

交互偏移支持矩阵机及其在滚动轴承故障诊断中的应用

徐海锋, 潘海洋, 郑近德, 童靳于

(安徽工业大学机械工程学院, 安徽 马鞍山 243002)

摘要: 针对滚动轴承微弱故障信息特征提取与分类的问题, 提出一种交互偏移支持矩阵机(Interactive Deviation Support Matrix Machine, IDSMM)。在 IDSMM 建模过程中, 将每个目标函数加入偏移参数, 从几何上看, 这使得每一类样本对应的超平面相对原位置进行偏移, 构造出一对具有偏移量的交互超平面, 且每一个超平面距离异类样本更远, 提高了模型的泛化能力。此外, IDSMM 引入多秩左右投影矩阵构造目标函数, 充分挖掘了矩阵中行列间的结构信息, 并使模型具有更好的数据拟合能力。采用两个滚动轴承数据集进行实验验证, 实验结果表明 IDSMM 在滚动轴承故障诊断中具有优异的分类性能。

关键词: 故障诊断; 滚动轴承; 交互偏移支持矩阵机; 多秩左右投影矩阵

中图分类号: TH165⁺.3; TH133.3 **文献标志码:** A **文章编号:** 1004-4523(2022)03-0760-11

DOI: 10.16385/j.cnki.issn.1004-4523.2022.03.026

引言

滚动轴承作为机械设备中的一类重要零部件, 其工作时往往处于高速旋转的运行状态, 长时间工作会致使滚动轴承等部件出现一定的疲劳损伤。因此, 如何对滚动轴承的健康状况进行智能监测和诊断, 已成为众多学者研究的热点内容与方向。

随着人工智能的发展, 基于模式识别的滚动轴承智能故障诊断方法得到了广泛的研究和应用, 如人工神经网络(Artificial Neural Networks, ANN)^[1]、K-最近邻(K-nearest Neighbor, KNN)^[2]和支持向量机(Support Vector Machine, SVM)^[3]等。上述方法通过不断完善, 已被应用于故障诊断、文本分类、图像识别等领域。SVM 作为其中最稳定的一种分类方法, 具有良好的泛化能力和稀疏性。在解决二分类问题时, 传统的 SVM 通过将原始数据映射到一定维度特征空间, 构造两个平行超平面来分割不同类型样本。当原始数据线性不可分时, 如异或(Exclusive OR, XOR)问题, SVM 很难找到一对线性超平面对数据样本进行有效分割。针对该类问题, Li 等^[4]提出多目标近似支持向量机(Multi-task Proximal Support Vector Machine, MPSVM), 其通过求解两个广义特征值问题, 构造

出两个非平行的分类面对不同数据进行分割, 能够较好应对异或问题; Chen 等^[5]以经验风险最小化作为基本框架提出了孪生支持向量机(Twin Support Vector Machine, TWSVM), 其将 SVM 的优化问题转变为求解两个较小的线性规划问题, 构造一对非平行的分类面, 实现了分类效率的提升。为提高 TWSVM 算法的泛化能力, Wang 等^[6]在构造目标函数的过程中引入正则项, 提出孪生有界支持向量机(Twin Bounded Support Vector Machine, TBSVM), 其使用连续过度松弛技术(Successive Over-Relaxation, SOR)^[7]和正则项解决模型凸优化和鲁棒性问题。

上述方法本质上都是基于向量进行建模和分类, 当处理微弱故障信号的诊断问题时, 仅提取若干故障可能无法完成故障的有效识别。为了解决该问题, Luo 等^[8]受核范数的启发提出支持矩阵机(Support Matrix Machine, SMM), 以矩阵样本作为直接输入, 充分挖掘了矩阵行列间的结构信息, 这为后续研究提供了新的思路, 如鲁棒支持矩阵机(Robust Support Matrix Machine, RSMM)^[9]、稀疏支持矩阵机(Sparse Support Matrix Machine, SSMM)^[10]、多分类支持矩阵机(Multiclass Support Matrix Machine, MSMM)^[11]等。但这类矩阵分类器通常是线性的, 无法解决矩阵数据线性不可分的问题, 同时,

收稿日期: 2021-04-12; **修订日期:** 2021-06-03

基金项目: 国家自然科学基金资助项目(51975004); 安徽省高校自然科学研究重点项目(KJ2020A0231); 液压振动与控制教育部工程研究中心(安徽工业大学)开放基金资助项目(HVC201904)。

还存在训练误差大、过拟合等不足。鉴于此,Gao等^[12]基于矩阵数据构造非平行分类器的思想,利用左右奇异值向量构造目标函数提出了一种孪生多秩支持矩阵机(Twin Multiple Rank Support Matrix Machine, TMRSMM);Jiang等^[13]利用多秩左右投影矩阵替换左右投影向量来构造目标函数,提出了多秩多线性孪生支持矩阵机(Multiple Rank Multi-Linear Twin Support Matrix Machine, MRMLTSMM)。TMRSMM和MRMLTSMM利用左右投影思想解决了非线性矩阵数据的分类问题,且作为非平行超平面分类器(Nonparallel Hyperplane Classifier, NHC)的一类方法,它们在构造分类超平面 $\text{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 0$ 和 $\text{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = 0$ 时,遵循“寻找一对非平行的超平面,使得每一类样本聚集在相应超平面的附近,同时,另一类样本尽可能地远离该平面”的思想,提高了模型的分类精度。但是,当数据样本存在XOR问题时,仅依靠非平行超平面 $\text{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 0$ 和 $\text{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = 0$ 难以完成复杂数据的有效分割。

基于上述研究,考虑矩阵拥有保持数据结构信息完整的优势以及提高模型泛化能力的问题,本文提出一种新的矩阵数据分类方法——交互偏移支持矩阵机(Interactive Deviation Support Matrix Machine, IDSMM)。IDSMM以矩阵作为输入与建模元素,通过构造一对交互超平面 $\text{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$ 和 $\text{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = -1$ 将两类复杂样本分割开来,使得正类样本靠近超平面 $\text{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$,负类样本尽可能远离此平面。同时,引入多秩左右投影矩阵构造目标函数,使IDSMM具有更好的数据拟合能力。与TMRSMM和MRMLTSMM方法相比,IDSMM构造两个交互偏移超平面 $\text{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$ 和 $\text{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = -1$,从几何学上看,两平面间的“间隔”将变大,每一个超平面距离异类样本更远,提高了模型的泛化能力。最后,为了验证所提方法的性能,采用两种不同的滚动轴承故障数据集进行验证,实验结果表明,与其他分类器相比,IDSMM在滚动轴承故障诊断中具有更优异的分类性能。

1 交互偏移支持矩阵机原理

1.1 支持矩阵机

支持矩阵机是一种基于经验风险最小化的矩阵分类器,其利用铰链损失和正则项组合解决一个凸优化问题。同时,SMM的核心思想是寻找两个平行超平面,将正负两类样本分割开来。为了得到两

个超平面,构造目标函数如下式所示:

$$\underset{\mathbf{W}, b}{\operatorname{argmin}} \frac{1}{2} \operatorname{Tr}(\mathbf{W}^T \mathbf{W}) + \tau \|\mathbf{W}\|_* + C \sum_{i=1}^n \{1 - y_i [\operatorname{Tr}(\mathbf{W}^T \mathbf{X}_i) + b]\}_+ \quad (1)$$

式中 $\mathbf{W} \in \mathbf{R}^{m \times m}$ 为权重矩阵, $\{\mathbf{X}_i | i=1, \dots, n\} \in \mathbf{R}^{m \times n}$ 为矩阵样本, $y_i = (-1, 1)$ 为对应类别标签, b 为偏移量, $\tau \|\mathbf{W}\|_*$ 为正则项, τ 为正则项系数, $C \sum_{i=1}^n \{1 - y_i [\operatorname{Tr}(\mathbf{W}^T \mathbf{X}_i) + b]\}_+$ 为铰链损失函数, C 为损失参数。

利用交替方向乘子法ADMM可以解决目标函数(1)的求解问题,得到权重矩阵 \mathbf{W} 和偏移量 b 。然后构建如下决策函数,对于新的未知样本 $\tilde{\mathbf{X}}$,有:

$$\tilde{y} = \operatorname{sign}(\operatorname{Tr}(\mathbf{W}^T \tilde{\mathbf{X}}) + b) \quad (2)$$

SMM是一种可以充分利用输入矩阵样本结构特征的分类器,其核心思想还是寻找两个平行的分类面,面对一些复杂数据时,很难有效地找到两个平行分类超平面去划分两类样本。

1.2 交互偏移支持矩阵机

SMM方法的核心在于寻找一个最优分类面 $\operatorname{Tr}(\mathbf{W}^T \mathbf{X}) + b = 0$ 对不同数据进行划分,根据VC间隔理论,在距离超平面 $1/\|\mathbf{W}\|$ 处,分别存在上下边界 $\operatorname{Tr}(\mathbf{W}^T \mathbf{X}) + b = 1$ 和 $\operatorname{Tr}(\mathbf{W}^T \mathbf{X}) + b = -1$,两个边界间的距离称为最大间隔。由于这种平行超平面的构造形式,使得SMM无法有效处理一些复杂数据。为了解决这个问题,IDSMM采用一对交互偏移超平面,同时,不同于以往NHC方法的超平面构造形式,IDSMM拟将交互偏移超平面构造为 $\operatorname{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$ 和 $\operatorname{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = -1$ 。为了实现这个目标,尽可能使目标函数中的 $(\operatorname{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 - 1)^2$ 和 $(\operatorname{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 + 1)^2$ 最小。这两项在目标函数中表示每一类样本与相应超平面的平方距离的和。通过最小化这两项,使得正类样本尽可能地接近 $\operatorname{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$ 并远离 $\operatorname{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = -1$,同时,负类样本尽可能的接近 $\operatorname{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = -1$ 并远离 $\operatorname{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$ 。IDSMM方法的详细步骤如下:

对训练集 $T_1 = \{(\mathbf{X}_1, y_1), \dots, (\mathbf{X}_r, y_r), \dots, (\mathbf{X}_l, y_l)\}$,IDSMM通过构造如下式所示的优化问题来寻找一对交互超平面 $G(\mathbf{X})_1 = \operatorname{Tr}(\mathbf{W}_1^T \mathbf{X}) + b_1 = 1$ 和 $G(\mathbf{X})_2 = \operatorname{Tr}(\mathbf{W}_2^T \mathbf{X}) + b_2 = -1$ 。

$$\min_{\mathbf{W}_1, b_1, \xi} \frac{1}{2} \sum_{i=1}^r [\operatorname{Tr}(\mathbf{W}_1^T \mathbf{X}_i) + b_1 - 1]^2 + C_1 \sum_{j=r+1}^l \xi_j + \frac{C_3}{2} (\|\mathbf{W}_1\|^2 + b_1^2)$$

$$\begin{aligned} \text{s.t. } & -[\text{Tr}(\mathbf{W}_1^T \mathbf{X}_j) + b_1] \geq 1 - \xi_j, \\ & \xi_j \geq 0, j = r+1, \dots, l \end{aligned} \quad (3)$$

$$\begin{aligned} \min & \frac{1}{2} \sum_{j=r+1}^l [\text{Tr}(\mathbf{W}_2^T \mathbf{X}_j) + b_2 + 1]^2 + C_2 \sum_{i=1}^r \xi_i + \\ & \frac{C_4}{2} (\|\mathbf{W}_2\|^2 + b_2^2) \\ \text{s.t. } & -\text{Tr}[(\mathbf{W}_2^T \mathbf{X}_i) + b_2] \geq 1 - \xi_i, \\ & \xi_i \geq 0, i = 1, \dots, r \end{aligned} \quad (4)$$

式中 $\mathbf{X}_i \in \mathbf{R}^{m \times n}$ 为矩阵输入, $y_i \in \{1, -1\}$ 是相应的输出, 前 r 个是正类输入, 对应标签 $y_i = 1, i = 1, \dots, r$, 后 $l-r$ 个是负类输入, 对应标签为 $y_i = -1, i = r+1, \dots, l$, $\mathbf{W}_1, \mathbf{W}_2 \in \mathbf{R}^{m \times n}$ 为权重矩阵, b_1, b_2 为偏移量, C_1, C_2 为非负惩罚系数, C_3, C_4 为非负正则项系数, ξ_i 和 ξ_j 为非负松弛变量。

根据文献[14], 引入多秩左右投影矩阵 \mathbf{P} 和 \mathbf{Q} , 于是优化问题(3)和(4)转化为下式, 两个分类超平面转化为 $G(\mathbf{X})_1 = \text{Tr}(\mathbf{P}_1^T \mathbf{X} \mathbf{Q}_1) + b_1 = 1$ 和 $G(\mathbf{X})_2 = \text{Tr}(\mathbf{P}_2^T \mathbf{X} \mathbf{Q}_2) + b_2 = -1$ 。

$$\begin{aligned} \min_{\mathbf{P}_1, \mathbf{Q}_1, b_1, \xi_1} & \frac{1}{2} \sum_{i=1}^r [\text{Tr}(\mathbf{P}_1^T \mathbf{X}_i \mathbf{Q}_1) + b_1 - 1]^2 + C_1 \sum_{j=r+1}^l \xi_j + \\ & \frac{C_3}{2} (\|\mathbf{P}_1 \mathbf{Q}_1^T\|^2 + b_1^2) \\ \text{s.t. } & -[\text{Tr}(\mathbf{P}_1^T \mathbf{X}_j \mathbf{Q}_1) + b_1] \geq 1 - \xi_j, \\ & \xi_j \geq 0, j = r+1, \dots, l \end{aligned} \quad (5)$$

$$\begin{aligned} \min_{\mathbf{P}_2, \mathbf{Q}_2, b_2, \xi_2} & \frac{1}{2} \sum_{j=r+1}^l [\text{Tr}(\mathbf{P}_2^T \mathbf{X}_j \mathbf{Q}_2) + b_2 + 1]^2 + C_2 \sum_{i=1}^r \xi_i + \\ & \frac{C_4}{2} (\|\mathbf{P}_2 \mathbf{Q}_2^T\|^2 + b_2^2) \\ \text{s.t. } & -[\text{Tr}(\mathbf{P}_2^T \mathbf{X}_i \mathbf{Q}_2) + b_2] \geq 1 - \xi_i, \\ & \xi_i \geq 0, i = 1, \dots, r \end{aligned} \quad (6)$$

式中 $\mathbf{P}_1, \mathbf{P}_2$ 和 $\mathbf{Q}_1, \mathbf{Q}_2$ 分别是左、右投影矩阵, $\mathbf{P}_1 = [\mathbf{p}_{1,1}, \mathbf{p}_{1,2}, \dots, \mathbf{p}_{1,K}]_{m \times K}$, $\mathbf{P}_2 = [\mathbf{p}_{2,1}, \mathbf{p}_{2,2}, \dots, \mathbf{p}_{2,K}]_{m \times K}$, $\mathbf{Q}_1 = [\mathbf{q}_{1,1}, \mathbf{q}_{1,2}, \dots, \mathbf{q}_{1,K}]_{n \times K}$, $\mathbf{Q}_2 = [\mathbf{q}_{2,1}, \mathbf{q}_{2,2}, \dots, \mathbf{q}_{2,K}]_{n \times K}$, K 为 \mathbf{P} 和 \mathbf{Q} 的秩。

采用交替迭代算法来求解优化问题(5)和(6)。

固定 \mathbf{Q}_1 , 求解 \mathbf{P}_1 。设 $\mathbf{q}_{1,1}, \mathbf{q}_{1,2}, \dots, \mathbf{q}_{1,K}$ 均为非零

$$\text{向量, 记 } \hat{\mathbf{f}}_{1,i} = \begin{bmatrix} \mathbf{X}_i \mathbf{q}_{1,1} \\ \mathbf{X}_i \mathbf{q}_{1,2} \\ \vdots \\ \mathbf{X}_i \mathbf{q}_{1,K} \end{bmatrix}_{mK \times 1}, \hat{\mathbf{p}}_1 = \begin{bmatrix} \mathbf{p}_{1,1} \\ \vdots \\ \mathbf{p}_{1,K} \end{bmatrix}_{mK \times 1}, \hat{\mathbf{q}}_1 =$$

$$\begin{bmatrix} \mathbf{q}_{1,1} \\ \vdots \\ \mathbf{q}_{1,K} \end{bmatrix}_{nK \times 1}, D_1 = \begin{bmatrix} \mathbf{q}_{1,1}^T \mathbf{q}_{1,1} \mathbf{I} & & \\ & \ddots & \\ & & \mathbf{q}_{1,K}^T \mathbf{q}_{1,K} \mathbf{I} \end{bmatrix}_{mK \times mK} \quad (D_1 =$$

$(\mathbf{Q}_1^T \mathbf{Q}_1) \otimes \mathbf{I}$, \mathbf{I} 为 $m \times m$ 阶单位矩阵。

此时, 优化问题(5)可化简为下式:

$$\begin{aligned} \min_{\hat{\mathbf{p}}_1, b_1, \xi_1} & \frac{1}{2} \sum_{i=1}^r (\hat{\mathbf{p}}_1^T \hat{\mathbf{f}}_{1,i} + b_1 - 1)^2 + C_1 \sum_{j=r+1}^l \xi_j + \\ & \frac{C_3}{2} (\hat{\mathbf{p}}_1^T D_1 \hat{\mathbf{p}}_1 + b_1^2) \\ \text{s.t. } & -(\hat{\mathbf{p}}_1^T \hat{\mathbf{f}}_{1,j} + b_1) \geq 1 - \xi_j, \\ & \xi_j \geq 0, j = r+1, \dots, l \end{aligned} \quad (7)$$

记 $\mathbf{p}_1 = D_1^{1/2} \hat{\mathbf{p}}_1, \mathbf{f}_1 = D_1^{-1/2} \hat{\mathbf{f}}_1$, 于是优化问题(7)转化为下式:

$$\begin{aligned} \min_{\mathbf{p}_1, b_1, \xi_1} & \frac{1}{2} \sum_{i=1}^r (\mathbf{p}_1^T \mathbf{f}_{1,i} + b_1 - 1)^2 + C_1 \sum_{j=r+1}^l \xi_j + \\ & \frac{C_3}{2} (\|\mathbf{p}_1\|^2 + b_1^2) \\ \text{s.t. } & -(\mathbf{p}_1^T \mathbf{f}_{1,j} + b_1) \geq 1 - \xi_j, \\ & \xi_j \geq 0, j = r+1, \dots, l \end{aligned} \quad (8)$$

式(8)可进一步化为下式:

$$\begin{aligned} \min_{\mathbf{p}_1, b_1, \xi_2} & \frac{1}{2} \|\mathbf{M} \mathbf{p}_1 + \mathbf{e}_1 b_1 - \mathbf{e}_1\|^2 + C_1 \xi_2^T \mathbf{e}_2 + \\ & \frac{C_3}{2} (\|\mathbf{p}_1\|^2 + b_1^2) \\ \text{s.t. } & -(\mathbf{N} \mathbf{p}_1 + \mathbf{e}_2 b_1) \geq \mathbf{e}_2 - \xi_2, \\ & \xi_2 > 0 \end{aligned} \quad (9)$$

式中 $\mathbf{M} = [\mathbf{f}_{1,1}, \dots, \mathbf{f}_{1,r}]_{r \times mK}$, $\mathbf{N} = [\mathbf{f}_{1,r+1}, \dots, \mathbf{f}_{1,l}]_{(l-r) \times mK}$, $\xi_2 = [\xi_{r+1}, \xi_{r+2}, \dots, \xi_l]^T$, $\mathbf{e}_1 \in \mathbf{R}^{r \times 1}$ 和 $\mathbf{e}_2 \in \mathbf{R}^{(l-r) \times 1}$ 是向量分量全为 1 的向量。

引入拉格朗日乘子 α, β , 优化问题(9)的拉格朗日函数为:

$$\begin{aligned} L(\mathbf{p}_1, b_1, \alpha, \beta, \xi_2) &= \frac{1}{2} \|\mathbf{M} \mathbf{p}_1 + \mathbf{e}_1 b_1 - \mathbf{e}_1\|^2 + \\ & C_1 \xi_2^T \mathbf{e}_2 - \alpha^T [-(\mathbf{N} \mathbf{p}_1 + \mathbf{e}_2 b_1) + \xi_2 - \mathbf{e}_2] + \\ & \frac{C_3}{2} (\|\mathbf{p}_1\|^2 + b_1^2) - \beta^T \xi_2 \end{aligned} \quad (10)$$

利用 Karush-Kuhn-Tucher(KKT) 条件可得:

$$\nabla_{\mathbf{p}_1} L = \mathbf{M}^T (\mathbf{M} \mathbf{p}_1 + \mathbf{e}_1 b_1 - \mathbf{e}_1) + C_3 \mathbf{p}_1 + \mathbf{N}^T \alpha = 0 \quad (11)$$

$$\nabla_{b_1} L = \mathbf{e}_1^T (\mathbf{M} \mathbf{p}_1 + \mathbf{e}_1 b_1 - \mathbf{e}_1) + C_3 b_1 + \mathbf{e}_2^T \alpha = 0 \quad (12)$$

$$\nabla_{\xi_2} L = C_1 \mathbf{e}_2 - \beta - \alpha = 0 \quad (13)$$

$$-(\mathbf{N} \mathbf{p}_1 + \mathbf{e}_2 b_1) \geq \mathbf{e}_2 - \xi_2, \xi_2 \geq 0 \quad (14)$$

$$\alpha^T (\mathbf{N} \mathbf{p}_1 + \mathbf{e}_2 b_1 - \xi_2 + \mathbf{e}_2) = 0, \beta^T \xi_2 = 0 \quad (15)$$

$$\alpha \geq 0, \beta \geq 0 \quad (16)$$

由式(13)和(16)可以得到:

$$0 \leq \alpha \leq C_1 \mathbf{e}_2 \quad (17)$$

设 $\mathbf{A} = [\mathbf{M} \ \mathbf{e}_1]$, $\mathbf{B} = [\mathbf{N} \ \mathbf{e}_2]$, 由式(11)和(12)可以得到:

$$(\mathbf{A}^T \mathbf{A} + C_3 \mathbf{I}) [\mathbf{p}_1, b_1]^T - \mathbf{A}^T \mathbf{e}_1 + \mathbf{B}^T \alpha = 0 \quad (18)$$

式中 \mathbf{I} 为 $(mK+1) \times (mK+1)$ 的单位矩阵。

进一步求解式(18)得到下式:

$$[p_1, b_1]^T = (A^T A + C_3 I)^{-1} (A^T e_1 - B^T \alpha) \quad (19)$$

由式(13)可得 $\beta = C_1 e_2 - \alpha$, 左右两边同时乘以 ξ_2 得 $\beta^T \xi_2 = C_1 e_2^T \xi_2 - \alpha^T \xi_2$, 代入式(10)得下式:

$$L = \frac{1}{2} \|M p_1 + e_1 b_1 - e_1\|^2 + \frac{C_2}{2} (\|p_1\|^2 + b_1^2) - \alpha^T [-(N p_1 + e_2 b_1) - e_2] \quad (20)$$

将式(11)和(12)左右两边分别乘以 $\frac{1}{2} p_1$ 和 $\frac{1}{2} b_1$, 所得结果代入式(20)得到下式:

$$L = -\frac{1}{2} (M p_1 + e_2 b_1 - e_1) + \frac{1}{2} \alpha^T (N p_1 + e_2 b_1) + \alpha^T e_2 \quad (21)$$

结合 $A = [M \ e_1]$, $B = [N \ e_2]$ 以及式(19), 可得下式:

$$L = \frac{1}{2} \alpha^T B (A^T A + C_3 I)^{-1} B^T \alpha + \frac{1}{2} e_1 - (e_2^T + A (A^T A + C_3 I)^{-1} B^T) \alpha - \frac{1}{2} A (A^T A + C_3 I)^{-1} B^T \quad (22)$$

最后, 消去常数项 $-\frac{1}{2} A (A^T A + C_3 I)^{-1} B^T + \frac{1}{2} e_1$ 得到对偶问题如下式所示:

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \alpha^T B (A^T A + C_3 I)^{-1} B^T \alpha - [e_2^T + e_1^T A (A^T A + C_3 I)^{-1} B^T] \alpha, \\ \text{s.t.} & \quad e_2^T \alpha \geq 0, \quad 0 \leq \alpha \leq C_1 e_2 \end{aligned} \quad (23)$$

求解对偶问题(23)得到 α , 结合式(19)得到 p_1, b_1 和 $P_1 = \text{Mat}(D_1^{-\frac{1}{2}} p_1) = \text{Mat}(((Q_1^T Q_1) \otimes I_{m \times m})^{-\frac{1}{2}} p_1)$, $\text{Mat}(\cdot)$ 表示将向量转化为矩阵。

接着, 固定 P_1 , 求解 Q_1 。 Q_1 的优化求解过程与 P_1 过程相同。

设 $\hat{g}_{1,i} = [X_i q_{1,1}, X_i q_{1,2}, \dots, X_i q_{1,K}]_{nK \times 1}^T$, $\hat{D}_1 = (P_1^T P_1) \otimes I_{n \times n}$, $\hat{M} = [\hat{g}_{1,1}, \hat{g}_{1,2}, \dots, \hat{g}_{1,l}]_{r \times mK}^T$, $\hat{N} = [\hat{g}_{1,r+1}, \hat{g}_{1,r+2}, \dots, \hat{g}_{1,r+2}]_{(l-r) \times mK}^T$, $\hat{A} = [\hat{M} \ e_1]$, $\hat{B} = [\hat{N} \ e_2]$ 。重复上述求解 Q_1 步骤, 可得优化问题(5)的对偶问题:

$$\begin{aligned} \min_{\hat{\alpha}} & \frac{1}{2} \hat{\alpha}^T \hat{B} (\hat{A}^T \hat{A} + C_3 I)^{-1} \hat{B}^T \hat{\alpha} - (e_2^T + e_1^T \hat{A} (\hat{A}^T \hat{A} + C_3 I)^{-1} \hat{B}^T) \hat{\alpha} \\ \text{s.t.} & \quad e_2^T \hat{\alpha} \geq 0, \quad 0 \leq \hat{\alpha} \leq C_1 e_2 \end{aligned} \quad (24)$$

且

$$\begin{bmatrix} q_1 \\ b_1 \end{bmatrix} = (\hat{A}^T \hat{A} + C_3 I)^{-1} (\hat{A}^T e_1 - \hat{B}^T \hat{\alpha}) \quad (25)$$

I 为 $(nk+1) \times (nk+1)$ 的单位矩阵。

求解对偶问题(24)得到 $\hat{\alpha}$, 结合式(25)得

到 q_1, b_1 和 $Q_1 = \text{Mat}(\hat{D}_1^{-\frac{1}{2}} q_1) = \text{Mat}(((P_1^T P_1) \otimes I_{n \times n})^{-\frac{1}{2}} q_1)$ 。

接着, 通过交替优化迭代求 P_1, Q_1 和 b_1 。

同样, 按照以上步骤, 可以求出 P_2, Q_2 和 b_2 。优化问题(6)的对偶问题如下式所示:

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \alpha^T A (B^T B + C_4 I)^{-1} A^T \alpha - (e_1^T + e_2^T B (B^T B + C_4 I)^{-1} A^T) \alpha \\ \text{s.t.} & \quad e_1^T \alpha \geq 0, \quad 0 \leq \alpha \leq C_2 e_1 \end{aligned} \quad (26)$$

$$\begin{bmatrix} p_2 \\ b_2 \end{bmatrix} = (B^T B + C_4 I)^{-1} (-B^T e_2 + A^T \alpha) \quad (27)$$

$$\begin{aligned} \min_{\hat{\alpha}} & \frac{1}{2} \hat{\alpha}^T \hat{A} (\hat{B}^T \hat{B} + C_4 I)^{-1} \hat{A}^T \hat{\alpha} - (e_1^T + e_2^T \hat{B} (\hat{B}^T \hat{B} + C_4 I)^{-1} \hat{A}^T) \hat{\alpha} \\ \text{s.t.} & \quad e_1^T \hat{\alpha} \geq 0, \quad 0 \leq \hat{\alpha} \leq C_2 e_1 \end{aligned} \quad (28)$$

$$\begin{bmatrix} q_2 \\ b_2 \end{bmatrix} = (B^T \hat{B} + C_4 I)^{-1} (-\hat{B}^T e_2 + A^T \hat{\alpha}) \quad (29)$$

最后, 对于新的未知样本 \tilde{X} , 其标签 \tilde{y} 可由决策函数得到, 如下式所示:

$$\tilde{y} = \underset{y}{\text{argmin}} \left\{ \frac{|\text{Tr}(P_1^T \tilde{X} Q_1) + b_1 - 1|}{\|\text{Vec}(P_1 Q_1^T)\|}, \frac{|\text{Tr}(P_2^T \tilde{X} Q_2) + b_2 + 1|}{\|\text{Vec}(P_2 Q_2^T)\|} \right\} \quad (30)$$

算法 交互偏移支持矩阵机

训练过程:

输入: 训练集 $T_1 = \{(X_1, y_1), \dots, (X_r, y_r), \dots, (X_l, y_l)\}$, 参数 K , 惩罚参数 C_1, C_2 , 正则项系数 C_3, C_4 ;

输出: $P_1, P_2, Q_1, Q_2, b_1, b_2$;

步骤 1. 初始化 $Q_1 \leftarrow Q_1^{(k)}$, 令 $k = 0$;

步骤 2. 求解 QP 问题(23)得到 α ,

计算 $\begin{bmatrix} p_1^{(k)} \\ b_1^{(k)} \end{bmatrix} = (A^T A + C_3 I)^{-1} (A^T e_1 - B^T \alpha)$

$P_1^{(k)} = \text{Mat}(D_1^{-\frac{1}{2}} p_1) = \text{Mat}(((Q_1^T Q_1) \otimes I_{m \times m})^{-\frac{1}{2}} p_1)$;

步骤 3. $P_1 \leftarrow P_1^{(k)}$, 求解 QP 问题(24)得到解 $\hat{\alpha}$,

计算 $\begin{bmatrix} q_1^{(k)} \\ b_1^{(k)} \end{bmatrix} = (\hat{A}^T \hat{A} + C_3 I)^{-1} (\hat{A}^T e_1 - \hat{B}^T \hat{\alpha})$

$Q_1^{(k+1)} = \text{Mat}(\hat{D}_1^{-\frac{1}{2}} q_1) = \text{Mat}(((P_1^T P_1) \otimes I_{n \times n})^{-\frac{1}{2}} q_1)$;

步骤 4. 如果 $\|P_1^{(k+1)} Q_1^{(k+1)} - P_1^{(k)} Q_1^{(k)}\| \leq \epsilon$, 或者 $k = s$ (s 为设定最高迭代步数) 时, 停止, 并且 $P_1 \leftarrow P_1^{(k+1)}, Q_1 \leftarrow Q_1^{(k+1)}, b_1 \leftarrow b_1^{(k+1)}$, 否则令 $k = k + 1$ 并且返回步骤 2;

步骤 5. 求解优化问题(4)得到 P_2, Q_2 和 b_2 。

测试过程:

输入: 测试集 $T_2 = \{X_i | i = 1, \dots, n\}$, 设置参数

$P_1, P_2, Q_1, Q_2, b_1, b_2;$

输出: 测试标签: $\tilde{y}_i, i = 1, \dots, n$

$\tilde{y}_i = \operatorname{argmin}$

$$\left\{ \frac{|\operatorname{Tr}(P_1^T X_i Q_1) + b_1 - 1|}{\|\operatorname{Vec}(P_1 Q_1^T)\|}, \frac{|\operatorname{Tr}(P_2^T X_i Q_2) + b_2 + 1|}{\|\operatorname{Vec}(P_2 Q_2^T)\|} \right\}$$

2 实验分析

为了验证本文提出的 IDSMM 的分类性能, 本节分别使用安徽工业大学滚动轴承故障振动信号数据集和湖南大学锥齿轮-滚动轴承故障振动信号数据集对其进行测试和分析, 采用 Accuracy, Recall, Precision, F_1 -score 和 Kappa 作为模型评价指标。Accuracy 为正确率, 是最常用的分类性能衡量指标, 表示所有测试样本中分类器能正确分类的占比; Recall 表示召回率, 指在实际的测试样本中, 被分类器正确识别出来的正类样本占比; Precision 表示精确率, 指测试的正类样本中, 真正能被模型识别出来的属于正类的样本占比; F_1 -score 为精确率和召回率的调和值; Kappa 系数常用于一致性检验, 用于判断模型的分类精度。以上五种指标在一定范围内, 值越大, 表示模型的性能越好。

同时, 为了验证所提方法的优越性, 选择 SMM, SSMM, RSMM 和 MRMLTSMM 作为对比(所有实验均在 Intel(R) Core(TM) i5 (2.20 GHz), RAM 8 GB 计算机上的 Matlab2019b 运行)。由于 SMM, RSMM, SSMM, MRMLTSMM 和 IDSMM 等方法的输入元素为矩阵, 需要构造输入矩阵来完成分类和建模。辛几何相似变换方法^[15]作为一种新的信号降噪与分析方法, 已被证明具有良好的特征提取能力, 因此, 本文拟采用辛几何相似变换分析原始信号, 以获得可以保存完整结构信息的辛几何系数矩阵。

IDSMM 的分类流程结构如图 1 所示: 1) 将不同状态的数据作为训练集, 通过辛几何相空间重构后

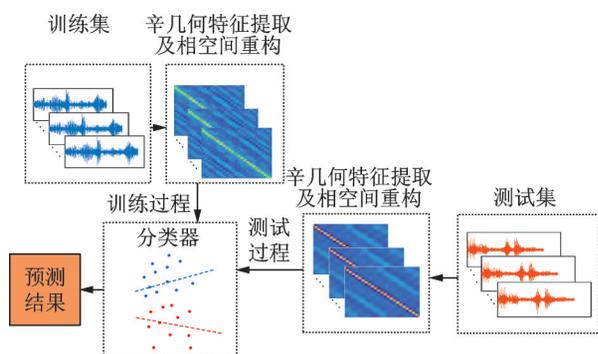


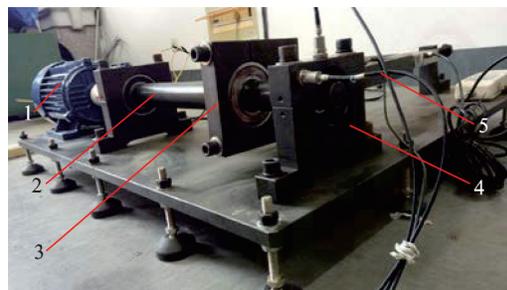
图 1 分类流程图

Fig. 1 Classification flow chart

输入到主程序, 得到决策函数(30); 2) 将不同状态的数据作为测试集输入到决策函数(30); 3) 对不同的模型采取同样的方法完成测试, 并进行对比分析。

2.1 实验 1

为了验证本文所提方法的分类性能, 选择安徽工业大学滚动轴承故障振动信号数据集进行分类测试。滚动轴承故障模拟试验台如图 2 所示。试验所用轴承型号为 SKF 6205 深沟球轴承, 使用电火花加工技术分别在轴承上布置出内圈故障、外圈故障和滚动体故障。试验过程中, 设置负载为 5 kN, 主轴转速为 900 r/min, 采样频率为 10240 Hz。同时, 利用加速度传感器分别采集滚动轴承在正常状态、内圈故障状态、外圈故障状态及滚动体故障状态下的加速度信号, 详细试验滚动轴承状态和试验参数见表 1(实验选择 260 组训练和测试样本)。



1. 变频电机; 2. 主轴; 3. 装载螺栓及其装载座; 4. 被测轴承及其轴承座; 5. 测力器。

图 2 滚动轴承故障模拟试验台

Fig. 2 Roller bearing fault simulation test bench

表 1 滚动轴承试验参数

Tab. 1 Experimental parameters of rolling bearings

故障类型	负载/kN	故障深度/mm	转速/(r·min ⁻¹)	采样频率/Hz	标签
正常状态		0			1
内圈故障	5	0.3	900	10240	2
外圈故障		0.2			3
滚动体故障		0.2			4

IDSMM 的性能与参数 K, C_1, C_2, C_3 和 C_4 相关, 因此需要在训练前选择最优值。首先, 根据文献[6, 13]相关取值方法, 通过实验取 $C_1=0.1, C_2=0.1, C_3=0.01$ 和 $C_4=0.01$ 。此外, 对于参数 K , 设置 K 的取值范围为 1~10, 选取 180 组训练和测试样本并采用 5 折交叉验证进行实验。图 3 为 IDSMM 在 K 为

1~10 范围内得到的五种指标对比,表 2 为训练时间。

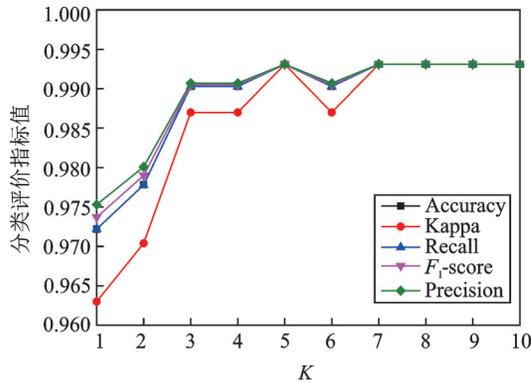


图 3 不同 K 的 IDSMM 五种指标对比

Fig. 3 Comparison of five indicators of IDSMM in different K

表 2 不同 K 的 IDSMM 训练时间

Tab. 2 IDSMM training time in different K

K	训练时间/s	K	训练时间/s
1	0.28	6	6.22
2	0.76	7	7.71
3	1.15	8	9.19
4	2.48	9	12.01
5	4.10	10	14.24

从图 3 中可以看出,IDSMM 分类器在不同的 K 下均能够保持较高的准确率。总体上,随着 K 的增大,各个指标逐渐增大,最终保持平稳状态,且 K=5 或 K≥6 时分类准确率最高。IDSMM 采用交替迭代求解左右投影矩阵,随着 K 值的增大, P₁, P₂, Q₁ 和 Q₂ 的矩阵尺寸增大,模型计算的复杂度增大,从而使得耗时越来越长。因此,为了使 IDSMM 分类性能达到最佳的同时缩短训练时间,选择 K=5 作为模型的输入参数,此时五种指标均处于较高水平,训练用时相对较短。

IDSMM 采用 SOR 方法求解对偶问题,通过交替迭代求解最优 P₁, P₂, Q₁ 和 Q₂。图 4 展示了该模型在训练过程中交替求解 P₁, P₂, Q₁ 和 Q₂ 的收敛过程曲线。由于 P₁, Q₁ 和 P₂, Q₂ 分别求解,所以会有两条不同的收敛曲线,其中, P₁ 和 Q₁ 表示求解 P₁ 和 Q₁ 过程中目标函数(5)的收敛曲线, P₂ 和 Q₂ 表示求解 P₂ 和 Q₂ 过程中目标函数(6)的收敛曲线。从图 4 可以看出,IDSMM 经过较少的迭代次数可以快速达到收敛状态。

SMM, RSMM, SSMM 和 MRMLTSMM 模型中均具有不同的参数,为了保障测试的客观性,采用 5 折交叉选择各个参数,五种分类器的最终参数选择如表 3 所示。然后,分别使用 IDSMM, SMM, RSMM, SSMM 和 MRMLTSMM 对 100 个样本进

行训练,得到五种预测模型,再用 160 个样本进行测试,识别结果如图 5 所示,从图中可以看出,IDSMM 将第 1 类和第 4 类的样本错分,其他类别全部被正确分类。SMM, RSMM 和 SSMM 存在较多第 4 类错分给第 3 类的样本,MRMLTSMM 方法错分率较低。表 4 为五种分类器的五种指标对比,从表中可以看出,IDSMM 五个指标均最高,表示 IDSMM 模型有着最好的分类性能。

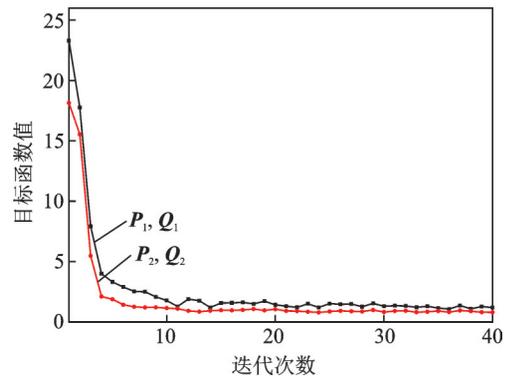


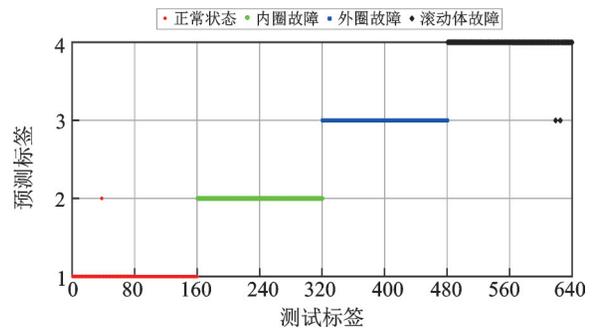
图 4 IDSMM 收敛曲线

Fig. 4 IDSMM convergence curve

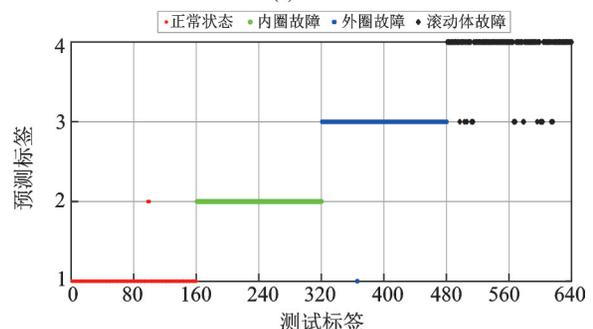
表 3 五种方法的参数

Tab. 3 Parameters of five methods

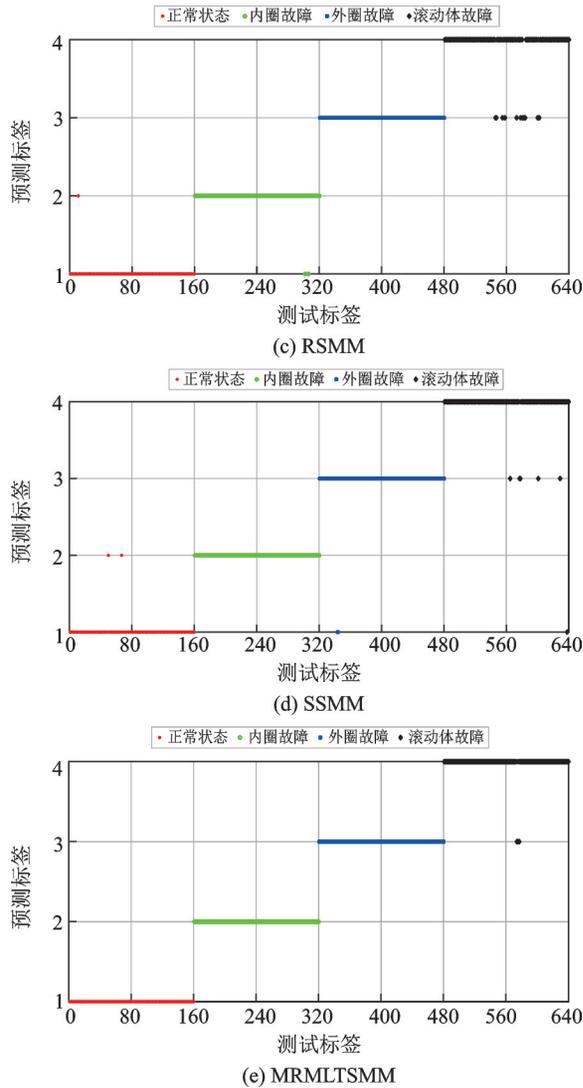
分类器	模型参数
IDSMM	C ₁ =C ₂ =0.1, C ₃ =C ₄ =0.01, K=5
MRMLTSMM	C ₁ =C ₂ =1, C ₃ =C ₄ =0.01, K=3
SMM	C=0.3, τ=10
RSMM	λ ₁ =0.5, λ ₂ =0.01, λ ₃ =0.001
SSMM	γ=0.001, τ=0.01



(a) IDSMM



(b) SMM



1, 2, 3, 4 分别表示正常状态、内圈故障、外圈故障和滚动体故障。

图5 五种方法分类结果

Fig. 5 Classification results of five methods

表4 五种方法的分类性能对比

Tab. 4 Comparison of classification performance of five methods

分类方法	Accuracy	Kappa	Recall	F_1 -score	Precision
SMM	0.9609	0.9479	0.9609	0.9636	0.9662
RSMM	0.9719	0.9625	0.9719	0.9733	0.9747
SSMM	0.9875	0.9833	0.9875	0.9878	0.9881
MRMLTSMM	0.9922	0.9896	0.9922	0.9923	0.9924
IDSMM	0.9953	0.9938	0.9953	0.9954	0.9954

为避免偶然因素的发生,得到更好的测试结果,选择260个样本采用5折交叉验证方法再次进行测试。图6展示了五种分类器分别在5折交叉验证下得到的结果(1~5表示每次的分类准确率)。从图6中可以看出,IDSMM分类性能最好,平均分类正确率达

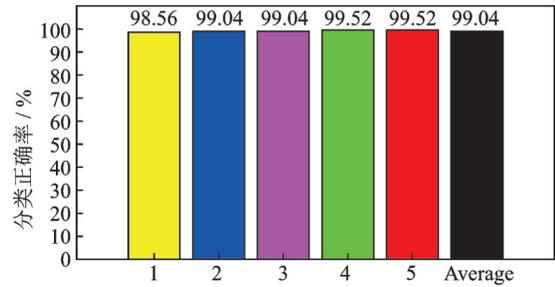
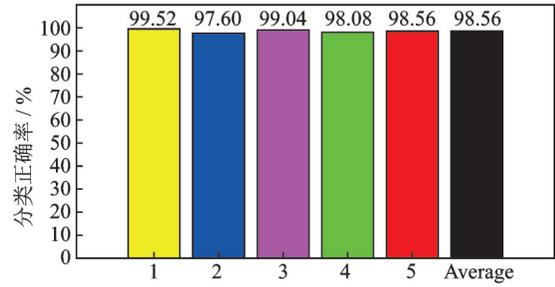
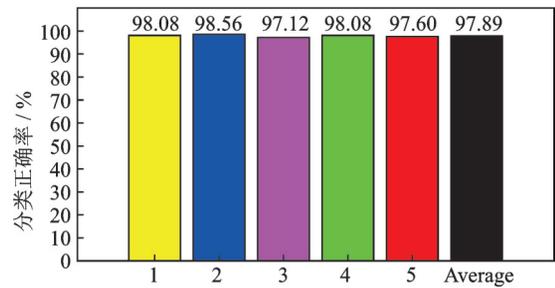
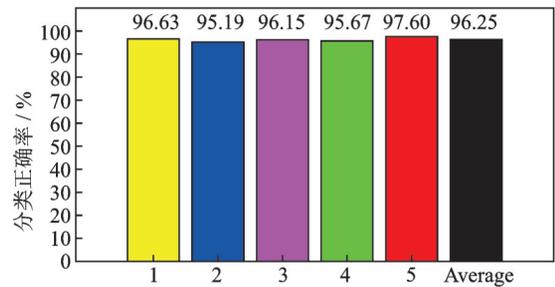
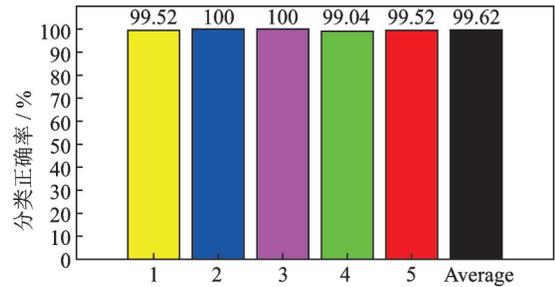


图6 5折交叉验证下五种方法的结果

Fig. 6 Results of five methods in 5-CV

到99.62%,SMM,RSMM,SSMM和MRMLTSMM的平均分类正确率分别为96.25%,97.89%,98.56%和99.04%。IDSMM方法在构造分类面时,其为不同的样本分别构造对应的分类面,得到一对交互偏移分类超平面,能够更有效地处理复杂数据,同时,

IDSMM 采用左右投影矩阵方法,降低了过拟合的影响。其次,相比 MRMLTSMM,IDSMM 的分类面间的间隔增大,从而使分类性能提高。SSMM 方法同样采用构造一对交互超平面进行分类,同时其模型本身具有稀疏属性,也实现了较为出色的分类性能。尽管 RSMM 在建构模型时提高了鲁棒性,但在此复杂数据下未表现出良好的分类性能。SMM 是一种基于经验风险最小化的分类器,其使用常规的方法来构造两个平行的分类超平面,模型的泛化能力较差,因此分类结果不佳。

综上所述,IDSMM 分类方法具有良好的分类性能,其模型参数在一定范围内,对整体模型的影响有限。同时,实验验证了其收敛的有效性,在较短的时间内,该方法能够得到最优解并完成训练和测试。

2.2 实验 2

IDSMM 在安徽工业大学滚动轴承故障振动信号数据集上展示了良好的分类性能,为了进一步验证论文提出方法的性能,选择湖南大学锥齿轮-滚动轴承故障振动信号数据集再次进行分类测试。试验采用的轴承为型号 SKF 6206。试验过程中,设置载荷为 4 kN,电机转速为 900 r/min,采样频率为 8192 Hz。轴承故障信息和试验参数如表 5 所示(选择 350 组训练和测试样本)。

表 5 滚动轴承试验参数

Tab. 5 Experimental parameters of rolling bearings

故障类型	负载/ kN	故障 深度/mm	转速/ (r·min ⁻¹)	采样 频率/Hz	标 签
正常状态		0			1
内圈故障 1		0.4			2
内圈故障 2		0.8			3
外圈故障 1	4	0.4	900	8192	4
外圈故障 2		0.8			5
滚动体故障 1		0.4			6
滚动体故障 2		0.8			7

图 7 展示了 IDSMM 方法在 K 为 1~10 范围内选择 180 组训练和测试样本进行实验得到的五种指标对比,表 6 为训练时间。从图 7 中可以看出,当 $K=4$,模型分类正确率达到 100%。随着 K 增大,IDSMM 分类性能明显提升, $K \geq 3$ 以后,分类准确率、召回率、 F_1 -score 和 Kappa 均处于较高水平。说明该模型在 $K \geq 4$ 的情况下能达到最佳的分类性能,但 K 值越大,训练用时越久。基于以上分析,选择 $K=4$ 作为模型的输入参数。

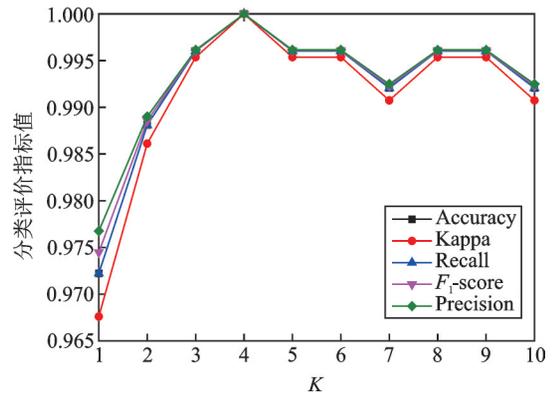


图 7 不同 K 的 IDSMM 的五种指标对比

Fig. 7 Comparison of five indicators of IDSMM in different K

表 6 不同 K 的 IDSMM 训练时间

Tab. 6 IDSMM training time in different K

K	训练时间/s	K	训练时间/s
1	0.10	6	3.27
2	0.38	7	4.30
3	0.97	8	6.41
4	1.59	9	8.27
5	2.35	10	10.59

选择 150 组样本作为训练集,剩余 200 组作为测试集,分别使用五种方法进行实验,实验结果如表 7、图 8 和 9 所示。从表 7 和图 8 中可以看出,IDSMM 五种指标均高于 SMM,RSMM,SSMM 和 MRMLTSMM,再次证明了该方法有着良好的分类性能。图 9 为根据分类结果得到的混淆矩阵。从图中可以清楚看出,在增加故障类别,提高数据复杂度后,IDSMM 存在较少的错分现象,但第 1 类和第 5 类仍被全部正确区分。考虑到偶然因素的存在,再次选择 350 组样本进行 5 折交叉验证测试,图 10 展示了五种方法在 5 折交叉验证下的测试结果。从图中可以看出,IDSMM 方法具有较高的分类正确率,平均正确率可以达到 99.80%,SMM,RSMM,SSMM 和 MRMLTSMM 的平均分类正确率分别为 97.43%,98.53%,99.14% 和 99.63%。相比其他四种方法,IDSMM 方法在 5 次分类中的正确率均稳定在 99% 以上,识别效果最佳。究其原因在于,SMM 利用输入的特征样本构建目标函数,特征数的增加对其模型的影响较大,使得分类准确率波动较大。SSMM 和 RSMM 通过增加稀疏属性和约束使得模型具有鲁棒性以及冗余特征的抗干扰能力,但是,SSMM 和 RSMM 本质上还是构造平行超平面,对复杂数据难以完成有效的样本分割。

为了判断 IDSMM 方法的小样本分析能力,随机抽取 10,20,30,40,50,60,70,80,90,100 个样本作为训练样本,150 个测试样本。IDSMM,SMM,

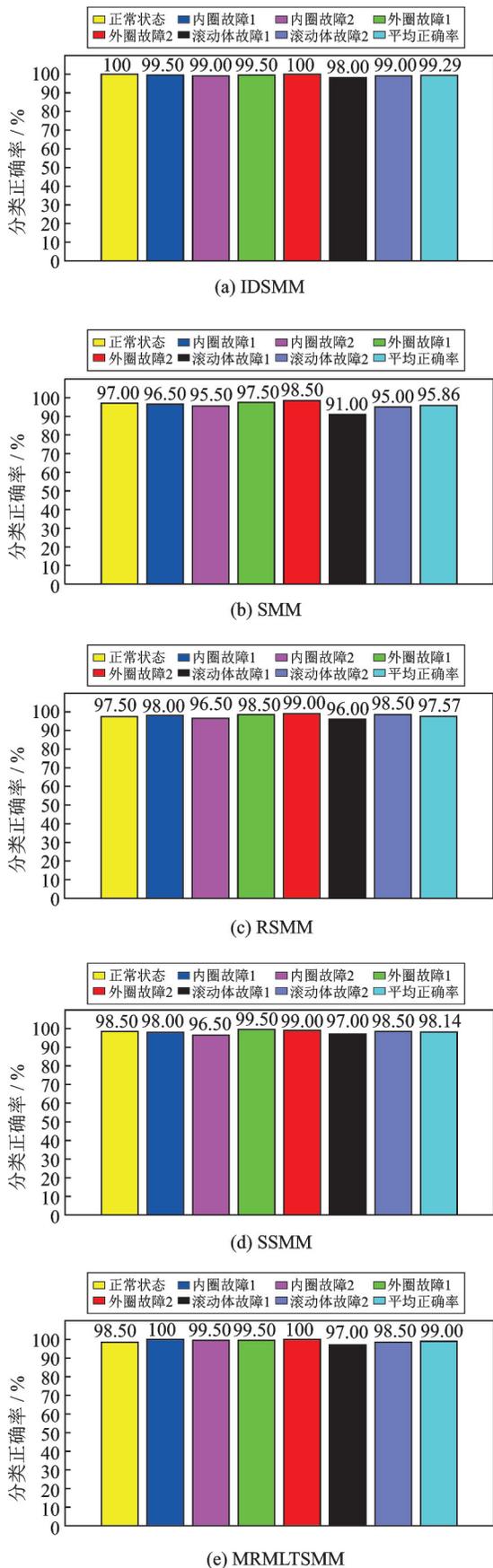


图8 五种方法的分类结果

Fig. 8 Classification results of the five methods

RSMM, SSMM 和 MRMLTSMM 五种方法得分类结果如图 11 所示。从图 11 中可以看出,IDSMM 和 MRMLTSMM 方法都具有优异的小样本数据分析

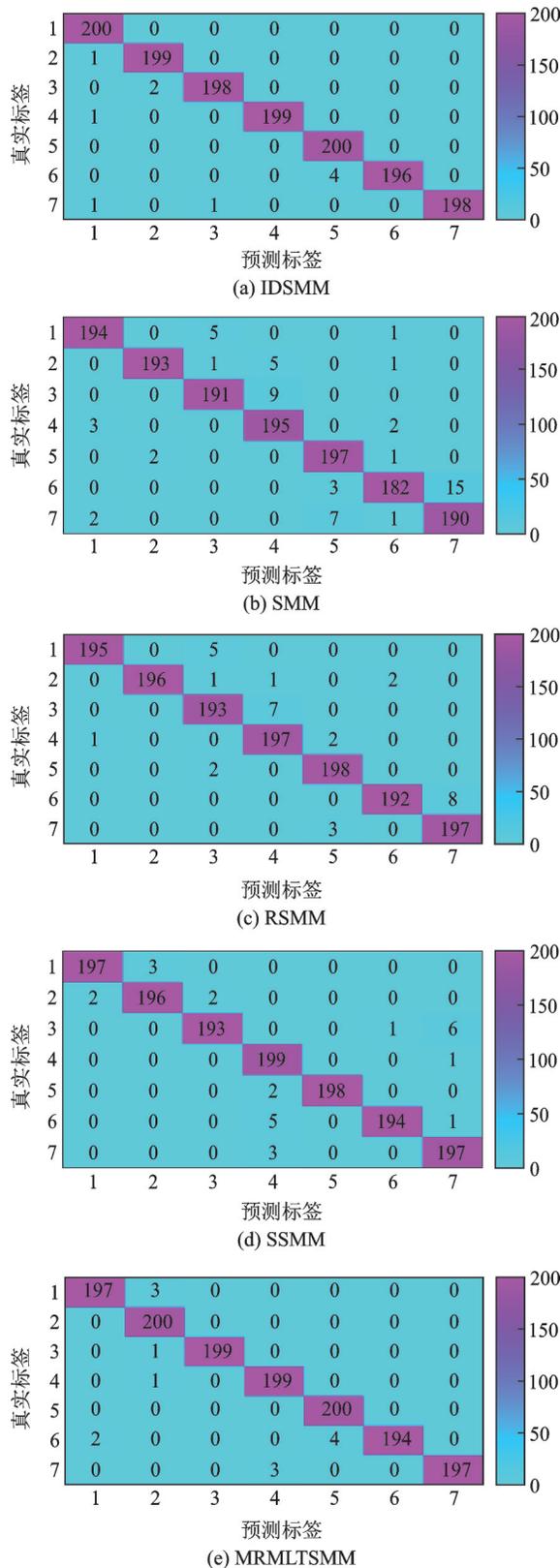


图9 五种方法的混淆矩阵

Fig. 9 Confusion matrix of five methods

能力,当训练样本数为10~40时,IDSMM 分类正确率比 MRMLTSMM 略高。总的来说,无论训练样本数多少,IDSMM 和 MRMLTSMM 分类准确率都高于其他三种方法。SSMM 和 RSMM 的识别性能低于 IDSMM,但也表现出了良好的小样本分析能力。由于 SMM 模型具有较差的鲁棒性和泛化能

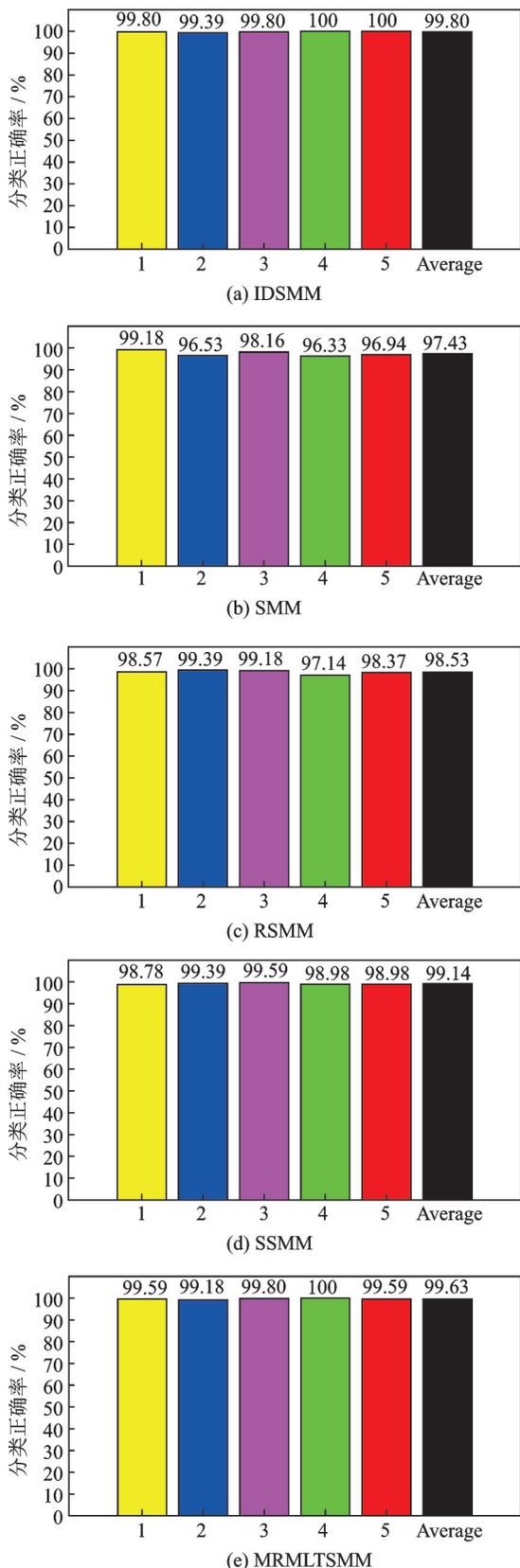


图 10 5折交叉验证五种方法的结果
Fig. 10 Results of five methods in 5-CV

力,其在任一训练样本数下的分类正确率都比其他四种方法低,表明SMM小样本分类性能最差。

综上所述,根据五种方法在5种衡量指标下的对比结果可知,IDSMM的诊断性能明显优于SMM,RSMM,SSMM和MRMLTSMM。IDSMM

表 7 五种方法的分类性能对比

Tab. 7 Comparison of classification performance of five methods

分类方法	Accuracy	Kappa	Recall	F ₁ -score	Precision
SMM	0.9586	0.9517	0.9586	0.9632	0.9679
RSMM	0.9757	0.9717	0.9757	0.9775	0.9792
SSMM	0.9814	0.9783	0.9814	0.9825	0.9836
MRMLTSMM	0.9900	0.9883	0.9900	0.9903	0.9907
IDSMM	0.9929	0.9917	0.9929	0.9990	0.9932

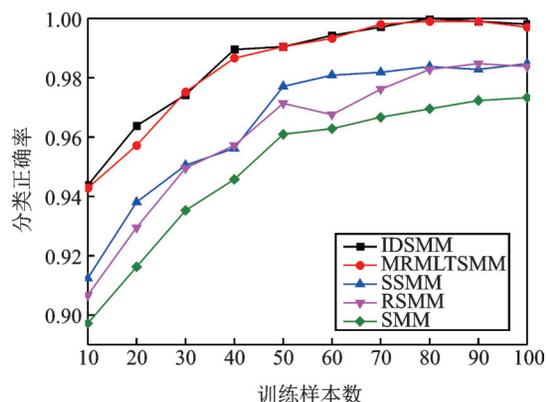


图 11 不同训练样本数的分类正确率

Fig. 11 Classification accuracy of different numbers of training samples

采用一种交互偏移超平面的方式构造两个超平面,当数据较为复杂时,可以为每一类特征样本构造一类超平面。同时,由于本类样本所属平面与异类样本的距离被拉大,因此具有更好的分类性能。SMM作为一种平行超平面分类器,当输入数据包含多种特征信息时,其无法有效构造合适的超平面。SSMM构建模型的前提是回归矩阵具有低秩特性,因此在面对大多数矩阵是多秩的情况下很难发挥模型本身的优势,而IDSMM引入了左右多秩投影矩阵,可以更好地解决该问题。相比其他平行超平面分类器,IDSMM因其借助NHC思想构造超平面的方式,在处理二分类问题尤其是复杂问题时具有明显的优势。

3 结 论

本文提出一种交互偏移支持矩阵机(IDSMM)的矩阵分类方法。不同于传统非平行超平面的构造形式,IDSMM在建模过程中,在每个目标函数中加入偏移参数,以改变约束,构造出一对交互偏移超平面 $\text{Tr}(\mathbf{W}_1^T \mathbf{X}^{(1)}) + b_1 = 1$ 和 $\text{Tr}(\mathbf{W}_2^T \mathbf{X}^{(2)}) + b_2 = -1$ 。交互偏移超平面克服了复杂数据间的有效分割问题,并且每一个超平面距离异类样本更远,提高了模

型的泛化能力。同时,IDSMM使用左右投影矩阵构造目标函数,避免了数据过拟合问题。通过对两个滚动轴承数据集进行实验验证,实验结果表明,与SMM,RSMM,SSMM和MRMLTSMM相比,IDSMM在滚动轴承故障诊断中具有更优异的性能。

参考文献:

- [1] Li Wendi, Ng Wing W Y, Wang Ting, et al. HELP: an LSTM-based approach to hyperparameter exploration in neural network learning [J]. *Neurocomputing*, 2021, 442: 161-172.
- [2] Mashaan Alshammari, John Stavrakakis, Masahiro Takatsuka. Refining a k-nearest neighbor graph for a computationally efficient spectral clustering [J]. *Pattern Recognition*, 2021, 114: 107869.
- [3] Lee Seunggyu. Monte Carlo simulation using support vector machine and kernel density for failure probability estimation [J]. *Reliability Engineering & System Safety*, 2021, 209: 107481.
- [4] Li Ya, Tian Xinmei, Song Mingli, et al. Multi-task proximal support vector machine [J]. *Pattern Recognition*, 2015, 48(10): 3249-3257.
- [5] Chen Weijie, Shao Yuanhai, Li Chunna, et al. MLTS-VM: a novel twin support vector machine to multi-label learning [J]. *Pattern Recognition*, 2016, 52: 61-74.
- [6] Wang Huiru, Zhou Zhijian. An improved rough margin-based ν -twin bounded support vector machine [J]. *Knowledge-Based Systems*, 2017, 128: 125-138.
- [7] Qi Zhiquan, Tian Yingjie, Shi Yong. Successive over-relaxation for Laplacian support vector machine [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, 26(4): 674-683.
- [8] Luo Luo, Xie Yubo, Zhang Zhihua, et al. Support matrix machines [C]. *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, Lille, France, 2015.
- [9] Zheng Qingqing, Zhu Fengyuan, Heng Pheng-Ann. Robust support matrix machine for single trial EEG classification [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, 26(3): 551-562.
- [10] Zheng Qingqing, Zhu Fengyuan, Qin Jing, et al. Sparse support matrix machine [J]. *Pattern Recognition*, 2018, 76: 715-726.
- [11] Zheng Qingqing, Zhu Fengyuan, Qin Jing, et al. Multi-class support matrix machine for single trial EEG classification [J]. *Neurocomputing*, 2018, 275: 869-880.
- [12] Gao Xizhan, Fan Liya, Xu Haitao. A novel method for classification of matrix data using Twin Multiple Rank SMMs [J]. *Applied Soft Computing*, 2016, 48: 546-562.
- [13] Jiang Rong, Yang Zhi-xia. Multiple rank multi-linear twin support matrix classification machine [J]. *Journal of Intelligent & Fuzzy Systems*, 2018, 35(5): 5741-5754.
- [14] Hou Chenping, Nie Feiping, Zhang Changshui, et al. Multiple rank multi-linear SVM for matrix data classification [J]. *Pattern Recognition*, 2014, 47(1): 454-469.
- [15] Pan Haiyang, Yang Yu, Li Xin, et al. Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis [J]. *Mechanical Systems and Signal Processing*, 2019, 114: 189-211.

Interactive deviation support matrix machine and its application in rolling bearing fault diagnosis

XU Hai-feng, PAN Hai-yang, ZHENG Jin-de, TONG Jin-yu

(School of Mechanical Engineering, Anhui University of Technology, Ma'anshan 243002, China)

Abstract: Aiming at the problem of feature extraction and classification of weak fault information of the roller bearing, this paper proposes an Interactive Deviation Support Matrix Machine (IDSMM). In the process of IDSMM modeling, the deviation parameter is added to each objective function. Geometrically, the hyperplane corresponding to each class of samples is deviation from the original position to construct a pair of interactive hyperplanes with deviation, and each hyperplane is further away from the heterogeneous samples, which improves the generalization ability of the model. In addition, IDSMM introduces a multi-rank left and right projection matrix to construct the objective function, which fully excavates the structural information between rows and columns in the matrix, and makes the model have better data fitting ability. Two data sets of roller bearings are used for experimental verification, the experimental results show that IDSMM has excellent classification performance in the fault diagnosis of the roller bearing.

Key words: fault diagnosis; rolling bearing; interactive deviation support matrix machine; multi-rank left and right projection matrix

作者简介: 徐海锋(1996—),男,硕士研究生。电话:18325690030;E-mail:haifenguuu@163.com。

通讯作者: 潘海洋(1989—),男,讲师。电话:18895558191;E-mail:pansea@sina.cn。