

奇异值分解下在线鲁棒正则化随机网络

于洋^{1†}, 邓瑞¹, 余刚^{2,3}, 庞新富⁴

(1. 沈阳航空航天大学 自动化学院, 辽宁 沈阳 110000; 2. 矿冶过程自动控制技术国家重点实验室, 北京 100160;

3. 矿冶自动控制技术北京市重点实验室, 北京 100160; 4. 沈阳工程学院 自动化学院, 辽宁 沈阳 110000)

摘要: 在线鲁棒随机神经网络(OR-RVFLN)具有较好的逼近性、较快的收敛速度、较高的鲁棒性能以及较小的存储空间。但是, OR-RVFLN 算法计算过程中会产生矩阵的不适定问题, 使得隐含层输出矩阵的精度较低。针对这个问题, 本文提出了奇异值分解下在线鲁棒正则化随机网络(SVD-OR-RRVFLN)。该算法在 OR-RVFLN 算法的基础上, 将正则化项引入到权值的估计中, 并且对隐含层输出矩阵进行奇异值分解; 同时采用核密度估计(KDE)法, 对整个SVD-OR-RRVFLN网络的权值矩阵进行更新, 并分析了所提算法的必要性和收敛性。最后, 将所提的方法应用于Benchmark数据集和磨矿粒度的指标预测中, 实验结果证实了该算法不仅可以有效地提高模型的预测精度和鲁棒性能, 而且具有更快的训练速度。

关键词: 随机神经网络; 正则化; 奇异值分解; 磨矿过程; 磨矿粒度

引用格式: 于洋, 邓瑞, 余刚, 等. 奇异值分解下在线鲁棒正则化随机网络. 控制理论与应用, 2024, 41(3): 407 – 415

DOI: 10.7641/CTA.2023.20857

Online robust regularized random networks under singular value decomposition

YU Yang^{1†}, DENG Rui¹, YU Gang^{2,3}, PANG Xin-fu⁴

(1. Automation College, Shenyang Aerospace University, Shenyang Liaoning 110000, China;

2. State Key Laboratory of Mining and Metallurgical Process Automatic Control Technology, Beijing 100160, China;

3. Beijing Key Laboratory of Mining and Metallurgy Automatic Control Technology, Beijing 110160, China;

4. Automation College, Shenyang Institute of Engineering, Shenyang Liaoning 110000, China)

Abstract: Online robust random vector functional link network (OR-RVFLN) has better approximation, faster convergence speed, higher robustness and smaller storage space. However, the OR-RVFLN algorithm can cause the ill-posed problem of the matrix in the calculation process, which makes the low precision of the hidden layer output matrix. To solve this problem, based on the singular value decomposition approach, this paper proposes the online robust regularized random vector functional link network (SVD-OR-RRVFLN). Firstly, the SVD-OR-RRVFLN introduces the regularization term into the OR-RVFLN algorithm, and the singular value decomposition approach is used for the hidden layer output matrix. Further, the kernel density estimation (KDE) method is used to update the matrix weight. Secondly, the necessity and convergence of the proposed algorithm are analyzed. Finally, the proposed method is applied to Benchmark data set and the index prediction of grinding particle size. The experimental results show that the proposed algorithm can not only effectively improve the prediction accuracy and robustness of the model, but also have faster training speed.

Key words: random vector functional link network; regularization; singular value decomposition; grinding process; grinding particle size

Citation: YU Yang, DENG Rui, YU Gang, et al. Online robust regularized random networks under singular value decomposition. *Control Theory & Applications*, 2024, 41(3): 407 – 415

收稿日期: 2022-09-29; 录用日期: 2023-10-23.

†通信作者. E-mail: 20180029@sau.edu.cn; Tel.: +86 24-89724448.

本文责任编辑: 周平.

矿冶过程自动控制技术国家重点实验室、矿冶过程自动控制技术北京市重点实验室项目(BGRIMM-KZSKL-2021-03), 国家自然科学基金项目(61773269), 辽宁省自然科学基金项目(2021-BS-189)资助.

Supported by the Open Foundation of State Key Laboratory of Process Automation in Mining and Metallurgy, Beijing Key Laboratory of Process Automation in Mining and Metallurgy (BGRIMM-KZSKL-2021-03), the National Natural Science Foundation of China (61773269) and the Natural Science Foundation of Liaoning Province of China (2021-BS-189).

1 引言

在复杂的工业过程中,为了实现运行指标的闭环优化控制,通常需要采用建模法^[1]对运行指标进行准确测量.建模方法的研究主要分为两个方面:机理分析建模方法和数据驱动建模方法^[2].传统的机理分析建模方法在城市电网电流耦合、医学图像诊断、极端天气的预测以及飞行控制^[3-6]等领域有重要的研究.但是,机理分析的建模方法因难以获取系统的基础知识和关键参数,会导致一定的偏差^[7].数据驱动的建模方法利用输入输出数据来建立对运行指标的估算模型,因此被广泛地用于解决高频电能转换、多机场系统运行、背景图像更新以及核反应堆中热流体现象等实际问题^[8-12].

目前,神经网络^[13]是最重要的数据驱动建模方法之一.神经网络可以通过学习大量的训练数据来拟合复杂非线性函数^[14],它具有处理复杂系统的自学习能力,可帮助解决增值评价中有关模型构建、数据处理及大规模应用的问题^[15].按照隐含层的层数,可以将其划分为多层前向型和单层前向型神经网络.多层前向型神经网络,拥有优良的逼近和泛化能力,被广泛研究和应用;单隐含层前向神经网络(single-hidden layer feed-forward neural networks, SLFN)^[16]亦具有较好的逼近性,反向传播神经网络(back propagation, BP)算法是主要的SLFN权值调整算法,已被用于补偿光纤陀螺温度、监测土地荒漠化趋势、人体动作识别以及网络安全评估^[17-20]等.然而,BP算法在学习过程中效率较低,限制了SLFN的发展与应用^[21].

随机权神经网络^[22](random vector functional link network, RVFLN)解决了BP算法学习效率低的问题.RVFLN的输入和输出是由一条直线链直接相连的,且输出层同时连接输入层和增强节点并汇集了SLFN网络的优点.由于RVFLN简单的训练过程以及万能逼近特性^[23],在实际工业过程中得到广泛的应用.

近年来,为了提高RVFLN的性能,学者们已提出许多改进策略.在权值变换方面,文献[24]将RVFLN的随机权重控制在 $[0, 1]$ 之内,建立了小范围RVFLN,提高了算法的泛化性能.文献[25]将合适的带权重因子的特征向量作为输入,提出了将经验小波变换与希尔伯特变换结合的加权RVFLN,可以准确预测故障位置.在此基础上,针对深度结构的神经网络大而时变的特性,文献[26]将经验小波变换和鲁棒最小方差RVFLN法相结合,以提高预测精度.文献[27]提出了鲁棒多核RVFLN,用于计算高度非平稳的组合电能质量事件,以获得突出的识别精度和稳健的抗噪性能.另外,为减少隐藏节点的数量,文献[28]利用增量学习的方法,使部分隐藏节点不被随机选择,提出了选择性集成深部双向RVFLN.同时,由于工业生产往往具有慢时变特性,所建立的数据模型的参数也应随之改

变.因此,文献[29]提出了自适应半监督的RVFLN,对模型参数进行了优化,并且其学习速度和精度均优于传统方法.文献[30]提出了在线鲁棒随机权神经网络(online robust RVFLN, OR-RVFLN)算法,该算法可以有效地减少内存,加快运算效率,并且具有良好的鲁棒性能.

目前,这些算法仍属于批处理最小二乘算法,输出权值矩阵的估计可能存在不适定问题,会导致计算误差偏大,从而影响预测精度;并且由于较小奇异值的存在,隐含层输出矩阵接近于奇异矩阵,从而影响模型的输出精度.因此,针对该问题,本文提出了基于奇异值分解的在线鲁棒正则化随机权神经网络(singular value decomposition online robust regularization RVFLN, SVD-OR-RRVFLN).

首先,该算法在OR-RVFLN的权值估计中引入带有惩罚项的损失函数,得到在线鲁棒正则化RVFLN(OR-RRVFLN),并且,对每一个隐含层输出矩阵进行奇异值分解,从而得到所提的SVD-OR-RRVFLN算法.其次,本文对所提算法进行必要性和收敛性的证明.最后,基于Benchmark数据集和实际的磨矿粒度数据对所提算法进行实验分析,证实该算法在提高了模型的预测精度的同时保持了更高的鲁棒性能.

2 相关算法

本节主要介绍在线鲁棒随机权神经网络和奇异值分解算法,以此来引出本文所提的基于奇异值分解的在线鲁棒正则化随机权神经网络.

2.1 在线鲁棒随机权神经网络

当 N 组数据样本 (x_i, y_i) 采样时,存在 L 个隐含层节点的SLFN的成本函数可以用误差平方和公式表示,即

$$J = \sum_{i=1}^N \|f(x_i) - y_i\|^2 = \sum_{i=1}^N \left\| \sum_{j=1}^L \omega_j \varphi_j(v_j, b_j, x_i) - y_i \right\|^2, \quad (1)$$

其中: v_j 是输入权值向量; b_j 是隐含层第 j 个节点阈值; ω_j 是第 j 个隐含层节点和输出层节点相连的权重.不同于SLFN需要调节所有参数,RVFLN可随机选取 v_j 和 b_j ,只需求权重 ω_j ,就可保证网络的逼近能力.因此,式(1)可转化为线性二次优化问题,即

$$\arg \min_{\omega_1, \dots, \omega_L} \sum_{i=1}^N \left\| \sum_{j=1}^L \omega_j \varphi_j(v_j, b_j, x_i) - y_i \right\|^2, \quad (2)$$

式(2)的矩阵形式表示为

$$\min_W \|HW - Y\|, \quad (3)$$

其中: $H = [h(x_1) \ \dots \ h(x_N)]^T \in \mathbb{R}^{N \times L}$ 为隐含层输出矩阵; $h(x_N) = [\varphi(v_1, b_1, x_N) \ \dots \ \varphi(v_L, b_L, x_N)]$ 为隐含层的特征映射; N 为样本数; $W = [\omega_1 \ \omega_2 \ \dots$

$\omega_L]^T \in \mathbb{R}^L$ 为输出权值向量; $Y = [y_1 \ y_2 \ \cdots \ y_N]^T \in \mathbb{R}^N$ 为实际的被测向量. 利用方程 $HW = Y$, 可以得到输出层的权重估计值 \hat{W} , 即

$$\hat{W} = H^+ Y, \quad (4)$$

式中 $H^+ = (H^T H)^{-1} H^T$ 为 H 的摩尔广义逆^[31], 将其代入到式(4)中, 就得到了最小二乘解, 即

$$\hat{W} = (H^T H)^{-1} H^T Y. \quad (5)$$

在RVFLN的权值估计中引入加权最小二乘^[32], 可提高RVFLN的鲁棒性, 即

$$J = (HW - Y)^T P (HW - Y), \quad (6)$$

权值 W 的估计如下:

$$\hat{W} = (H^T P H)^{-1} H^T P Y, \quad (7)$$

其中 $P = \text{diag}\{p_1, p_2, \dots, p_N\}$ 为惩罚权重矩阵, 可根据样本数据的可靠性进行调整, 采用核密度估计(kernel density estimation, KDE)方法^[33]进行估计.

为了减小存储空间和提高运算速度, 在线鲁棒RVFLN^[23]受到关注, 具体算法如下: 分别定义初始输出权值 \hat{W}_0 和设置初始权重矩阵为 P_0 , 即

$$\hat{W}_0 = (H_0^T P_0 H_0)^{-1} H_0^T P_0 Y_0, \quad (8)$$

令 $K_0 = (H_0^T P_0 H_0)^{-1}$.

当新的数据 $\{(x_i, y_i)\}_{i=N+1}^{N+N_1}$ 产生时, N_1 是新的样本数量, 式(3)可用如下方程表示:

$$\min_W \left\| \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} - \begin{bmatrix} Y_0 \\ Y_1 \end{bmatrix} \right\|^2, \quad (9)$$

根据式(7)和式(9)的解, 可以写为

$$\left(\begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} P_0 & 0 \\ 0 & P_1 \end{bmatrix} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \right)^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} P_0 & 0 \\ 0 & P_1 \end{bmatrix} \begin{bmatrix} Y_0 \\ Y_1 \end{bmatrix} = K_1 (H_0^T P_0 Y_0 + H_1^T P_1 Y_1), \quad (10)$$

其中: $H_1 = [h(x_{N+1}) \ \cdots \ h(x_{N+N_1})]^T \in \mathbb{R}^{N_1 \times L}$, $h(x_{N+N_1}) = [\varphi(v_1, b_1, x_{N+N_1}) \ \cdots \ \varphi(v_L, b_L, x_{N+N_1})]$, $Y_1 = [y_{N+1} \ y_{N+2} \ \cdots \ y_{N+N_1}]^T$, $P_1 = \text{diag}\{p_{N+1}, p_{N+2}, \dots, p_{N+N_1}\}$, 并且

$$K_1 = (H_0^T P_0 H_0 + H_1^T P_1 H_1)^{-1} = (K_0^{-1} + H_1^T P_1 H_1)^{-1}. \quad (11)$$

根据反演公式 $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$, 可得

$$K_1 = K_0 - K_0 H_1^T (P_1^{-1} + H_1^T K_0 H_1)^{-1} H_1 K_0 = (I - Q_1 H_1) K_0, \quad (12)$$

其中

$$Q_1 = K_0 H_1^T (P_1^{-1} + H_1 K_0 H_1^T)^{-1}. \quad (13)$$

把式(12)代入式(10)中, 可得

$$\begin{aligned} \hat{W}_1 &= (I - Q_1 H_1) K_0 (H_0^T P_0 Y_0 + H_1^T P_1 Y_1) = \\ &= \hat{W}_0 - Q_1 H_1 \hat{W}_0 + Q_1 P_1^{-1} P_1 Y_1 = \\ &= \hat{W}_0 + Q_1 (Y_1 - H_1 \hat{W}_0). \end{aligned} \quad (14)$$

综上所述, 当第 k 个数据块产生时, 网络参数的权值更新算法如下:

$$\begin{cases} \hat{W}_k = \hat{W}_{k-1} + Q_k (Y_k - H_k \hat{W}_{k-1}), \\ K_k = (I - Q_k H_k) K_{k-1}, \\ Q_k = K_{k-1} H_k^T (P_k^{-1} + H_k K_k H_k^T)^{-1}. \end{cases} \quad (15)$$

2.2 奇异值分解算法

随着大数据技术的飞速发展, 矩阵分解特别是矩阵的奇异值分解(singular value decomposition, SVD)在数据检索、图像压缩、人脸识别和故障诊断等领域有着广泛应用^[34-35]. 如今已经衍生出许多SVD的变体, 包括随机SVD^[36]、顺序截断高阶SVD^[37]、改进SVD^[38]、Tikhonov正则化改进随机截断SVD^[39]、局部均值分解SVD^[40].

假设存在一矩阵 $A \in \mathbb{R}^{m \times n}$, 使得

$$A = U \Sigma D^T = \sum_{i=1}^n \sigma_i u_i d_i^T, \quad (16)$$

其中: $U = [u_1 \ u_2 \ \cdots \ u_m] \in \mathbb{R}^{m \times m}$ 是 $m \times m$ 阶酉矩阵; $D = [d_1 \ d_2 \ \cdots \ d_n] \in \mathbb{R}^{n \times n}$ 是 $n \times n$ 阶对角矩阵; $\Sigma = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\} \in \mathbb{R}^{n \times n}$ 是 $n \times n$ 阶酉矩阵; u_i 是第 i 个正交向量; d_i 是第 i 个奇异值; m 和 n 分别是矩阵 A 的行和列.

通过奇异值分解, 可以去除矩阵中存在的极小的奇异值, 从而提高矩阵的输出精度.

3 基于奇异值分解的在线鲁棒正则化随机神经网络

通过第2节对相关算法的说明, 本节主要介绍了所提算法的具体形式, 描述了算法的具体步骤, 并且对算法的必要性和收敛性进行了分析.

3.1 SVD-OR-RRVFLN算法具体形式

隐含层输出矩阵 H 求摩尔广义逆 H^+ 的过程可能会产生矩阵的不适定问题, 为此, 本文将 l_2 正则化引入到式(7)中, 得到

$$\hat{W}' = (H^T P H + \mu I)^{-1} H^T P Y, \quad (17)$$

其中 μ 为正则化参数.

根据式(8)-(14)(17), 得到 OR-RRVFLN 递推公式, 即

$$\hat{W}'_k = \hat{W}'_{k-1} + Q'_k (Y_k - H_k \hat{W}'_{k-1}), \quad (18a)$$

$$K'_k = (I - Q'_k H_k) K'_{k-1}, \quad (18b)$$

$$Q'_k = K'_{k-1} H_k^T (P_k^{-1} + H_k K'_k H_k^T)^{-1}, \quad (18c)$$

$$K'_0 = (H_0^T P_0 H_0 + \mu I)^{-1}. \quad (18d)$$

奇异值分解法可以提高隐含层输出矩阵的精度. 因此, 本文对式(18a)–(18c)中每一个隐含层输出矩阵 H 进行奇异值分解, 其表达式为

$$H_k = \sum_{i=1}^L \sigma_{i,k} u_{i,k} d_{i,k}^T. \quad (19)$$

式(18a)–(18d)–(19)构成了本文所提的SVD-OR-RRVFLN算法. 在式(18d)中引入了正则化项, 解决了由逆矩阵引起的不适定性; 同时, 根据式(19), 奇异值分解算法去除了隐含层矩阵分解后产生的极小的奇异值, 进一步提高了模型输出的精度.

3.2 SVD-OR-RRVFLN算法描述

SVD-OR-RRVFLN算法包括离线和在线两个学习过程, 具体学习过程如下:

步骤1 离线学习阶段.

1) 选取隐含层节点数 L , 随机分配输入权重 v_j 和 b_j ;

2) 从数据中选择 M 组训练样本 ($M/5 > L$), 采用5折交叉验证法, 分别进行5次实验, 每次选择其中一组作为测试样本, 其余为训练样本;

3) 在每次实验中, 根据式(17)估计权重 \hat{W}' ;

4) 计算5次实验估计误差的均值, 判断是否具有最小的误差均值, 满足则给出 \hat{W}'_0 和 K'_0 , 然后, 进入在线学习阶段; 不满足则继续选取其他 L , 重新进行验证试验.

步骤2 在线学习阶段.

1) 选取最大迭代次数 K_{\max} , 令 $k = 1$;

2) 当新的数据产生时, 根据式(19)求分解后的隐含层输出矩阵 H_k ;

3) 根据式(18a)–(18c)更新矩阵 \hat{W}'_k , K'_k 和 Q'_k ;

4) 如果 $k = K_{\max}$, 结束; 否则, 令 $k \leftarrow k + 1$, 返回2).

3.3 SVD-OR-RRVFLN算法必要性的证明

定理1 设 ΔY 和 ΔW 分别代表 Y 和 W 的偏差, 可满足如下不等式:

$$\frac{\Delta W}{W} \leq E(H) \frac{\Delta Y}{Y}, \quad (20)$$

其中 $E(H) = \|H\| \|H^{-1}\|$ 为条件数.

证 根据公式

$$Y = HW, \quad (21)$$

令

$$H(W + \Delta W) = Y + \Delta Y, \quad (22)$$

所以, 由式(22)可得

$$H\Delta W = \Delta Y, \quad (23)$$

即

$$\Delta W = H^{-1} \Delta Y. \quad (24)$$

根据范数的相容性^[41], 式(21)–(24)的表达如下:

$$\|H\| \|W\| \geq \|Y\|, \quad (25)$$

$$\|\Delta W\| \leq \|H^{-1}\| \|\Delta Y\|, \quad (26)$$

用式(26)除以式(25), 可以得到如下不等式:

$$\frac{\Delta W}{\|H\| \|W\|} \leq \frac{\|H^{-1}\| \|\Delta Y\|}{\|Y\|}, \quad (27)$$

所以

$$\frac{\Delta W}{W} \leq \|H\| \|H^{-1}\| \frac{\Delta Y}{Y}. \quad (28)$$

证毕.

假设隐含层输出矩阵 H 特征值的最大值和最小值分别为 σ_{\max} 和 σ_{\min} , 根据文献[42], 有如下关系式:

$$E(H) = \|H\| \|H^{-1}\| = \frac{\sigma_{\max}}{\sigma_{\min}}, \quad (29)$$

根据式(29)可知, 当奇异值接近于0时, 条件数变得非常大, 矩阵接近于奇异矩阵, 影响计算精度, 导致最小二乘方法不容易对普通的函数进行拟合, 并且限制了RVFLN神经网络的学习能力. 因此, 本文将正则化参数和奇异值分解方法引入到OR-RVFLN中.

3.4 SVD-OR-RRVFLN算法收敛性的证明

定理2 设 $\varepsilon = Y - HW$ 是零均值的白噪声, 则式(18a)给出的 \hat{W}'_k 是一致收敛的, 即

$$\lim_{k \rightarrow \infty} \hat{W}'_k = \alpha, \quad (30)$$

其中 α 是模型参数真值.

证 令 $\hat{W}''_k = \alpha - \hat{W}'_k$, 设

$$\lim_{k \rightarrow \infty} \hat{W}''_k = \hat{W}''_k, \quad (31)$$

通过 $\varepsilon = Y - HW$, 可得

$$Y_k = H_k \alpha + \varepsilon_k. \quad (32)$$

根据式(18a)–(32), 可得

$$\begin{aligned} \hat{W}''_k &= (I - Q_k H_k) \hat{W}''_{k-1} - Q_k \varepsilon_k = \\ &K_k'^{-1} K_k' \hat{W}''_{k-1} - Q_k \varepsilon_k, \end{aligned} \quad (33)$$

令 $A_k = K_k'^{-1} K_k'$, 则通过式(12)可得

$$A_k = K_k'^{-1} K_k' = (K_{k-1}'^{-1} + H_k^T P_k H_k)^{-1}, \quad (34)$$

则 \hat{W}''_k 可由下式表示:

$$\hat{W}''_k = A_k \hat{W}''_{k-1} - Q_k \varepsilon_k, \quad (35)$$

设矩阵 A_k 的特征值是 η , 则下式成立:

$$A_k x = \eta x, \quad (36)$$

式中 x 是非零特征向量. 把式(33)代入到式(35)中, 可得

$$(1 - \eta) K_k'^{-1} x = \eta H_k^T P_k H_k x, \quad (37)$$

根据式(19), 进一步有

$$(1 - \eta)x^T K_k'^{-1}x = \eta x^T \sum_{i=1}^t p_i \sigma_{i,k}^2 x. \quad (38)$$

因为 $K_k'^{-1}$ 是正定矩阵, 且 $\sum_{i=1}^t p_i \sigma_{i,k}^2 > 0$. 因此, 对于所有的非零向量 x , 式(37)中的 $(1 - \eta)$ 和 η 必须同号, 即

$$\frac{1 - \eta}{\eta} = \frac{1}{\eta} - 1 > 0. \quad (39)$$

可见, 矩阵 A_k 的特征值满足 $0 < \eta < 1$, 所以式(35)在平衡点 $\hat{W}'' = 0$ 处是稳定的, 然后根据式(31), 可得

$$\lim_{k \rightarrow \infty} \hat{W}_k'' = 0. \quad (40)$$

由于 $\hat{W}_k'' = \alpha - \hat{W}_k'$, 所以可得式(30). 证毕.

4 性能评估

为了验证算法的有效性, 本文将所提的SVD-OR-RRVFLN, OR-RVFLN^[30] 和 OR-RRVFLN^[43] 算法进行比较. OR-RVFLN 算法^[43]中逆矩阵的求解采用 MATLAB提供的pinv(·)函数来实现, 神经网络隐含层的节点数采用交叉验证法进行确定.

本实验先把输入和输出的数据归一化到区间[0, 1]上; 再将不同的算法分别独立运行50次; 最后, 通过选取测试数据的均方根误差 (root mean square error, RMSE)、标准差 (standard deviation, Dev.)以及平均相对误差3种性能指标, 验证算法的预测精度和鲁棒性能.

4.1 Benchmark数据集

本实验采用加利福尼亚大学欧文分校所提供的 Benchmark数据集验证所提算法性能¹, Benchmark数据集属性如表1所示.

表 1 Benchmark数据集属性^[30]

数据集	变量格式	输出变量	样本数	训练数据个数	测试数据个数
SinC	2	1	300	200	100
Airfoil Self-noise	6	1	1503	703	800
Yacht	7	1	308	200	108
Concrete	9	1	1030	800	230

4.2 比较实验

对于每一个数据集, 首先在固定的范围内, 即0.01, 0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 选取正则化参数, 所提算法在数据集上对不同正则化参数的估计结果如表2所示. 通过表2发现, 当正则化参数分别为3, 5, 8和0.5时, 所对应的均方根误差和标准差最小.

选取的最优正则化参数用于所提算法和 OR-RRVFLN算法^[43], 并且与OR-RVFLN算法^[30]进行比较,

其结果如表3所示, 对于每一个数据集, 所提算法可获得最小的均方根误差和标准差, OR-RRVFLN^[43]次之, OR-RVFLN^[30]最大, 实验结果证实了所提算法具有较高的鲁棒性能. 同时, 为了便于观察所提算法在预测性能上的优势, 本文以SinC数据集为例加以说明. 从图1可以看出, 相比于另外两种算法, SVD-OR-RRVFLN算法的输出值更加接近真实值; 根据相对误差图(图1)和平均相对误差表(表4)可知, 与另外两种算法相比, 本文所提算法的相对误差以及平均相对误差都是最小的. 因此, 实验验证了所提算法具有较高的预测性能.

表 2 不同正则化参数的估计结果

Table 2 Estimation results of different regularization parameters

μ	SinC	Airfoil Self-noise	Yacht	Concrete
	RMSE Dev.	RMSE Dev.	RMSE Dev.	RMSE Dev.
0.01	0.0088	0.4853	0.1593	11.5478
	0.0008	0.0320	0.0220	1.6132
0.1	0.0086	0.4834	0.1583	11.5443
	0.0007	0.0320	0.0216	1.6128
0.5	0.0085	0.4928	0.1594	11.5425
	0.0007	0.0319	0.0221	1.6126
1	0.0073	0.4943	0.1587	11.5454
	0.0006	0.0323	0.0217	1.6129
2	0.0117	0.4898	0.1655	11.5487
	0.0009	0.0339	0.0229	1.6133
3	0.0057	0.4885	0.1645	11.5504
	0.0005	0.0332	0.0228	1.6135
4	0.0079	0.4876	0.1632	11.5513
	0.0007	0.0327	0.0226	1.6136
5	0.0128	0.4823	0.1603	11.5535
	0.0010	0.0318	0.0223	1.6139
6	0.0096	0.4854	0.1596	11.5547
	0.0008	0.0320	0.0221	1.6142
7	0.0132	0.4886	0.1587	11.5501
	0.0011	0.0334	0.0217	1.6134
8	0.0063	0.4903	0.1579	11.5503
	0.0006	0.0341	0.0215	1.6135
9	0.0062	0.4924	0.1593	11.5475
	0.0006	0.0348	0.0220	1.6132
10	0.0087	0.4883	0.1587	11.5489
	0.0007	0.0329	0.0217	1.6133

上述实验说明了, 在Benchmark数据集上, 所提的SVD-OR-RRVFLN算法不仅具有较高的鲁棒性, 在预

¹<http://archive.ics.uci.edu/ml/machine-learning-databases/>

测精度方面具有明显的优势.

表 3 3种算法在Benchmark数据集上的比较

Table 3 Comparison of three algorithms on Benchmark datasets

数据集	算法	RMSE	Dev.
SinC	OR-RVFLN ^[30]	0.0068	0.0007
	OR-RRVFLN ^[43]	0.0059	0.0006
	SVD-OR-RRVFLN	0.0057	0.0005
Airfoil Self-noise	OR-RVFLN ^[30]	0.5069	0.0364
	OR-RRVFLN ^[43]	0.4902	0.0341
	SVD-OR-RRVFLN	0.4823	0.0318
Yacht	OR-RVFLN ^[30]	0.2036	0.0292
	OR-RRVFLN ^[43]	0.1826	0.0248
	SVD-OR-RRVFLN	0.1579	0.0215
Concrete	OR-RVFLN ^[30]	11.8402	1.8215
	OR-RRVFLN ^[43]	11.8128	1.6213
	SVD-OR-RRVFLN	11.5425	1.6126

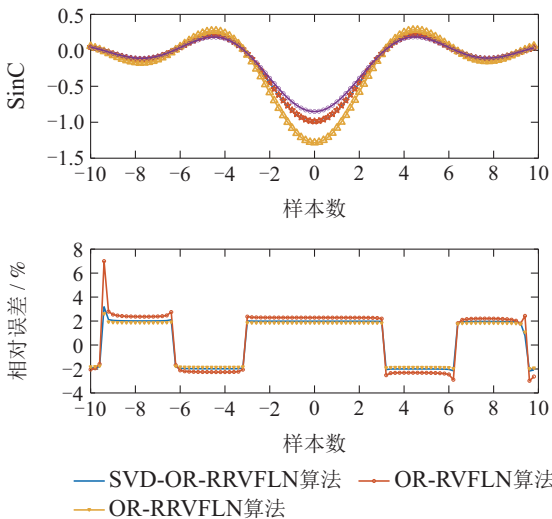


图 1 3种算法在SinC上的估计结果和相对误差

Fig. 1 Estimation results and relative error of three algorithms on SinC

表 4 3种算法在SinC上的平均相对误差比较

Table 4 Comparison of average relative error of three algorithms on SinC

算法	训练相对误差/%	测试相对误差/%
OR-RVFLN ^[30]	13.66	14.58
OR-RRVFLN ^[43]	4.89	5.17
SVD-OR-RRVFLN	2.25	4.81

4.3 SVD-OR-RRVFLN算法在磨矿过程指标预测中的应用

赤铁矿一段磨矿过程的工艺流程如图2所示. 该过程中原矿先经过传送带送入球磨机进行研磨, 再通过

钢球不断破碎, 生产出一定浓度的矿浆. 同时, 向磨机中加入一定量的水, 以将矿浆浓度调节在限制范围内. 粉碎后得到的浆液将被连续地进入到螺旋分级机内进行分级, 并且把一定量浓度的研磨水直接添加到螺旋分级机腔中, 以便随时调节其浓度; 把检验不合格的矿浆再送入球磨机中继续研磨, 而从分级机溢流出的标准矿则送入选别工艺.

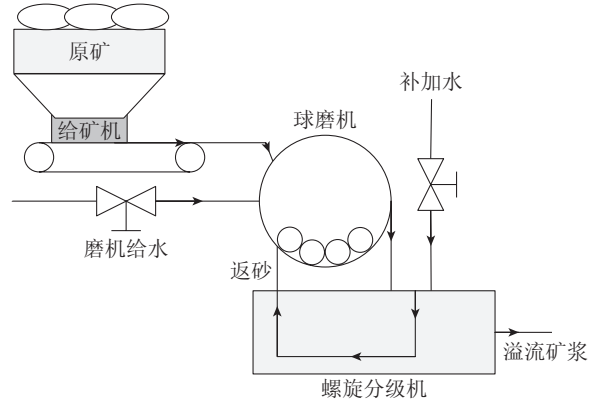


图 2 赤铁矿一段磨矿过程的工艺流程

Fig. 2 Process flow of primary grinding process of hematite

由于赤铁矿矿石性质不稳定且矿石颗粒具有磁团聚现象, 实验室化验方法无法满足磨矿粒度闭环优化控制的要求. 传统的模型驱动的建模方法会因为一些条件的忽略而引起较大的误差^[30]. 数据驱动建模方法是一种黑箱建模方法, 可有效地提高模型的精度和鲁棒性. 本文采用 70 组测试数据, 将所提的 SVD-OR-RRVFLN 算法应用于磨矿粒度软测量模型的建模中, 以此来验证所提算法性能.

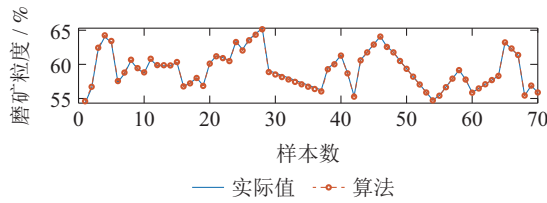
表5为所提算法对不同正则化参数的误差结果. 在表5中选取使RMSE和Dev.值最小的正则化参数; 图3是3种算法下磨矿粒度估计的预测结果. 为进一步分析3种算法的性能, 分别计算不同算法的结果以及相对误差, 并在表6和图4中给出. 从图4可以看出, 除了在样本数30~40区间内, 所提算法总体上具有最小的相对误差; 同时根据表6, 进一步验证了所提算法具有最小的均方根误差, 标准差, 以及平均相对误差, 说明了SVD-OR-RRVFLN算法在磨矿过程数据中整体上具有良好的预测精度和鲁棒性能.

为了验证所提算法的时效性, 本文对比了所提算法与另外两种算法接收新样本时的在线训练时间. 比较结果如图5所示, 所提方法的平均训练时间为0.008 s; OR-RVFLN^[30]算法的平均训练时间为0.02 s, OR-RRVFLN^[43]算法的平均训练时间为0.015 s. 伴随着样本的累积, 所提算法的训练时间相对于另外两种算法更加稳定. 通过以上的实验对比发现, 所提算法不仅有较高的鲁棒性能和预测性能, 而且有更少的时间消耗. 因此, 所提的SVD-OR-RRVFLN算法提高了训练速度.

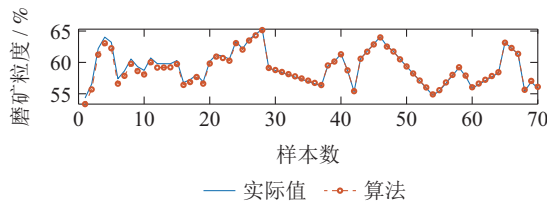
表5 不同正则化参数的误差结果

Table 5 The error results of the different regularization parameters

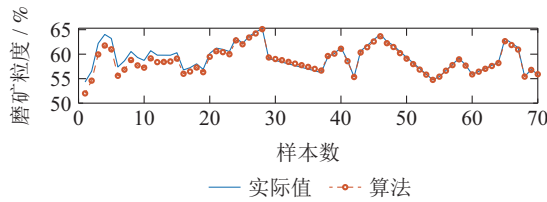
μ	RMSE	Dev.	μ	RMSE	Dev.
0.01	0.2423	0.0612	5	0.2409	0.0592
0.1	0.2413	0.0605	6	0.2432	0.0621
0.5	0.2392	0.0587	7	0.2451	0.0634
1	0.2388	0.0572	8	0.2384	0.0571
2	0.2375	0.0541	9	0.2396	0.0582
3	0.2382	0.0556	10	0.2401	0.0593
4	0.2398	0.0572	—	—	—



(a) SVD-OR-RRVFLN算法



(b) OR-RRVFLN算法



(c) OR-RVFLN算法

图3 3种算法在磨矿粒度上的估计预测结果

Fig. 3 Estimation prediction results of three algorithms on grinding particle size

表6 3种算法在磨矿粒度上的估计误差结果

Table 6 Estimation error results of three algorithms on the grinding particle size

算法	RMSE	Dev.	平均相对误差/%
OR-RVFLN ^[30]	0.2786	0.0748	1.32
OR-RRVFLN ^[43]	0.2542	0.0634	0.45
SVD-OR-RRVFLN	0.2375	0.0541	0.15

5 结论

本文针对逆矩阵的不适定问题提出了奇异值分解的在线鲁棒正则化随机神经网络, 即SVD-OR-RRVFLN方法. 将正则项添加到输出权值的估计矩阵中, 并采用奇异值分解法对整个网络的隐含层输出矩阵进行更新, 从而提高了模型的预测精度. 同时采用核

密度估计法, 对整个SVD-OR-RRVFLN网络的权值矩阵进行更新, 以提高算法的鲁棒性能. 基于Benchmark数据集的仿真实验表明: 与其他方法相比, 本文所提的方法能有效提高预测精度和鲁棒性能. 同时, 这个方法应用于磨矿过程中, 进一步说明所提算法不仅可以提高模型的预测精度和鲁棒性能, 而且加快了训练速度.

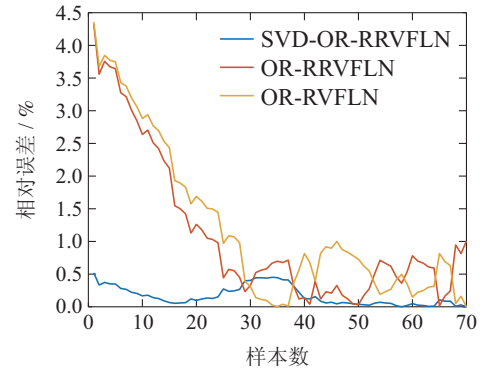


图4 3种算法在磨矿粒度上的相对误差

Fig. 4 Relative error of three algorithms on grinding particle size

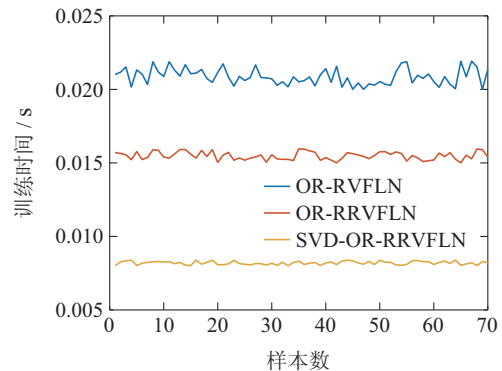


图5 在线训练时间

Fig. 5 Online training time

参考文献:

- [1] WANG Qianjin, DAI Wei, LU Qun, et al. A sparse learning method for randomly assigned network soft sensor model. *Control and Decision*, 2022, 37(12): 1 – 13.
(王前进, 代伟, 陆群, 等. 一种随机配置网络软测量模型的稀疏学习方法. *控制与决策*, 2022, 37(12): 1 – 13.)
- [2] CHEN Long, LIU Quanli, WANG Linqing, et al. A survey of production process index prediction methods in process industry based on data. *Acta Automatica Sinica*, 2017, 43 (6): 944 – 954.
(陈龙, 刘全利, 王霖青, 等. 基于数据的流程工业生产过程指标预测方法综述. *自动化学报*, 2017, 43(6): 944 – 954.)
- [3] JIANG Youhua, QIAN Yan, CAO Yilong. Analysis and research on harmonic current coupling mechanism of urban power grid. *Control Theory & Applications*, 2017, 34 (2): 233 – 242.
(江友华, 钱彦, 曹以龙. 城市电网谐波电流耦合机理分析与研究. *控制理论与应用*, 2017, 34(2): 233 – 242.)
- [4] OKUWOBI I, DING Z, WAN J, et al. Artificial intelligence model driven by transfer learning for image-based medical diagnosis. *Journal of Intelligent & Fuzzy Systems*, 2022, 43(4): 4601 – 4612.

- [5] GAO S, ZHU S, YU H. Dynamical downscaling of temperature extremes over China using the WRF model driven by different lateral boundary conditions. *Atmospheric Research*, 2022, 278: 106348.
- [6] CHEN Xuanwen, MA Chao, MA Qian, et al. Research on flight control software factory based on model driven component library. *Measurement and Control Technology*, 2022, 42(2): 1 – 8.
(陈宣文, 马超, 马倩, 等. 基于模型驱动构件库的飞行控制软件工厂研究. 测控技术, 2022, 42(2): 1 – 8.)
- [7] XIE S, XIE Y F, HUANG T, et al. Generalized predictive control for industrial processes based on neuron adaptive splitting and merging RBF neural network. *IEEE Transactions on Industrial Electronics*, 2019, 66(2): 1192 – 1202.
- [8] WANG B, LI Z, DAI Z, et al. Data-driven mode identification and unsupervised fault detection for nonlinear multimode processes. *IEEE Transactions on Industrial Informatics*, 2020, 16(6): 3651 – 3661.
- [9] WANG Runnan, XIE Fan, ZHANG Bo. Data driven modeling of peak current controlled boost converter based on neural network method. *Journal of Electrical Engineering*, 2022, 17(2): 142 – 150.
(王闰南, 谢帆, 张波. 基于神经网络方法的峰值电流控制Boost变换器数据驱动建模. 电气工程学报, 2022, 17(2): 142 – 150.)
- [10] CHENG Binghe, ZHAO Zheng. Departure flow analysis of multi airport system based on data driven modeling. *Aerocomputing Technology*, 2022, 52(3): 67 – 69, 73.
(程炳贺, 赵征. 基于数据驱动建模的多机场系统离场流分析. 航空计算技术, 2022, 52(3): 67 – 69, 73.)
- [11] SUN Guoqing, HOU Zhongsheng. Data driven background image updating algorithm in camera dithering scene. *Control Theory & Applications*, 2022, 39(5): 840 – 933.
(孙国庆, 侯忠生. 相机抖动场景中数据驱动的背景图像更新算法. 控制理论与应用, 2022, 39(5): 840 – 933.)
- [12] LIU Y, HU R, KRAUS A, et al. Data-driven modeling of coarse mesh turbulence for reactor transient analysis using convolutional recurrent neural networks. *Nuclear Engineering and Design*, 2022, 390: 111716.
- [13] WANG Dianhui, CHAI Tianyuo. A robust adaptive control algorithm based on neural network model. *Acta Automatica Sinica*, 1996, 22(4): 447 – 451.
(王殿辉, 柴天佑. 一个基于神经网络模型的鲁棒自适应控制算法. 自动化学报, 1996, 22(4): 447 – 451.)
- [14] WANG Jingyi, WANG Zhiguo, CHEN Yumin, et al. The application of depth artificial neural network in seismic inversion. *Progress in Geophysics*, 2022, 38(1): 1 – 26.
(王竞仪, 王治国, 陈宇民, 等. 深度人工神经网络在地震反演中的应用进展. 地球物理学进展, 2022, 38(1): 1 – 26.)
- [15] LI Junfei, TAN Dingliang, LI Gefei. The application of artificial neural network in value-added evaluation is analyzed. *China Examination*, 2022, 7: 77 – 84.
(李俊飞, 谭顶良, 李格非. 试析人工神经网络在增值评价中的应用. 中国考试, 2022, 7: 77 – 84.)
- [16] HORNIK K, STINCHCOMBE M H, WHITE H. Multilayer feedforward networks are universal approximates. *Neural Network*, 1989, 2(5): 359 – 366.
- [17] QIU Haitao, XU Mengtong, LIU Wei, et al. Research on temperature compensation method of FOG based on ACO-BP neural network. *Elector Optics and Control*, 2022, 30(7): 1 – 7.
(仇海涛, 徐梦桐, 刘伟, 等. 基于ACO-BP神经网络的光纤陀螺温度补偿方法研究. 电光与控制, 2022, 30(7): 1 – 7.)
- [18] LU G, XU D, MENG Y. Dynamic evolution analysis of desertification images based on BP neural network. *Computational Intelligence and Neuroscience*, 2022, DOI: 10.1155/2022/5645535.
- [19] ZHAO Yan, GUO Ming, SUN Jianqiang, et al. Inertial sensor position correction algorithm in human motion recognition. *Control Theory & Applications*, 2021, 38 (11): 1883 – 1890.
(赵琰, 郭明, 孙建强, 等. 人体动作识别中的惯性传感器位置校正算法. 控制理论与应用, 2021, 38(11): 1883 – 1890.)
- [20] LI J. Application of BP NN algorithm in computer network security evaluation. *Computer Informatization and Mechanical System*, 2023, 6(4): 31 – 34.
- [21] YU Benguo. Research on the limitations and improvement of BP neural network. *Journal of Shanxi Agricultural University (Natural Science Edition)*, 2009, 29(1): 89 – 93.
(余本国. BP神经网络局限性及其改进的研究. 山西农业大学学报(自然科学版), 2009, 29(1): 89 – 93.)
- [22] PAO Y, PARK G, SOBAJIC D. Learning and generalization characteristics of the random vector functional-link net. *Neurocomputing*, 1994, 6(2): 163 – 180.
- [23] IGELNIK B, PAO Y. Stochastic choice of basis functions in adaptive function approximation and the functional-link net. *IEEE Transactions on Neural Network*, 1995, 6(6): 1320 – 1329.
- [24] XU Linfengyi, OU Guofu, JIN Haozhe. H₂S concentration prediction of delayed coking start-up line based on optimized RVFLN model. *Journal of Petroleum (Petroleum Processing)*, 2018, 34(6): 1127 – 1135.
(许霖凤一, 偶国富, 金浩哲. 基于优化RVFLN模型的延迟焦化开工线H₂S浓度预测. 石油学报(石油加工), 2018, 34(6): 1127 – 1135.)
- [25] MRUTYUNJAYA S, DASH P. Fault location estimation for series-compensated double-circuit transmission line using EWT and weighted RVFLN. *Engineering Applications of Artificial Intelligence*, 2020, 88: 1 – 11.
- [26] DASH D, DASH P, BISOI R. Short term solar power forecasting using hybrid minimum variance expanded RVFLN and sine-cosine levy flight PSO algorithm. *Renewable Energy*, 2021, 174: 513 – 537.
- [27] SAHANI M, DASH P. Deep convolutional stack autoencoder of process adaptive VMD data with robust multikernel RVFLN for power quality events recognition. *IEEE Transactions on Instrumentation and measurement*, 2021, 70: 3054673.
- [28] YU X, LIAN C, SU Y, et al. Selective ensemble deep bidirectional RVFLN for landslide displacement prediction. *Natural Hazards*, 2022, 112: 1 – 21.
- [29] WEI D, HU J, CHENG Y, et al. RVFLN-based online adaptive semi-supervised learning algorithm with application to product quality estimation of industrial processes. *Journal of Central South University*, 2019, 26(12): 3338 – 3350.
- [30] DAI W, LIU Q, CHAI T. Particle size estimate of grinding processes using random vector functional link networks with improved robustness. *Neurocomputing*, 2015, 169(112): 361 – 372.
- [31] RAO C, MITRA S. Further contributions to the theory of generalized inverse of matrices and its applications. *The Indian Journal of Statistics, Series A (1961–2002)*, 1971, 33(3): 289 – 300.
- [32] VENKATARAMANAN B. System identification: Theory for the user (second edition). *Automatica*, 2002, 38(2): 375 – 378.
- [33] LATECKI L, LAZAREVIC A, POKRAJAC D. Outlier detection with kernel density functions. *The 5th International Conference on Machine Learning and Data Mining in Pattern Recognition*. Leipzig, Germany: Springer-Verlag, 2007, (4571): 61.
- [34] LI Shunli, YAO Tingfu, YU Ping, et al. The application of singular value decomposition of matrix in image compression. *Computer Knowledge and Technology*, 2022, 18 (19): 1 – 2, 8.
(李顺利, 姚廷富, 余萍, 等. 矩阵的奇异值分解在图像压缩中应用. 电脑知识与技术, 2022, 18(19): 1 – 2, 8.)
- [35] YIN Xunlong, MOU Zonglei, WANG Youqing. Fault diagnosis of rotating machinery based on DVMD noise reduction. *Control Theory & Applications*, 2022, 39(7): 1324 – 1334.
(尹逊龙, 牟宗磊, 王友清. 基于DVMD降噪的旋转机械故障诊断. 控制理论与应用, 2022, 39(7): 1324 – 1334.)

- [36] LIU Yaming, GUO Hongxiang, CHEN Yanhu, et al. Random singular value decomposition based on photon computation. *Acta Optica Sinica*, 2022, 42(19): 1 – 12.
(刘雅名, 郭宏翔, 陈彦虎, 等. 基于光子计算的随机奇异值分解. 光学学报, 2022, 42(19): 1 – 12.)
- [37] LI Y, SHEN Q, JIANG M, et al. Accelerating dynamic MRI reconstruction using adaptive sequentially truncated higher-order singular value decomposition. *Current Medical Imaging*, 2022, 18(12): 1 – 12.
- [38] JIN Kejun, YU Hongtao, WU Yiteng, et al. A graph convolution network defense method based on improved singular value decomposition. *Computer Applications*, 2022, 43(5): 1 – 9.
(金柯君, 于洪涛, 吴翼腾, 等. 基于改进奇异值分解的图卷积网络防御方法. 计算机应用, 2022, 43(5): 1 – 9.)
- [39] HUANG G, LIU Y, YIN F. Tikhonov regularization with MTRSVD method for solving large-scale discrete ill-posed problems. *Journal of Computational and Applied Mathematics*, 2022, 405: 113969.
- [40] LIU Shiqi, WANG Yajing, MEI Yu, et al. Inter harmonic detection method based on ELMD-SVD and Prony. *Electric Drive*, 2022, 52(13): 48 – 55.
(刘士琦, 王雅静, 梅宇, 等. 基于ELMD-SVD和Prony的谐波间谐波检测方法. 电气传动, 2022, 52(13): 48 – 55.)
- [41] SHENG Z, ZENG Z, QU H, et al. Optical fiber intrusion signal recognition method based on TSVD-SCN. *Optical Fiber Technology*, 2019, 48: 270 – 277.
- [42] YIN Xiaoyan, YANG Dandan. Properties of singular values of matrices. *Studies in College Mathematics*, 2021, 24(1): 56 – 58.
(尹小艳, 杨丹丹. 关于矩阵奇异值的若干性质. 高等数学研究, 2021, 24(1): 56 – 58.)
- [43] TANG J. Optimal selection of learning parameters for regularised random vector functional-link networks-based soft measuring model. *International Journal of Machine Intelligence and Sensory Signal Processing*, 2018, 2(2): 140 – 158.

作者简介:

于洋 博士, 讲师, 硕士生导师, 目前研究方向为复杂工业过程建模与优化, E-mail: 20180029@sau.edu.cn;

邓瑞 硕士研究生, 目前研究方向为复杂过程建模与优化, E-mail: dengrui1@stu.sau.edu.cn;

余刚 博士, 研究员, 目前研究方向为选冶过程建模与优化, E-mail: yugang@bgrimm.com;

庞新富 博士, 副教授, 硕士生导师, 目前研究方向为复杂过程建模与优化, E-mail: pangxf@sie.edu.cn.