

3.3 性能分析

定理 1 假设 $\text{span}(\Gamma)$ 在 L_2 空间稠密, 且 $\forall g \in \Gamma$, $0 < \|g\| < b_g$, 其中 $b_g \in \mathbb{R}^+$. 对于一个给定的实数 c , 存在一个足够小的正实数 ε . 给定 $1 - \varepsilon < r < 1$, 一个非负实数序列 μ_L 满足 $1 - \varepsilon - r \leq \mu_L \leq 1 - r$. 对于 $L = 1, 2, \dots$ 和 $q = 1, 2, \dots, m$, 做出如下定义^[24]:

$$\delta_{L,q}^* = (1 - r - \mu_L) \left\| y - \sum_{j=1}^L \beta_j^* g_j \right\| = (1 - r - \mu_L) \|e_{L,q}^*\|, \quad (21)$$

$$\xi = \lim_{L \rightarrow \infty} (1 - \varepsilon)^L < 1, \quad (22)$$

$$\gamma = \lim_{L \rightarrow \infty} \prod_{k=1}^L (r + \mu_k) < 1. \quad (23)$$

若激活函数 g_L 满足下式:

$$\langle e_{L-1,q}^*, g_L \rangle^2 \geq \frac{(b_g^2 + \frac{1}{\lambda_1})}{(b_g^2 + \frac{1}{\lambda_1})} \delta_{L,q}^*. \quad (24)$$

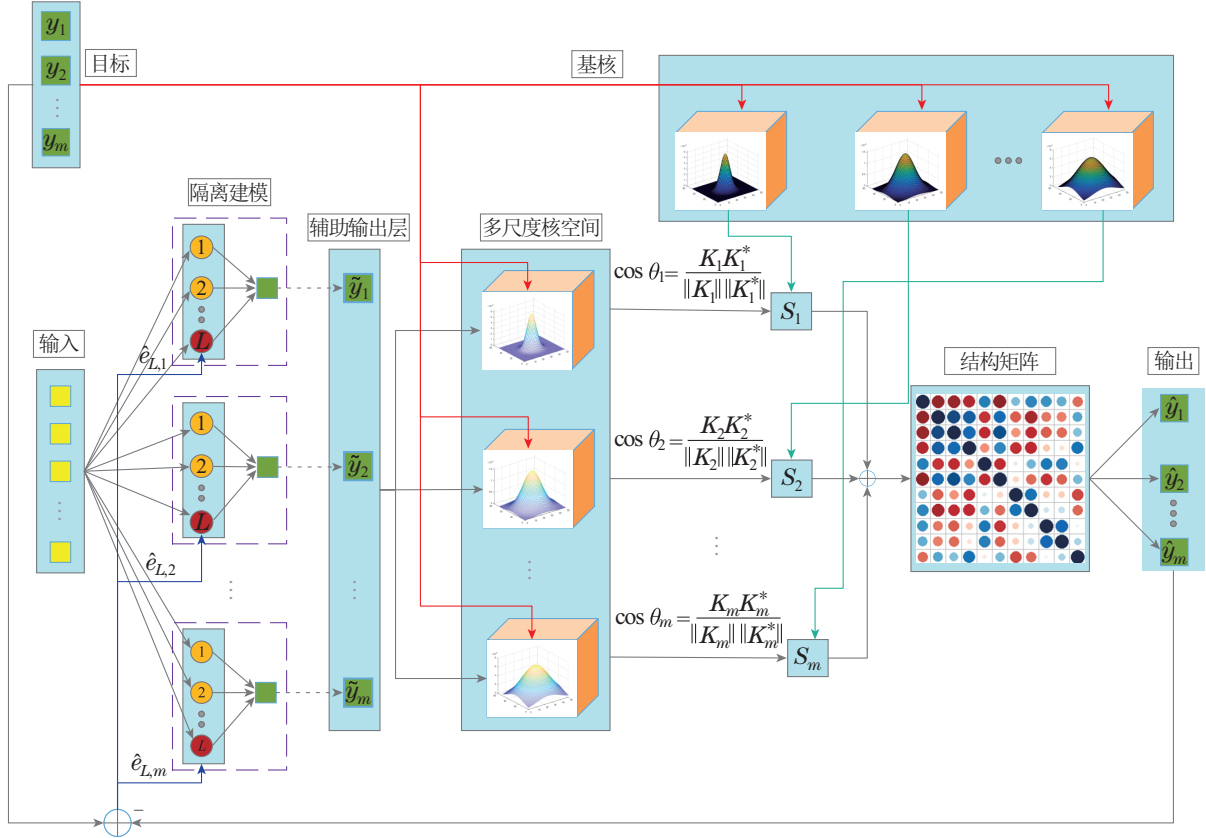


图 1 多尺度核随机配置网络结构

Fig. 1 Multi-scale kernel stochastic configuration network structure

输出权值以如下公式计算:

$$[\beta_1 \ \beta_2 \ \dots \ \beta_L] = \arg \min_{\beta} \|y - \sum_{j=1}^L \beta_j g_j\|^2 + \frac{1}{2} \|\beta\|_2^2. \quad (25)$$

那么有 $\sqrt{\xi} \|y\| \leq \lim_{L \rightarrow \infty} \|e_L^*\| \leq \sqrt{\gamma} \|y\|$.

在隔离建模机制下, 每个单目标的辅助输出对应的输出权值都是采用全局更新策略的, 如式(12)所示. 以第*i*个目标为例, 根据定理1, 在全局更新输出权值下的误差最终会稳定在一个误差带内, 即

$$\sqrt{\xi} \|y_i\| \leq \lim_{L \rightarrow \infty} \|e_{i,L}^*\| \leq \sqrt{\gamma} \|y_i\|. \quad (26)$$

式(26)对于隔离建模下的其他目标仍然成立.

通过高维映射后得到的核矩阵提高了模型对多目标非线性结构信息的挖掘能力, 再经过结构矩阵*U*的编码得到最终的预测输出, 在此过程中仅需要优化结构矩阵*U*. 在模型的增量构建中, 对*U*进行梯度寻优策略, 其更新方式如式(19)–(20)所示.

引理 1 给定函数*f*, 若存在常数 $M \geq 0$, 使得任意*u*和*v*有 $\|f(u) - f(v)\| \leq M \|u - v\|$, 则*f*满足利普希茨连续条件^[25].

定理 2 令*F*是MSK-SCN的目标函数(式(15)), 且它的梯度满足利普希茨连续条件. 当 $0 \leq \eta \leq \frac{2}{M}$

(*M*为利普希茨常数), 采用式(19)–(20)的方式更新结构矩阵*U*, 则输出误差 $\hat{E} = \|Y - KU\|$ 是单调下降的.

证 由于目标函数*F*是可微的凸函数, 故将*F*围绕*F(v)*进行泰勒二次展开可得如下不等式:

$$F(u) \leq F(v) + \nabla F(v)^T (u - v) + \frac{1}{2} \nabla^2 F(v) \|u - v\|^2, \quad (27)$$

又因其梯度 ∇F 满足利普希茨连续, 进而有 $\nabla^2 F(v) \leq MI$, 其中*I*是单位矩阵, 则式(27)进一步有

$$F(u) \leq F(v) + \nabla F(v)^T (u - v) + \frac{M}{2} \|u - v\|^2. \quad (28)$$

式(14)中*U*的第*t* + 1次迭代记为*U*^{*t*+1}, 令

$$\begin{cases} u = U^{t+1} = U^t - \eta \nabla F(U^t), \\ v = U^t, \end{cases} \quad (29)$$

其中 η 为梯度迭代的步长.

将式(29)代入式(28), 则有

$$F(U^{t+1}) \leq F(U^t) + \nabla F(U^t)^T (U^{t+1} - U^t) + \frac{M}{2} \|U^{t+1} - U^t\|^2 =$$

$$\begin{aligned}
& F(U^t) - \eta \|\nabla F(U^t)\|^2 + \\
& \frac{M\eta^2}{2} \|\nabla F(U^t)\|^2 = \\
& F(U^t) + \eta \left(\frac{M\eta}{2} - 1 \right) \|\nabla F(U^t)\|^2.
\end{aligned} \tag{30}$$

可知 $\|\nabla F(U^L)\|^2 \geq 0$, 因此当 $\eta \left(\frac{M\eta}{2} - 1 \right) \leq 0$ 时 (即 $0 \leq \eta \leq \frac{2}{M}$), 目标函数值始终保持单调递减, 即

$$F(U^{L+1}) \leq F(U^L) \leq \dots \leq F(U^1), \tag{31}$$

其中 U^L 是模型节点增加到 L 时对应的结构矩阵. 同时随着迭代次数的增加, U 会渐近稳定而不再变化, 即

$$\lim_{t \rightarrow \infty} \|U^{t+1} - U^t\| = 0. \tag{32}$$

因此在模型增量构造中, 式(14)的多目标误差会单调下降

$$\hat{E}_1 \geq \hat{E}_2 \geq \dots \geq \hat{E}_L, \tag{33}$$

其中 \hat{E}_L ($[\hat{e}_{L,1} \ \hat{e}_{L,2} \ \dots \ \hat{e}_{L,m}]$) 是隐层节点达到 L 时模型的多目标误差. 基于全局更新策略和梯度下降方式, 保证了辅助输出和预测输出的残差为单调递减.

证毕.

3.4 复杂度分析

假设训练具有 L 个节点的 SCN 的时间复杂度为 $O(L)$, 隔离建模机制为每个目标分配一个专属于模型构建辅助输出层, 使得 MSK-SCN 的时间复杂度为 $O(mL)$. 辅助输出层由 m 个子模型的输出组成, 其空间复杂度为 $O(n)$, 其中 n 表示样本数. 同时每个尺度参数对应的核矩阵的空间复杂度是 $O(n)$, 因此多尺度核空间的空间复杂度为 $O(mn)$.

4 仿真实验

为充分证明所提方法在多目标建模任务上的有效性, 本部分从虚拟数据集和公开数据集两方面比较 MSK-SCN 与当前 MTR 算法的仿真表现, 包括 SST (stacked single target)^[11], ERC (ensemble of regressor chains)^[11], MTSCN (multi-target regression stochastic configuration network)^[9], MSLR (multitarget sparse latent regression), MMR (multi-layer multi-target regression)^[18] 以及 MTR-ITC (multi-target sparse regression with instances and targets correlations)^[19]. 并且通过均方根误差 RMSE 和决定系数 R^2 评价指标以衡量不同方法的性能, 所有实验在 Intel Core i5 CPU (3.0 GHz) 8 GB RAM 的计算机中通过 Python 3.8 进行实现.

4.1 虚拟数据集实验

本部分设计了线性与非线性关系混合的自构数据集, 验证 MSK-SCN 对输入到输出之间的联系以及多输出内在的相关性的建模效果. 并揭示了多尺度核探

索机制可以充分挖掘到多输出内在的线性关系以及复杂的非线性结构知识.

4.1.1 虚拟数据集构造

假设共享输入 $X \in \mathbb{R}^d$ 来源于多元高斯分布, 即 $X \sim N(\mu, \Sigma)$, 其中 μ 是均值向量, Σ 是协方差矩阵. 为了构建多目标之间的非线性结构知识, 基于非线性函数映射输入并以串级方式构造输出目标, 即

$$\begin{cases}
y_1 = f_1(X), \\
y_2 = f_2(f_1(X)), \\
\vdots \\
y_m = f_m(f_{m-1}(\dots f_1(X))),
\end{cases} \tag{34}$$

其中: f 是使用的非线性函数 e^{-x^2} , $\cos x$ 及 $\sin x$; m 是多目标的数量, 虚拟数据集在实验中进行归一化.

4.1.2 参数设置

单核函数在捕捉数据非线性关系方面存在局限, 尤其在面对数据的复杂非线性结构时. 多尺度核学习通过灵活选择尺度参数, 构建高维特征空间, 增强了模型对数据内在复杂性的学习能力. 然而, 核矩阵的计算和存储成本较高, 不匹配的尺度参数可能导致资源浪费^[26]. 目前, 多核学习在选择核函数和尺度参数方面缺乏成熟的理论指导, 通常依赖实验确定. 设置与目标个数相等的多尺度核可能在理论上增强模型的学习能力, 但也增加了引入对建模帮助不大或产生负面影响的核的概率. 因此, 核函数和尺度参数的优化选择是提升多核学习性能的关键. 为平衡建模速度与模型性能, 根据实验的结果, 在仿真中最终只保留两个尺度参数对应的核矩阵. 根据文献[23]中, 较小的尺度参数适合剧烈变化的样本分类, 而较大的尺度参数偏向平缓变化的样本分类任务, 因此, 多尺度核空间的尺度参数经过预设参数 $\{\frac{1}{n}, 0.1, 0.5, 1, 2\}$ 的线性搜索而确定. 而 MSK-SCN 的其他参数的设置如表 1 所示, 隐层节点数的最大值 L_{\max} 设为 120, 每个节点对应最大的配置次数 T_{\max} 设为 20. 正则化参数 λ_1 和 λ_2 通过线性搜索获得, 且其范围在 $[0 : 0.1 : 10]$, 而梯度下降的步长范围在 $[0 : 0.01 : 1]$. 在虚拟数据集仿真中, SST 和 ERC 使用 RBF 核的支持向量回归模型作为基础学习器, 并且核带宽设置为 0.1, 其中根据文献[11]的结论, ERC 生成 10 条随机链; SCN 的 $L_{\max} = 120$, $T_{\max} = 10$, γ 的范围是 $\{0.3, 0.8, 1, 3, 10\}$; MTS-SCN 的最大节点数 $L_{\max} = 120$, γ 的范围设置为 $\{0.4, 0.6, 1, 3, 10\}$, λ_1 与 λ_2 在 $[0 : 0.1 : 10]$ 线性搜索得到; MSLR 的 λ 和 β 分别由 $10^{[-4:1]}$ 线性搜索; MMR 的 λ , β 和 γ 分别在区间 $10^{[-3:1:0]}$ 线性搜索; MTR-ITC 的正则化参数 λ_1 和 λ_2 分别固定为 10^{-2} 和 10^{-2} , λ_3 和 λ_4 在区间 $10^{[-3:1:4]}$ 线性搜索得到.

表 1 MSK-SCN在虚拟数据集上参数设置

Table 1 Parameters settings of MSK-SCN on the synthetic dataset

| 目标个数 | 参数 | | | | | |
|------|---------------|------------|-------------------|-------------|-------------|--------|
| | σ_1 | σ_2 | \mathcal{Y} | λ_1 | λ_2 | η |
| 5 | $\frac{1}{n}$ | 0.1 | {0.3,0.5,0.8,1,3} | 2 | 0.1 | 0.05 |
| 10 | $\frac{1}{n}$ | 0.1 | {0.4,0.6,0.8,1,3} | 2 | 0.1 | 0.05 |
| 15 | $\frac{1}{n}$ | 0.1 | {0.4,0.8,1,3,10} | 4 | 0.1 | 0.05 |

4.1.3 结果分析

虚拟数据集的相关系数如图2所示, 其目标间非线性关系比重较高, 这为MTR算法准确建模带来极大挑战. 表2展示了多种MTR方法在3组虚拟数据集上的RMSE结果, 其中SST和ERC由于简单地将预测值堆叠到输入空间期望获得多目标的相关性, 难以保证模型能够学习到准确的非线性关系, 使得其预测效果并不理想. 而SCN在建模中未考虑多目标间的相关性, 导致SCN在几种MTR法中拟合表现最差. MTSCN引入的结构矩阵利用线性方式编码多目标的内在联系, 使得建立的模型难以充分挖掘非线性结构信息, 导致其对SCN在MTR应用上的提升有限, 并且MSK-SCN在前两组虚拟数据集上相较于MTSCN的准确度有较高提升. MSLR与MMR的拟合多目标效果与基于堆叠方法的SST和ERC相差不大. 而MTR-ITC在几种MTR方法中表现较好, 但MSK-SCN相对于其建模准确度仍然有较大提升. 基于上述分析, MSK-SCN构建多尺度核空间有效提高了模型挖掘非线性结构信息的能力, 其拟合多目标的效果优于其他MTR算法.

表 2 不同算法在虚拟数据集上以RMSE(%)为指标的结果

Table 2 Results of different algorithms with RMSE(%) as the index on the synthetic dataset

| 目标个数 | 方法 | | | | | | | |
|------|-------|-------|-------|-------|-------|--------|---------|---------|
| | SST | ERC | SCN | MTSCN | MSLR | MMRSCN | MTR-ITC | MSK-SCN |
| 5 | 5.91 | 6.26 | 11.90 | 5.74 | 6.45 | 6.82 | 4.96 | 1.72 |
| 10 | 6.31 | 7.93 | 12.91 | 6.62 | 6.38 | 8.48 | 5.21 | 1.63 |
| 15 | 30.86 | 30.42 | — | — | 33.11 | 34.04 | 38.78 | 8.25 |

当建模目标个数达到15时, 所有对比算法的性能剧烈下降, 虽然MSK-SCN的效果也有所下降, 但仍然保持较高的准确度. 图3进一步展示了SCN类模型在其上的表现结果, 其中SCN和MTSCN分别在第20个节点和第50个节点附近发生过拟合. 这也进一步证实了原始SCN无法直接用于MTR问题上, 以及结构矩阵对于SCN的提升作用有限. 而基于多尺度核空间的MSK-SCN在节点数为100左右时模型达到收敛, 并且训练误差和测试误差非常接近, 模型具有良好的泛化

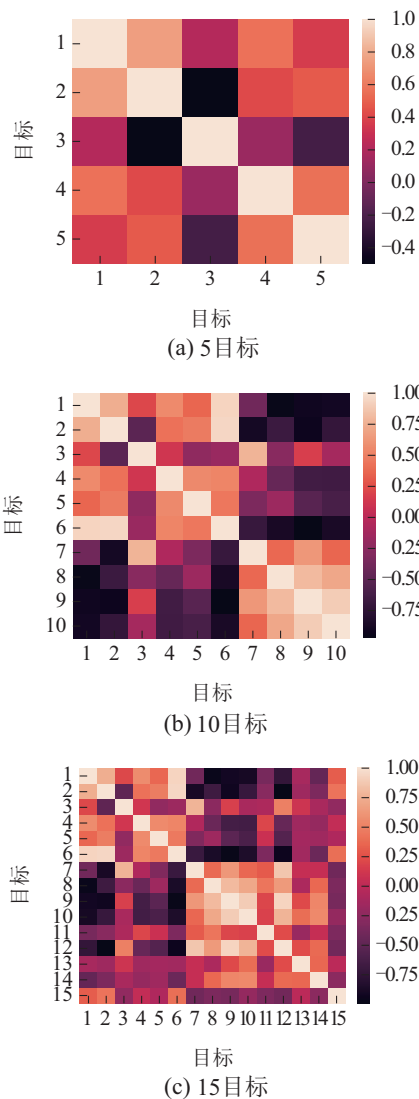


图 2 虚拟数据集的相关系数

Fig. 2 Correlation coefficient of the synthetic dataset

能力. 这有效验证了多尺度核探索机制对于多输出之间的复杂非线性关系的学习能力.

表2反映了MTR模型的宏观逼近效果, 为进一步分析不同算法对每个目标的拟合程度, 图4展示了4种MTR模型对虚拟数据集以 R^2 为指标的拟合效果, 其决定系数越接近1, 表明模型拟合程度越高. 基于结构矩阵与堆叠预测值的方法并不能对每个目标都有效逼近, 甚至在某些目标上其 R^2 为负值, 这也进一步表明其对于多目标之间的复杂关系无法有效学习. 然

而MSK-SCN在3组虚拟数据上,一直可以准确拟合每个目标,充分证明多尺度核探索机制对于挖掘多输出内在相关性的有效性.

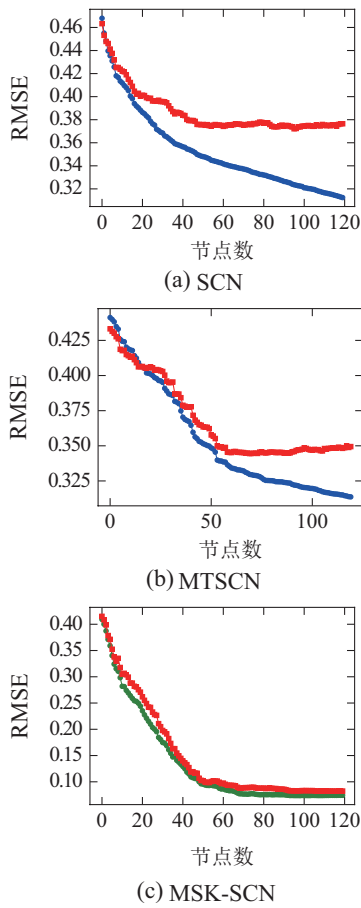


图3 SCN类模型在第3组虚拟数据集上的误差曲线

Fig. 3 The error curves of the SCN series models on the third set of synthetic datasets

4.2 真实数据集实验

本部分在真实数据集上进行实验,与当前MTR算法对比,进一步证明所提方法在建模高维输入与多输出之间的联系以及多输出内在相关性上的有效性.

4.2.1 真实数据集

真实数据集涉及价格与能源预测等应用领域,为MTR建模的普适性带来一定挑战,其来源于 <http://mulan.sourceforge.net/datasets-mtr.html>. 真实数据集目标间相关系数如图5所示,其统计信息如表3所示,并且所有数据在实验中进行归一化处理.

由于核矩阵的计算开销过大,为平衡建模速度与模型性能,与在虚拟数据集上的实验一样,仍然保留两个尺度参数对应的核矩阵. 进而MSK-SCN的重要参数在不同数据集上的设置如表4所示. 隐层节点数的最大值 L_{max} 设为100,每个节点对应的最大的配置次数 T_{max} 设为20. 正则化参数 λ_1 和 λ_2 在 $[0 : 0.1 : 10]$ 区间线性搜索获得,而梯度下降的步长在 $[0 : 0.01 : 1]$ 基于线性搜索得到. 在公开数据集上, SST 和 ERC 使

用带宽0.1的RBF核的支持向量回归模型作为基础学习器,并且ERC生成10条随机链; SCN的 $L_{max} = 120$, $T_{max} = 10$, γ 的范围是 $\{0.5, 0.8, 1, 3, 10\}$; MTSCN的最大节点数 $L_{max} = 120$, γ 的范围设置为 $\{0.3, 0.5, 1, 3, 10\}$, λ_1 与 λ_2 在 $[0 : 0.1 : 10]$ 线性搜索得到; MSLR的 λ 和 β 分别由 $10^{[-4:1:1]}$ 线性搜索; MMR的 λ , β 和 γ 分别在区间 $10^{[-3:1:0]}$ 线性搜索; MTR-ITC的正则化参数 λ_1 与 λ_2 分别固定为 10^{-1} 和 10^{-1} , λ_3 和 λ_4 在区间 $10^{[-3:1:4]}$ 线性搜索得到.

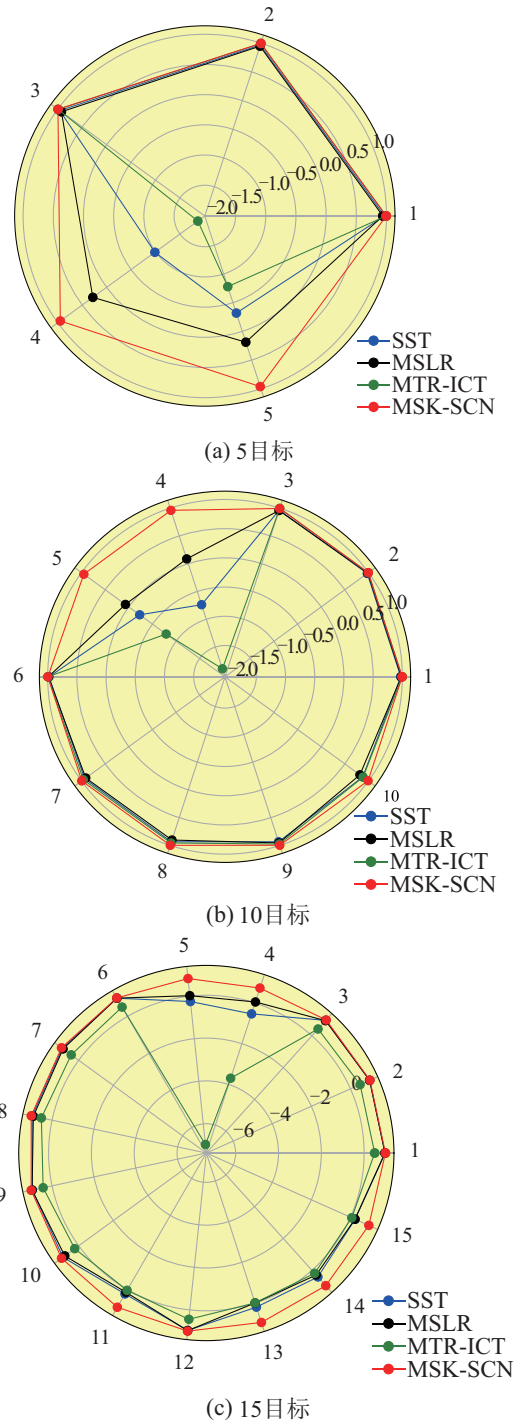


图4 4种MTR算法以 R^2 为目标拟合表现

Fig. 4 The target fitting performances of the four MTR algorithms with R^2 scores

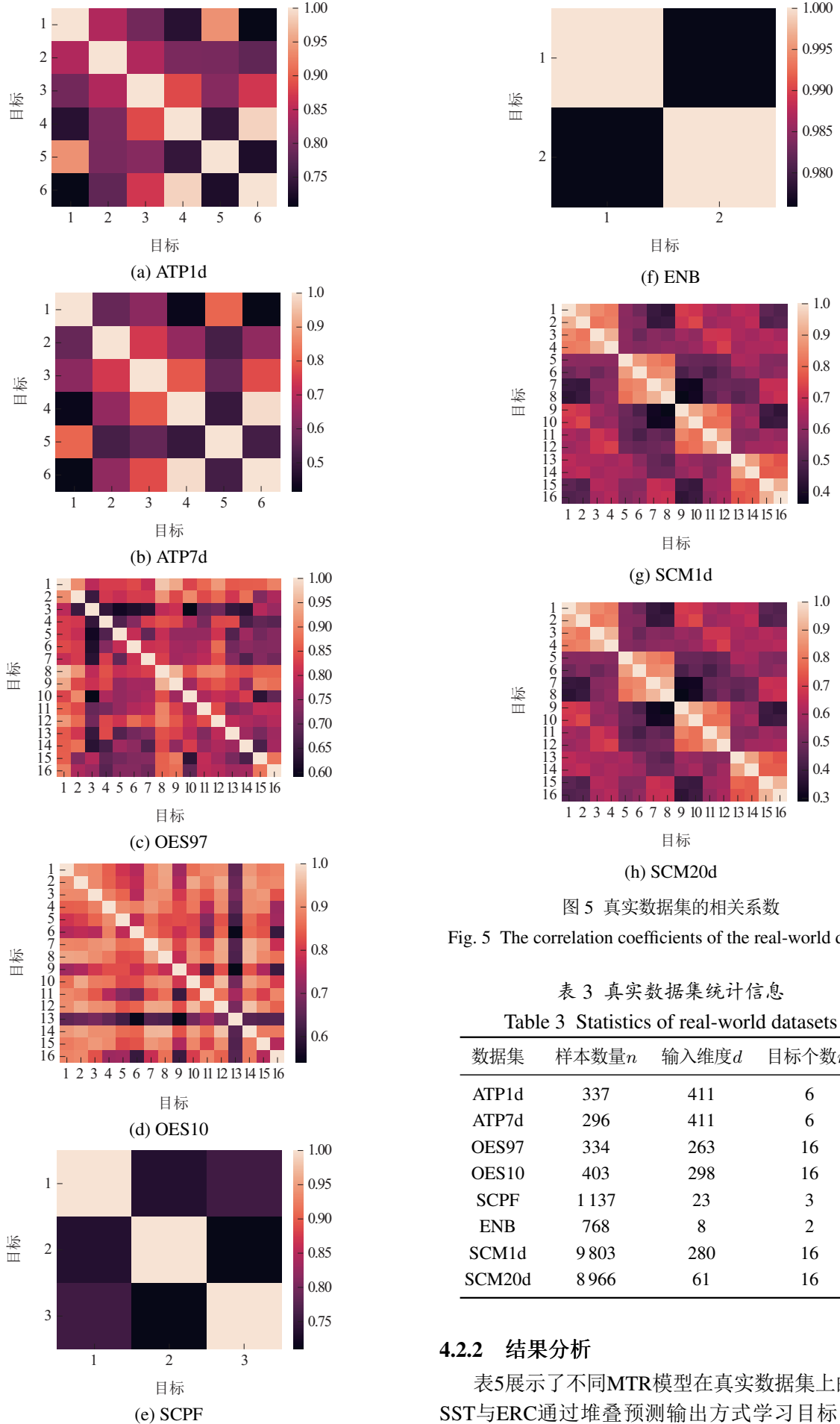


图 5 真实数据集的相关系数

Fig. 5 The correlation coefficients of the real-world datasets

表 3 真实数据集统计信息

Table 3 Statistics of real-world datasets

| 数据集 | 样本数量 n | 输入维度 d | 目标个数 m |
|--------|----------|----------|----------|
| ATP1d | 337 | 411 | 6 |
| ATP7d | 296 | 411 | 6 |
| OES97 | 334 | 263 | 16 |
| OES10 | 403 | 298 | 16 |
| SCPF | 1 137 | 23 | 3 |
| ENB | 768 | 8 | 2 |
| SCM1d | 9 803 | 280 | 16 |
| SCM20d | 8 966 | 61 | 16 |

4.2.2 结果分析

表5展示了不同MTR模型在真实数据集上的结果, SST与ERC通过堆叠预测输出方式学习目标间相关

性,然而依赖于第1次建模的模型效果,导致其仅在一定程度上提高了模型的拟合表现.其中SCN在建模过程中未考虑多目标的相关性,导致其在各数据集上的拟合表现较差.而MTSCN利用线性编码方式提高了SCN学习多目标相关信息的能力,进而提高了模型拟合多目标的效果.MSLR,MMR和MTR-ITC同样使用结构矩阵方式学习多目标的结构信息,然而受限于结构矩阵的学习能力,导致其难以有效挖掘到多目标之间更加复杂的非线性结构知识,使得其对多目标

的预测效果与SST和ERC相当.但是MSK-SCN在真实数据集上的表现超越其他MTR模型,相较于MTSCN的准确度分别提升了19.19%,16.93%,24.21%,17.09%,3.4%,8.9%,17.17%,16.33%,这充分证明了隔离建模对于构建高维输入到多输出间的联系以及多尺度核空间对于挖掘多目标间的相关性尤其是复杂的非线性结构知识的有效性.同时,所提方法在不同的多目标建模任务上均取得出色表现,具有一定通用性,能够胜任不同应用领域下的建模任务.

表4 MSK-SCN在真实数据集的参数设置
Table 4 Parameter settings of MSK-SCN on the realworld datasets

| 目标 | 参数 | | | | | |
|--------|---------------|------------|--------------------|-------------|-------------|--------|
| | σ_1 | σ_2 | γ | λ_1 | λ_2 | η |
| ATP1d | $\frac{1}{n}$ | 0.1 | {0.4,0.8,1,3,10} | 1 | 0.5 | 0.01 |
| ATP7d | $\frac{1}{n}$ | 0.1 | {0.4,0.8,1,3,10} | 0.5 | 0.5 | 0.01 |
| OES97 | $\frac{1}{n}$ | 0.1 | {0.3,0.5,0.8,3,10} | 1 | 0.1 | 0.03 |
| OES10 | $\frac{1}{n}$ | 0.1 | {0.2,0.5,0.8,1,3} | 1 | 0.1 | 0.02 |
| SCPF | $\frac{1}{n}$ | 0.1 | {0.3,0.6,1,3,10} | 4 | 0.2 | 0.05 |
| ENB | $\frac{1}{n}$ | 0.1 | {0.4,0.6,0.8,1,3} | 2 | 0.1 | 0.05 |
| SCM1d | $\frac{1}{n}$ | 0.1 | {0.4,0.6,0.8,1,3} | 2 | 0.1 | 0.02 |
| SCM20d | $\frac{1}{n}$ | 0.1 | {0.4,0.6,0.8,1,3} | 2 | 0.1 | 0.05 |

表5 不同算法在真实数据集上以RMSE(%)为指标的结果
Table 5 Results of different algorithms on the real-world datasets with RMSE(%) as index

| 数据集 | 方法 | | | | | | | |
|--------|-------|-------|-------|-------|-------|-------|---------|---------|
| | SST | ERC | SCN | MTSCN | MSLR | MMR | MTR-ITC | MSK-SCN |
| ATP1d | 8.74 | 8.75 | 12.12 | 9.90 | 9.14 | 8.44 | 8.37 | 8.00 |
| ATP7d | 12.14 | 12.19 | 14.02 | 12.52 | 11.95 | 11.66 | 13.62 | 10.40 |
| OES97 | 7.61 | 7.67 | 8.64 | 5.70 | 5.38 | 5.13 | 4.90 | 4.32 |
| OES10 | 7.67 | 7.73 | 7.20 | 5.50 | 5.07 | 5.06 | 5.05 | 4.56 |
| SCPF | 8.35 | 8.46 | 3.19 | 2.59 | 3.49 | 2.60 | 2.92 | 2.50 |
| ENB | 9.20 | 9.48 | 10.81 | 8.61 | 8.55 | 8.84 | 10.92 | 7.84 |
| SCM1d | 6.35 | 6.64 | 7.12 | 6.58 | 9.51 | 10.99 | 10.24 | 5.45 |
| SCM20d | 7.97 | 9.08 | 8.19 | 9.12 | 11.04 | 10.72 | 10.31 | 7.63 |

图6则进一步展示了MTR模型在4个真实数据集上对每个目标的拟合程度.从图5呈现的输出间相关系数可以看出,真实数据集整体上线性程度较高,因此其他算法与MSK-SCN的逼近效果相差不大.基于结构矩阵的模型难以对每个数据集上所有目标进行准确的逼近,如MMR在某些目标的拟合上甚至出现负值,这也表明了结构矩阵难以学习到目标之间的非

线性相关知识.同样,SST的多目标拟合效果也不理想,因为其依赖初次模型的效果,同时基于特征扩展的方式在学习复杂的相关性上能力孱弱.所提的MSK-SCN在真实数据集上对每个目标的拟合效果达到最佳,这归功于多尺度核探索机制能够有效挖掘到多目标之间的联系进而提高模型性能.从宏观以及微观层面呈现出的多目标的拟合结果,MSK-SCN有效

拓展了SCN在MTR上的应用能力.

5 结论

本文所提的MSK-SCN能够同时对高维输入与多输出之间的联系以及多目标内在的相关性进行探索建模, 有效提高了SCN解决MTR问题的能力. 隔离建模机制能够为每个目标提供个性化参数, 从而避免隐

层参数陷入局部最优. 多尺度核探索机制通过不同尺度参数的核组合形式, 可以充分挖掘到多目标内部的复杂的相关知识如非线性结构关系, 有效提高了模型的学习能力. 最后, 通过虚拟数据集以及真实数据集的实验仿真, 所提方法的表现优于当前的MTR模型, 进一步证明了其可以有效获取多输出内在关系, 提高了模型在MTR问题上的应用能力.

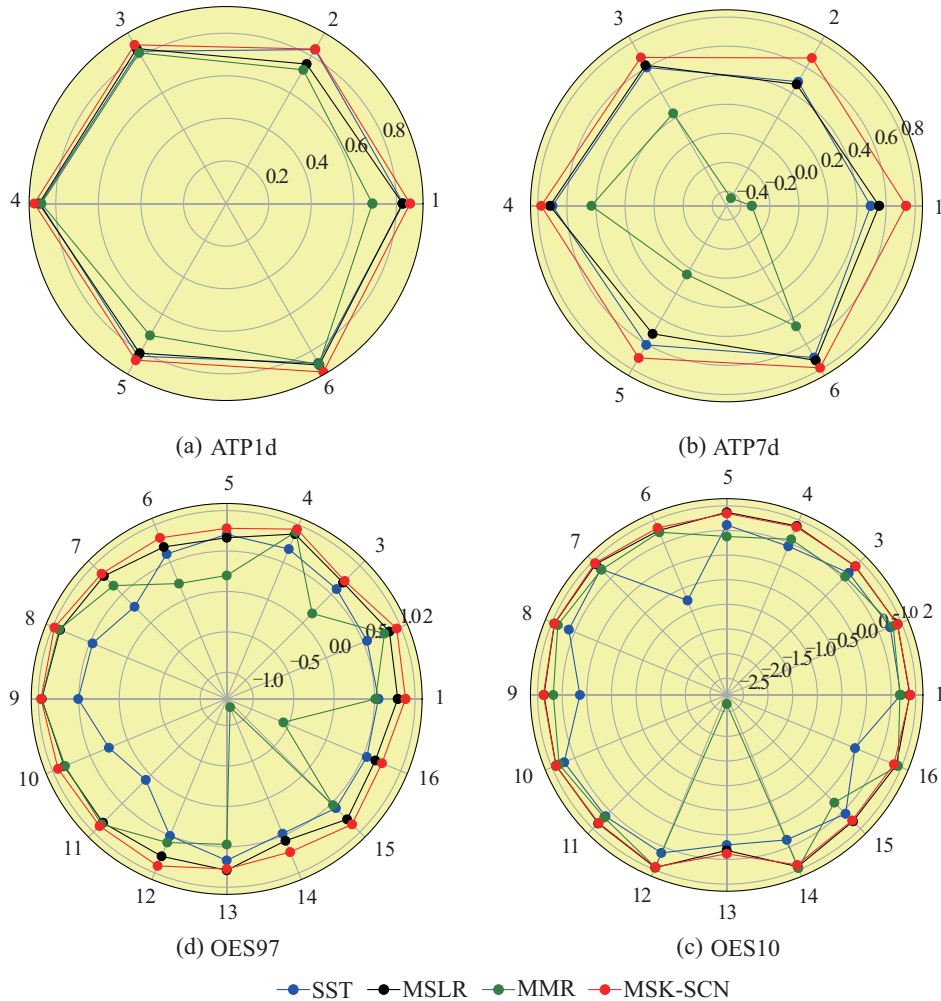


图6 4种MTR算法对每个目标以 R^2 为指标的拟合表现

Fig. 6 The target fitting performances of the 4 MTR algorithms with R^2 scores

参考文献:

[1] HU Gensheng, DENG Feiqi. Support vector regression with piecewise loss function and its application in investment decision. *Control Theory & Applications*, 2006, 23(2): 315 – 318.
(胡根生, 邓飞其. 具分段损失函数的支持向量机回归及在投资决策中的应用. *控制理论与应用*, 2006, 23(2): 315 – 318.)

[2] ZHOU Ping, ZHANG Li, LI Wenpeng, et al. Autoencoder and PCA based RVFLNs modeling for multivariate molten iron quality in blast furnace ironmaking. *Acta Automatica Sinica*, 2018, 44(10): 1799 – 1811.
(周平, 张丽, 李温鹏, 等. 集成自编码与PCA的高炉多元铁水质量随机电神经网络建模. *自动化学报*, 2018, 44(10): 1799 – 1811.)

[3] PAO Y H, TAKEFUJI Y. Functional-link net computing: Theory, system, architecture, and functionalities. *Computer*, 1992, 25(5): 76 – 79.

[4] LI Wenpeng, ZHOU Ping. Robust regularized RVFLNs modeling of molten iron quality in blast furnace ironmaking. *Acta Automatica Sinica*, 2020, 46(4): 721 – 733.
(李温鹏, 周平. 高炉铁水质量鲁棒正则化随机神经网络建模. *自动化学报*, 2020, 46(4): 721 – 733.)

[5] TYUKIN I Y, PROKHOROV D V. Feasibility of random basis function approximators for modeling and control. *IEEE Control Applications, (CCA) & Intelligent Control, (ISIC)*. St.Petersburg, Russia: IEEE, 2009: 1391 – 1396.

[6] WANG D H, LI M. Stochastic configuration networks: Fundamentals and algorithms. *IEEE Transactions on Cybernetics*, 2017, 47(10): 3466 – 3479.

[7] REYES O, VENTURA S. Performing multi-target regression via a parameter sharing-based deep network. *International Journal of Neural Systems*, 2019, 29: 1950014.

- [8] XU D N, SHI Y X, TSANG I W, et al. Survey on multi-output learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(7): 2409 – 2429.
- [9] WANG Q J, HONG Q Q, WU S, et al. Multitarget stochastic configuration network and applications. *IEEE Transactions on Artificial Intelligence*, 2023, 4(2): 338 – 348.
- [10] BORCHANI H, VARANDO G, BIELZA C, et al. A survey on multi-output regression. *Wires Data Mining and Knowledge Discovery*, 2015, 5(5): 216 – 223.
- [11] SPYROMITROS X E, TSOUMAKAS G, GROVES W, et al. Multi-target regression via input space expansion: Treating targets as inputs. *Machine Learning*, 2016, 104: 55 – 98.
- [12] SANTANA E J, MASTELINI S M, BARBON S J. Deep regressor stacking for air ticket prices prediction. *Brazilian Symposium on Information Systems*, 2017, 1: 25 – 31.
- [13] WANG J, CHEN Z, SUN K, et al. Multi-target regression via target specific features. *Knowledge-Based Systems*, 2019, 170: 70 – 78.
- [14] LIU Hongtao, LI Hang, WANG Jin, et al. Multi-target regression via sparse integration and label-specific features. *Acta Electronica Sinica*, 2020, 48(5): 906 – 913.
(刘洪涛, 李航, 王进, 等. 基于标签特定特征的多目标回归稀疏集成方法. *电子学报*, 2020, 48(5): 906 – 913.)
- [15] WU S, LIU X, YU G, et al. Multi-target regression via stochastic configuration networks with modular stacked structure. *International Journal of Machine Learning and Cybernetics*, 2024, DOI: 10.1007/s13042-023-02075-2.
- [16] ROTHMAN A J, LEVINA E, ZHU J. Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics*, 2010, 19(4): 947 – 962.
- [17] ZHEN X T, YU M Y, ZHENG F, et al. Multitarget sparse latent regression. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, 29(5): 1575 – 1586.
- [18] ZHEN X T, YU M Y, HE X F, et al. Multi-target regression via robust low-rank learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(2): 497 – 504.
- [19] HE Dubo, SUN Shengxiang. Multi-target sparse regression with instance and target correlations. *Control and Decision*, 2024, 39(5): 1478 – 1486.
(何杜博, 孙胜祥. 基于实例与目标相关性的多目标稀疏回归算法. *控制与决策*, 2024, 39(5): 1478 – 1486.)
- [20] YUAN H L, ZHENG J J, LAI L L, et al. Sparse structural feature selection for multitarget regression. *Knowledge-based Systems*, 2018, 6: 200 – 209.
- [21] DINUZZO F, ONG C, GEHLER P, et al. Learning output kernels with block coordinate descent. *International Conference on Machine Learning*. Bellevue, WA, USA: International Machine Learning Society, 2011: 49 – 56.
- [22] ARASHLOO S, KITTLER J. Multi-target regression via non-linear output structure learning. *Neurocomputing*, 2022, 492: 572 – 580.
- [23] WANG Hongqiao, CAI Yanning, SUN Fuchun, et al. Adaptive sequence learning and applications for multi-scale kernel method. *Pattern Recognition and Artificial Intelligence*, 2011, 24(1): 72 – 81.
(汪洪桥, 蔡艳宁, 孙富春, 等. 多尺度核方法的自适应序列学习及应用. *模式识别与人工智能*, 2011, 24(1): 72 – 81.)
- [24] WANG Qianjin, YANG Chunyu, MA Xiaoping, et al. Underground airflow quantity modeling based on SCN. *Acta Automatica Sinica*, 2021, 47(8): 1963 – 1975.
(王前进, 杨春雨, 马小平, 等. 基于随机配置网络的井下供给风量建模. *自动化学报*, 2021, 47(8): 1963 – 1975.)
- [25] KIRSZBRAUN M D. Über die zusammenziehenden und lipschitzchen transformationen. *Fundamenta Mathematicae*, 1934, 22: 77 – 108.
- [26] LI Junbao, GAO Huijun. Data-dependent kernel function based kernel optimization algorithm. *Pattern Recognition and Artificial Intelligence*, 2010, 23(3): 300 – 306.
(李君宝, 高会军. 基于数据依赖核函数的核优化算法. *模式识别与人工智能*, 2010, 23(3): 300 – 306.)

作者简介:

代伟 教授, 博士生导师, 目前研究方向为人工智能驱动的复杂流程工业过程建模与运行优化控制, E-mail: weidai@cumt.edu.cn;

吴尚 硕士研究生, 目前研究方向为数据驱动建模和机器学习算法, E-mail: wushangcumt@163.com;

南静 博士研究生, 目前研究方向为随机神经网络、轻量化模型、工业物联网, E-mail: jingn@cumt.edu.cn;

刘鑫 副教授, 博士生导师, 目前研究方向为复杂数据驱动的系统辨识方法及其应用、工业数据分析与建模, E-mail: 15B904027@hit.edu.cn.