

DOI:10.11992/tis.2016030.
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20160513.0924.024.html>

一种改进的投影孪生支持向量机

花小鹏¹, 孙一颗¹, 丁世飞²

(1.盐城工学院 信息工程学院, 江苏 盐城 224051; 2. 中国矿业大学 计算机学院, 江苏 徐州 221116)

摘 要:针对投影孪生支持向量机(PTSVM)在训练阶段欠考虑样本空间局部结构和局部信息的缺陷,提出一种具有一定局部学习能力的有监督分类方法:加权投影孪生支持向量机(weighted PTSVM, WPTSVM)。相比于 PTSVM, WPTSVM 优势在于:通过构造类内近邻图为每个样本获取特定的权值,并且以加权均值取代标准均值,在一定程度上提高了算法的局部学习能力;选取异类样本集中少量边界点构造优化问题的约束条件,很大程度上降低了二次规划求解的时间复杂度;继承了 PTSVM 的优点,可以看成 PTSVM 的推广算法。理论分析及其在人造数据集和真实数据集上的测试结果表明该方法具有上述优势。

关键词:分类;投影孪生支持向量机;局部信息;加权均值;近邻图;二次规划;约束条件;时间复杂度

中图分类号:TP391.4 **文献标志码:**A **文章编号:**1673-4785(2016)03-0384-06

中文引用格式:花小鹏,孙一颗,丁世飞.一种改进的投影孪生支持向量机[J]. 智能系统学报, 2016, 11(3): 384-391.

英文引用格式:HUA Xiaopeng, SUN Yike, DING Shifei. An improved projection twin support vector machine[J]. CAAI transactions on intelligent systems, 2016, 11(3): 384-391.

An improved projection twin support vector machine

HUA Xiaopeng¹, SUN Yike¹, DING Shifei²

(1. School of Information Engineering, Yancheng Institute of Technology, Yancheng 224051, China; 2. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: A supervised classification method having a local learning ability, called weighted projection twin support vector machine (WPTSVM), is proposed. This method aims to improve upon a defect that projection twin support vector machines (PTSVMs) have, namely, that PTSVMs do not take account of the local structure and local information of a sample space in the training process. Compared with PTSVM, WPTSVM improves its local learning ability to some extent by attaching different weights for each sample according to the within-class neighborhood graph and replacing the standard mean with a weighted mean. Moreover, to reduce computational complexity, WPTSVM chooses a small number of boundary points in the contrary-class based on the between-class neighborhood graph to construct constraints of the original optimization problems. The method inherits the merits of PTSVM and can be regarded as an improved version of PTSVM. Experimental results on artificial and real datasets indicate the effectiveness of the WPTSVM method.

Keywords: classification; projection twin support vector machine; local information; weighted mean; neighborhood graph; quadratic programming; constraint condition; time complexity

在分类问题中,经典支持向量机(SVM)依据间

隔最大化准则生成分类决策面,存在训练时间复杂度偏高和欠考虑样本分布情况的缺陷^[1-2]。近年来,作为经典 SVM 的改进方法,非平行超平面分类器(nonparallel hyperplane classifiers, NHCs)^[3]已经成为模式识别领域新的研究热点之一。孪生支持向

收稿日期:2016-03-20. 网络出版日期:2016-05-13.

基金项目:国家重点基础研究计划项目(2013CB329502);国家自然科学基金项目(61379101);江苏省自然科学基金项目(BK20151299).

通信作者:花小鹏. E-mail: xp_hua@163.com.

量机 (twin support vector machine, TWSVM)^[4] 是 NHCs 方法中主要代表性算法之一,其主要思想源于泛化特征值中心支持向量机 (generalized eigenvalue proximal SVM, GEPSVM)^[5]。TWSVM 将 GEPSVM 中两个优化子问题转换成两个形如 SVM 的小规模二次规划问题,从而使其训练时间复杂度缩减为经典 SVM 的 1/4。除了训练速度上的优势,TWSVM 还继承了 GEPSVM 能够在线性模式很好地处理异或 (XOR) 样本分类问题的优势。然而,当两类样本具有不同散度分布时,TWSVM 的泛化性能欠佳^[6]。文献[7]提出一种新的非平行超平面分类器:投影孪生支持向量机 (projection twin support vector machine, PTSVM)。与 TWSVM 不同的是,PTSVM 优化目的是为每类样本寻找最佳投影轴,而且通过递归迭代算法,PTSVM 能够生成多个正交投影轴。实验结果表明,PTSVM 对复杂的 XOR 问题具有更好的分类能力。为解决非线性分类问题。文献[8]进一步提出 PTSVM 的非线性方法。

然而分析发现,PTSVM 在训练过程中仅仅考虑样本空间的全局结构和全局信息,忽视了样本空间的局部结构和局部信息。许多研究结果表明同类数据集中大部分样本在局部上是关联的(即数据集中存在潜藏的局部几何结构),而这种内在的局部信息对数据分类又是至关重要的^[9]。这种潜在的局部信息可以通过数据集中样本间的 k 近邻关系进行挖掘^[9-11]。

基于上述分析,本文基于 PTSVM 提出一种新的具有一定局部学习能力的非平行超平面分类器算法:加权投影孪生支持向量机 (weighted PTSVM, WPTSVM)。相比于 PTSVM,WPTSVM 优势体现在以下 4 个方面:1)通过构造类内近邻图为每个样本获取特定的权值,并且以加权均值取代标准均值,在一定程度上提高了算法的局部学习能力;2)选取异类样本集中少量边界点构造优化问题的约束条件,很大程度上降低了二次规划求解的时间复杂度;3)WPTSVM 继承了 PTSVM 的优点,可以看成 PTSVM 的推广算法。4)WPTSVM 具有更好分类性能。

$$W_{ij}^s = \begin{cases} \exp(-\| \mathbf{x}_i^{(c)} - \mathbf{x}_j^{(c)} \|^2 / t), & \mathbf{x}_j^{(c)} \in \text{Ne}(\mathbf{x}_i^{(c)}) \text{ 或 } \mathbf{x}_i^{(c)} \in \text{Ne}(\mathbf{x}_j^{(c)}), \\ 0, & \text{其他} \end{cases}$$

(3)

式中: t 为热核参数,Ne(\mathbf{x}) 表示 \mathbf{x} 的 k 近邻样本集。

定义 2^[12] 考虑第 c 类样本 $\mathbf{x}_i^{(c)}$,给定相反类中任意样本 $\mathbf{x}_l^{(c)}$ ($l=1, 2, \dots, m_c$),则类间近邻图 G_d 的相似矩阵 \mathbf{W}^d (\mathbf{W}_{il}^d) _{$m_1 \times m_2$} 可定义为

1 投影孪生支持向量机(PTSVM)

假定两类学习样本集分别表示为实数矩阵 $\mathbf{A} \in \mathbb{R}^{m_1 \times n}$ 和 $\mathbf{B} \in \mathbb{R}^{m_2 \times n}$ 。 n 为样本维度, m_1 和 m_2 分别为第 1 类(+1 类)和第 2 类(-1 类)样本数目,并且令 $m=m_1+m_2$ 。PTSVM 算法的优化目标可以看作是在实数空间中寻找两个最佳投影轴为 \mathbf{w}_1 和 \mathbf{w}_2 的决策超平面:

$$\mathbf{x}^T \mathbf{w}_1 + b_1 = 0, \mathbf{x}^T \mathbf{w}_2 + b_2 = 0.$$

(1)

需要注意的是,这里的偏置 $b_1 = -\mathbf{e}_1^T \mathbf{A} \mathbf{w}_1 / m_1$, $b_2 = -\mathbf{e}_2^T \mathbf{B} \mathbf{w}_2 / m_2$, \mathbf{e}_1 和 \mathbf{e}_2 是两个实体为 1 的列向量, $\mathbf{A} = [\mathbf{x}_1^{(1)} \ \mathbf{x}_2^{(1)} \ \cdots \ \mathbf{x}_{m_1}^{(1)}]^T$, $\mathbf{B} \mathbf{x} = [\mathbf{x}_1^{(2)} \ \mathbf{x}_2^{(2)} \ \cdots \ \mathbf{x}_{m_2}^{(2)}]^T$, $\mathbf{x}_j^{(i)}$ 表示第 i 类的第 j 个样本。

第 1 类超平面的优化准则 PTSVM-1:

$$\min \frac{1}{2} \sum_{i=1}^{m_1} \left(\mathbf{w}_1^T \mathbf{x}_i^{(1)} - \mathbf{w}_1^T \frac{1}{m_1} \sum_{j=1}^{m_1} \mathbf{x}_j^{(1)} \right)^2 + C_1 \sum_{l=1}^{m_2} \xi_l$$
$$\text{s.t.} \quad - \left(\mathbf{w}_1^T \mathbf{x}_l^{(2)} - \mathbf{w}_1^T \frac{1}{m_1} \sum_{j=1}^{m_1} \mathbf{x}_j^{(1)} \right) + \xi_l \geq 1, \xi_l \geq 0,$$

(2)

式中 C_1 是惩罚参数, ξ_l 为损失变量。

显然,PTSVM 在优化目标函数中考虑的训练样本集内在的散度 $\sum_{i=1}^{m_1} \left(\mathbf{w}_1^T \mathbf{x}_i^{(1)} - \mathbf{w}_1^T \frac{1}{m_1} \sum_{j=1}^{m_1} \mathbf{x}_j^{(1)} \right)^2$,体现的是样本集内在的全局分布。故该方法忽视了潜在在训练样本集内部的局部几何结构。

2 加权投影孪生支持向量机(WPTSVM)

2.1 算法构造

为刻画同类样本集内在的紧凑型和异类样本集间的分散性,依据图论^[10,12] 为每类决策面构建类内近邻图 G_s 和类间近邻图 G_d 。

定义 1 给定第 c 类中的任意两个样本 $\mathbf{x}_i^{(c)}$ 和 $\mathbf{x}_j^{(c)}$, ($c=1,2;i,j=1,2, \dots, m_c$), 则类内近邻图 G_s 的相似矩阵 \mathbf{W}^s (\mathbf{W}_{ij}^s) _{$m_1 \times m_1$} 可定义为

$$W_{il}^d = \begin{cases} 1, & \mathbf{x}_l^{(c)} \text{ is } k \text{ nearest neighbors of } \mathbf{x}_i^{(c)} \\ 0, & \text{其他} \end{cases}$$

(4)

依据定义 2,第 c 类中每一个样本定义权重:

$$f_l^{(c)} = \begin{cases} 1, & \exists i, w_{il}^{(c)} \neq 0 \\ 0, & \text{其他} \end{cases} \quad (5)$$

显然,权重 $f_l^{(c)} = 1$ 样本是第 c 类样本集的边界样本。

定义 3 WPTSVM 算法中,第 1 类训练样本集相应决策超平面的优化准则 WPTSVM-1:

$$\begin{aligned} \min & \frac{1}{2} \sum_{i=1}^{m_1} \rho_i^{(1)} \left(\mathbf{w}_1^T \mathbf{x}_i^{(1)} - \mathbf{w}_1^T \sum_{j=1}^{m_1} \lambda_j^{(1)} \mathbf{x}_j^{(1)} \right)^2 + C_1 \sum_{l=1}^{m_2} \xi_l \\ \text{s.t.} & -f_l^{(2)} \left(\mathbf{w}_1^T \mathbf{x}_l^{(2)} - \mathbf{w}_1^T \sum_{j=1}^{m_1} \lambda_j^{(1)} \mathbf{x}_j^{(1)} \right) + \xi_l \geq f_l^{(2)}, \xi_l \geq 0 \end{aligned} \quad (6)$$

第 2 类超平面优化准则 WPTSVM-2:

$$\begin{aligned} \min & \frac{1}{2} \sum_{l=1}^{m_2} \rho_l^{(2)} \left(\mathbf{w}_2^T \mathbf{x}_l^{(2)} - \mathbf{w}_2^T \sum_{j=1}^{m_2} \lambda_j^{(2)} \mathbf{x}_j^{(2)} \right)^2 + C_2 \sum_{i=1}^{m_1} \eta_i, \\ \text{s.t.} & f_i^{(1)} \left(\mathbf{w}_2^T \mathbf{x}_i^{(1)} - \mathbf{w}_2^T \sum_{j=1}^{m_2} \lambda_j^{(2)} \mathbf{x}_j^{(2)} \right) + \eta_i \geq f_i^{(1)}, \eta_i \geq 0 \end{aligned} \quad (7)$$

式中: C_1 和 C_2 是惩罚参数, ξ_l 和 η_i 为损失变量,

$$\rho_i^{(c)} = \frac{m_c}{\sum_{j=1}^{m_c} W_{ij}^s}, \lambda_j^{(c)} = \rho_j^{(c)} / \sum_{i=1}^{m_c} \rho_i^{(c)}, c = 1, 2.$$

式(6)中, $\rho_i^{(1)}$ 代表样本 $\mathbf{x}_i^{(1)}$ 的权重, $\rho_i^{(1)}$ 值越大,表示 $\mathbf{x}_i^{(1)}$ 越重要,对保持样本空间局部信息的贡献程度越大; $\sum_{j=1}^{m_1} \lambda_j^{(1)} \mathbf{x}_j^{(1)}$ 为第 1 类样本空间的加权均值,比起式(2)中的标准均值更能体现样本空间的局部结构^[13];约束条件表明 WPTSVM-1 仅仅考虑第 2 类样本中 $f_l^{(2)} = 1$ 的边界样本。显然,式(6)优化目标是为第 1 类样本寻找最佳投影轴 $\mathbf{w}_1 \mathbf{x}_1$,使得权重较大的样本投影后尽可能聚集在加权均值中心附近,而第 2 类中 $f_l^{(1)} = 1$ 的边界样本离中心尽可能远。式(7)有类似的几何解释式。式(6)矩阵形式为

$$\begin{aligned} \min & \frac{1}{2} (\mathbf{A} \mathbf{w}_1 - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A} \mathbf{w}_1)^T \\ & \mathbf{D}^{(1)} (\mathbf{A} \mathbf{w}_1 - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A} \mathbf{w}_1) + C_1 \mathbf{e}_2^T \boldsymbol{\xi} \\ \text{s.t.} & -\mathbf{F}^{(2)} (\mathbf{B} \mathbf{w}_1 - \mathbf{e}_2 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A} \mathbf{w}_1) + \boldsymbol{\xi} \geq \mathbf{F}^{(2)} \mathbf{e}_2, \boldsymbol{\xi} \geq 0 \end{aligned} \quad (8)$$

式中: $\boldsymbol{\xi}_2 = (\xi_1, \dots, \xi_{m_2})^T$, $\mathbf{D}^{(1)} = \text{diag}(\rho_1^{(1)}, \dots, \rho_{m_1}^{(1)})$, $\mathbf{E}^{(1)} = \text{diag}(\lambda_1^{(1)}, \dots, \lambda_{m_1}^{(1)})$, $\mathbf{F}^{(2)} = \text{diag}(f_1^{(2)}, \dots, f_{m_2}^{(2)})$ 。

类似于传统 SVM 求解方法,通过引入拉格朗日函数生成式(8)的对偶问题

$$\begin{aligned} \min & \frac{1}{2} \boldsymbol{\alpha}^T (\mathbf{F}^{(2)} (\mathbf{B} - \mathbf{e}_2 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})) \times \\ & ((\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})^T \mathbf{D}^{(1)} (\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A}))^{-1} \\ & (\mathbf{F}^{(2)} (\mathbf{B} - \mathbf{e}_2 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A}))^T \boldsymbol{\alpha} - \mathbf{e}_2^T \boldsymbol{\alpha} \\ \text{s.t.} & 0 \leq \boldsymbol{\alpha} \leq C_1 \mathbf{e}_2 \end{aligned} \quad (9)$$

式中 $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_{m_2}]^T$ 。

由此对偶问题的求解可得原问题式(6)中投影轴 \mathbf{w}_1 。类似的方法,可求原问题式(7)中投影轴 \mathbf{w}_2 。

对于未知样本 \mathbf{x} , WPTSVM 的决策函数为

$$\text{label}(\mathbf{x}) = \underset{c \in \{1, 2\}}{\text{argmin}} \{d_c\} = \begin{cases} d_1 \Rightarrow \mathbf{x} \in \text{第 1 类} \\ d_2 \Rightarrow \mathbf{x} \in \text{第 2 类} \end{cases} \quad (10)$$

式中: $d_c = |\mathbf{w}_c^T \mathbf{x} - \mathbf{w}_c^T \sum_{j=1}^{m_c} \lambda_j^{(c)} \mathbf{x}_j^{(c)}|$, $|\cdot|$ 表示绝对值。

2.2 算法分析

这里以 WPTSVM 的第 1 类样本优化问题 WPTSVM-1 为例进行算法分析,第 2 类样本优化问题 WPTSVM-2 有类似的算法分析。

考虑对偶形式式(9),假设实对称矩阵 $((\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})^T \mathbf{D}^{(1)} (\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A}))$ 可逆(若不可逆,可采用类似文献[3]方法引入正则化项 $\mathbf{I}(\epsilon > 0)$, ϵ 尽可能的小)。式(14)中的核函数 $\mathbf{K}(\cdot)$ 只有保证 Mercer 核时,才能保证其是二次凸规划,所求的解才为全局最优解。为验证这一问题,给出如下定理。

定理 1 式(9)为凸二次规划。

证明 式(9)核矩阵:

$$\begin{aligned} \tilde{\mathbf{K}} &= [\tilde{k}_{ij} = (\mathbf{F}^{(2)} (\mathbf{B} - \mathbf{e}_2 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})) \cdot \\ & ((\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})^T \mathbf{D}^{(1)} (\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A}))^{-1} \cdot \\ & (\mathbf{F}^{(2)} (\mathbf{B} - \mathbf{e}_2 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A}))^T \end{aligned}$$

是实对称矩阵。下面证明矩阵 $\tilde{\mathbf{K}}$ 是半正定的。设任意 $\mathbf{x} \in \mathbf{R}^{m_2}$, 则有 $\mathbf{x}^T \tilde{\mathbf{K}} \mathbf{x} = (\mathbf{x}^T (\mathbf{F}^{(2)} (\mathbf{B} - \mathbf{e}_2 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})) ((\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A})^T \mathbf{D}^{(1)} (\mathbf{A} - \mathbf{e}_1 \mathbf{e}_1^T \mathbf{E}^{(1)} \mathbf{A}))^{-1/2})^2 \geq 0$, 因此 $\tilde{\mathbf{K}}$ 半正定且为 Mercer 核函数。根据文献[14]中引理 2 可知,式(9)为凸二次规划。证毕。

定理 2 假定 $\boldsymbol{\alpha}$ 是对偶问题式(9)的解,则 \mathbf{w}_1 是原优化问题式(6)的全局最优解。

证明 由定理 1 的证明可得,式(9)为凸二次规划问题,又依据文献[14]中引理 3 的满足条件可知,该二次规划的解为全局最优解。证毕。

2.3 算法比较

2.3.1 泛化性能

定理 3 PTSVM 是 WPTSVM 的特例。

证明 考虑 WPTSVM 的第 1 类样本的优化准则(7)。令 $\mathbf{D}^{(1)} = \mathbf{I} \in \mathbf{R}^{m_1 \times m_1}$, $\mathbf{F}^{(2)} = \mathbf{I} \in \mathbf{R}^{m_2 \times m_2}$, 则式(7)转化为 PTSVM 的第 1 类样本的优化准则(2)。对于 WPTSVM 的第 2 类样本的优化准则(8)有类似的特性。因此, PTSVM 是 WPTSVM 的特例,而

WPTSVM 是 PTSVM 的推广算法。证毕。

定理 3 说明了本文提出的 WPTSVM 算法继承了 PTSVM 的优点。进一步比较式 (7) 和 (2) 可知, PTSVM 仅仅考虑类内样本的全局信息,而 WPTSVM 用加权均值代替 PTSVM 中标准均值,可以在一定程度上提高算法的局部学习能力,因为基于近邻图的加权均值比起标准均值更能体现样本空间的局部结构^[13]。除此之外,WPTSVM 还在优化目标函数中引入了样本权值,权值越大,说明该样本越重要,对保持训练样本集潜在局部信息的贡献程度越大。图 1 给出了 WPTSVM 与 PTSVM 在人造数据集上的分类决策面。不难看出,WPTSVM 的决策面能够在一定程度上体现样本集内在局部流形结构,而 PTSVM 相对较弱。

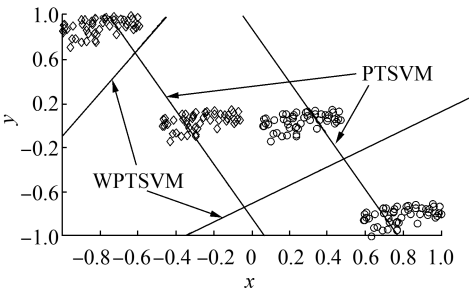


图 1 两种算法人造数据集上的比较

Fig.1 Comparison of two algorithms on artificial dataset

2.3.2 训练时间复杂度

从二次规划求解角度分析,PTSVM 在训练阶段要针对每类中全部样本进行求解,所以计算复杂度为 $o(m_1^3+m_2^3)$,而 WPTSVM 优化准则中约束条件指明只对 $f_i^{(c)}=1$ 的样本(边界样本)进行二次规划求解,计算复杂度为 $o(m_{1-SV}^3+m_{2-SV}^3)$,其中 m_{1-SV},m_{2-SV} 分别为第 1 类样本及第 2 类样本中相应边界样本点数。诚然,WPTSVM 在训练阶段要事先求出每个样本的类内权重及类间权重,计算复杂度分别为 $o(m_1^2+m_2^2)$ 和 $o(2m_1 \cdot m_2)$ 。

2.3.3 构造非线性分类算法

针对线性不可分情况,本文进一步提出基于核空间的非线性 WPTSVM 算法——NWPTSVM。

定义 4 NWPTSVM 的第 1 类超平面的优化准则为

$$\begin{aligned} \min \quad & \frac{1}{2} (K(A,C^T)u_1 - e_1 e_1^T E^{(1)} K(A,C^T)u_1)^T D^{(1)} \cdot \\ & (K(A,C^T)u_1 - e_1 e_1^T E^{(1)} K(A,C^T)u_1) + C_1 e_1^T \xi, \\ \text{s.t.} \quad & -F^{(2)} (K(B,C^T)u_1 - e_2 e_1^T E^{(1)} K(A,C^T)u_1) + \\ & \xi \geq F^{(2)} e_2, \xi \geq 0 \end{aligned} \tag{11}$$

第 2 类超平面的优化准则为

$$\begin{aligned} \min \quad & \frac{1}{2} (K(B,C^T)u_2 - e_2 e_2^T E^{(2)} K(B,C^T)u_2)^T D^{(2)} \cdot \\ & (K(B,C^T)u_2 - e_2 e_2^T E^{(2)} K(B,C^T)u_2) + C_2 e_1^T \eta, \\ \text{s.t.} \quad & F^{(1)} (K(A,C^T)u_2 - e_1 e_2^T E^{(2)} K(B,C^T)u_2) + \\ & \eta \geq F^{(1)} e_1, \eta \geq 0 \end{aligned} \tag{12}$$

通过引入拉格朗日函数,可按照类似上述 WPTSVM 算法的推导过程分别得出式 (11) 和 (12) 的对偶形式,然后通过二次规划求解可求得投影矢量 u_1 和 u_2 。NWPTSVM 的决策函数为

$$\text{label}(x) = \underset{c \in 1,2}{\operatorname{argmin}} \{d_c\} = \begin{cases} d_1 \Rightarrow x \in \text{第 1 类} \\ d_2 \Rightarrow x \in \text{第 2 类} \end{cases} \tag{13}$$

式中 $d_c = \left| K(x^T,C^T)u_i - \sum_{j=1}^{m_c} \lambda_j^{(c)} K((x_j^{(c)})^T,C^T)u_i \right|$ 。

3 实验分析

实验选用人造数据集和真实数据集,进一步验证本文 WPTSVM 方法的有效性。实验环境: 2.3 GHz CPU,2 GB 内存,实验软件 MATLAB 7.1。

3.1 复杂交叉数据集

相对于经典 SVM,线性模式下对 XOR 问题的测试能力是非平行超平面分离器优势之一^[3-5]。因此,本节首先验证 MPTSVM 测试交叉数据集的能力。图 2 给出一种较复杂的人造交叉数据集:Comxor^[7]。表 1 给出了 TWSVM、PTSVM 和 WPTSVM 三种算法在该测试数据集上 10 折交叉验证结果。参数 C_1 与 C_2 的搜索范围为 $\{2^i | i=-8,-6,\cdots,+8\}$; WPTSVM 中类内近邻参数 k_1 的搜索范围为 $\{1,2,\cdots,9\}$,类间近邻参数 $k_2=5$,热核参数 t 的搜索范围为 $\{2^i | i=-1,0,\cdots,8\}$ 。从表 1 实验结果来看,PTSVM 分类性能优于 TWSVM,而本文 WPTSVM 则具有更佳分类性能。

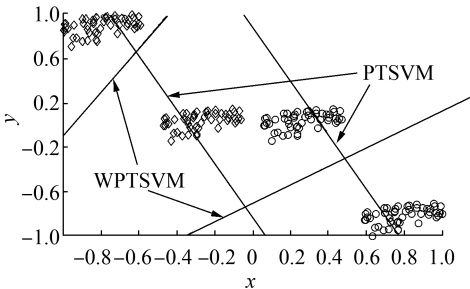


图 2 复杂交叉数据集
Fig.2 Compxor dataset

表 1 3 种算法在复杂交叉数据集上的分类精度 (%) 比较

Table 1 Classification comparison of three algorithms on comxor dataset

Dataset	TWSVM	PTSVM	WPTSVM
Comxor	93.76±6.70	95.67±4.73	97.36±2.90

3.2 流形数据集

数据集 two-moons (如图 3 所示) 的结构具有明显局部流形,所以该数据集多用于测试算法的局部学习能力^[2,13]。这里通过测试 two-moons 数据集,并与 TWSVM 和 PTSVM 方法进行比较,验证本文 WPTSVM 方法局部学习能力。

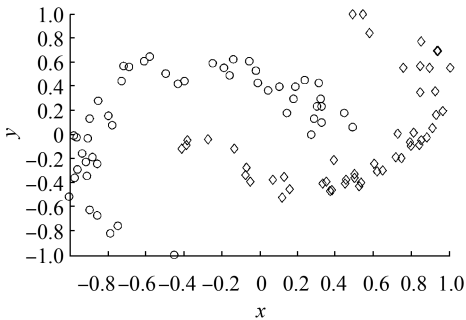


图 3 Two-moons 数据集

Fig.3 Two-moons dataset

实验选择正负类样本各为 50 的 two-moons 数据集。随机抽取 40% 训练集和 60% 测试集进行实验,重复此过程 10 次并记录实验结果的平均值

表 3 线性模式下 3 种算法的实验比较

Table 3 Experimental comparison of three algorithms on UCI datasets with linear kernel

Dataset	TWSVM		PTSVM		WPTSVM	
	正确率/%	训练时间/s	正确率/%	训练时间/s	正确率/%	训练时间/s
Hepatitis (155×19)	84.17±7.79	3.84	83.67±10.80	3.46	87.33±9.17	1.36
Cleve (296×13)	82.42±4.21	5.40	82.08±3.38	5.26	83.46±3.71	2.56
Sonar (208×60)	66.00±16.85	3.93	66.36±14.27	3.82	68.07±13.60	1.58
Bupa_liver(345×6)	68.69±3.23	2.42	67.59±5.30	2.84	70.20±6.42	2.17
Wdbc (569×30)	96.99±1.63	19.77	96.99±1.63	19.95	97.53±1.65	1.14
Haberman (306×3)	74.50±2.11	4.09	74.56±7.90	5.74	75.44±5.03	1.36
Heart (270×13)	84.07±5.25	5.30	84.07±5.51	5.38	85.19±5.49	2.40
Monks2 (432×6)	67.14±4.98	2.84	67.84±3.15	2.56	67.13±3.20	0.78
Monks3 (432×6)	77.41±11.32	6.07	80.74±13.05	6.32	81.69±12.85	3.76
Australian(690×14)	86.26±4.76	13.37	86.28±4.91	11.90	86.83±5.25	2.81
Pima_india(768×8)	77.17±2.80	40.91	76.41±3.53	42.81	76.79±4.05	10.25
Tic_tac_toe(958×9)	64.93±2.31	18.07	63.11±5.33	18.21	78.72±6.13	6.24

(见表 2)。显然,WPTSVM 方法具有更好的测试效果,这说明本文加权措施的确能够在一定程度上提高 PTSVM 算法的局部学习能力。

表 2 3 种算法在双月形数据集上的分类精度 (%) 比较

Table 2 Classification comparison of three algorithms on two-moons dataset

Dataset	TWSVM	PTSVM	WPTSVM
two-moons	77.17±13.72	76.33±12.36	78.83±8.86

3.3 UCI 数据集

UCI 数据集经常被用来验证算法的分类精度^[3-7,15-16]。该数据集包含若干数据子集,涉及诸多应用领域。

实验中选取部分数据子集,利用 10-折交叉验证法测试算法分类性能。非线性算法实验中,选择高斯核 $\exp(-\|\mathbf{x}_i-\mathbf{x}_j\|^2/\sigma)$ 作为核函数,参数 σ 的搜索范围为 $\{2^i|i=-1,0,\cdots,7\}$,其他参数搜索范围同 3.1 节。线性模式及非线性模式下实验结果如表 3 和表 4 分别所示。

从训练时间上看:TWSVM 与 PTSVM 相当,WPTSVM 明显比这两种算法快。原因是 WPTSVM 选用少量边界样本进行二次规划求解,而其他两种算法则选择异类中全部样本。

从分类性能上看:本文提出的 WPTSVM 算法具有更好的分类能力,这也进一步佐证了本文提高 PTSVM 算法局部学习能力的措施确实有效。

表 4 非线性模式下 3 种算法的实验比较

Table 4 Experimental comparison of three algorithms on UCI datasets with nonlinear kernel

Dataset	TWSVM		PTSVM		WPTSVM	
	正确率/%	训练时间/s	正确率/%	训练时间/s	正确率/%	训练时间/s
Spectf (267×46)	83.43±7.15	14.63	82.49±8.36	14.65	84.03±5.17	1.73
Cleve (296×13)	85.12±5.77	7.22	85.93±4.90	7.82	85.24±4.56	3.40
Wpbc (198×33)	79.36±5.52	4.12	78.03±6.59	1.72	79.98±6.62	0.59
P_gene (106×57)	65.50±18.23	2.25	64.13±21.17	2.22	68.13±16.71	1.19
Sonar (208×60)	63.29±18.65	7.55	64.50±13.50	8.63	68.43±13.26	2.78
Spect (267×22)	84.42±9.04	11.42	84.42±7.43	4.01	84.80±8.27	5.54
Monks2 (432×6)	68.77±15.24	5.02	68.71±3.43	6.43	68.99±11.07	4.57
Vertebral (310×6)	85.81±6.32	5.63	84.52±6.42	6.02	86.13±6.46	2.87
Monks3 (432×6)	98.70±3.89	11.09	98.52±4.44	9.97	99.26±2.22	8.03
Breast_gy (277×9)	75.86±5.89	5.24	73.57±4.10	4.04	75.94±6.69	0.87

4 结束语

本文基于投影孪生支持向量机(PTSVM)提出一种新的非平行超平面分类器方法:加权投影孪生支持向量机(WPTSVM)。WPTSVM 不仅继承了 PTSVM 方法的优点,而且在一定程度上提高了算法局部学习能力。除此之外,WPTSVM 通过类间近邻图选择少量边界样本构造优化问题约束条件,相当程度上降低了二次规划求解时间复杂度。理论分析及实验结果均验证了本文所提算法的有效性。诚然,WPTSVM 在构造类内及类间近邻图时,需要花费额外的计算开销,特别是在学习样本数目较大时,算法计算复杂度会偏高,这也是今后进一步研究的目标。

参考文献:

[1] 皋军, 王士同, 邓赵红. 基于全局和局部保持的半监督支持向量机[J]. 电子学报, 2010, 38(7): 1626-1633.

GAO Jun, WANG Shitong, DENG Zhaohong. Global and local preserving based semi-supervised support vector machine [J]. Acta electronica sinica, 2010, 38(7): 1626-1633.

[2] 花小朋, 丁世飞. 局部保持对支持向量机[J]. 计算机研究与发展, 2014, 51(3): 590-597.

HUAXIAOPENG, DING Shifei. Locality preserving twin support vector machines [J]. Journal of computer research and development, 2014, 51(3): 590-597.

[3] DING Shifei, HUA Xiaopeng, YU Junzhao. An overview on nonparallel hyperplane support vector machines[J]. Neural computing and applications, 2014, 25(5): 975-982.

[4] JAYADEVA, KHEMCHAND R, CHANDRA S. Twin support vector machines for pattern classification [J]. IEEE transaction on pattern analysis and machine intelligence, 2007, 29 (5): 905-910.

[5] MANGASARIAN O L, WILD E W. MultisurFace proximal support vector machine classification via generalized eigenvalues [J]. IEEE transactions on pattern analysis and machine intelligence, 2006, 28 (1): 69-74.

[6] PENG Xinjun, XU Dong. Bi-density twin support vector machines for pattern recognition [J]. Neurocomputing, 2013, 99: 134-143.

[7] CHEN Xiaobo, YANG Jian, YE Qiaolin, et al. Recursive projection twin support vector machine via within-class variance minimization[J]. Pattern recognition, 2011, 44 (10/11): 2643-2655.

[8] SHAO Yuanhai, WANG Zhen, CHEN Weijie, et al. A regularization for the projection twin support vector machine [J]. Knowledge-based systems, 2013, 37 (1): 203-210.

[9] YANG Xubing, CHEN Songcan, CHEN Bin, et al. Proximal support vector machine using local information [J]. Neurocomputing, 2009, 73(1): 357-365.

[10] COVER T M, HART P E. Nearest neighbor pattern classification [J]. IEEE transactions on information theory, 1967, 13 (1): 21-27.

[11] WANG Xiaoming, CHUNG Fulai, WANG Shitong. On minimum class locality preserving variance support vector machine[J]. Patter recognition, 2010, 43 (8): 2753-2762.

[12] YE Qiaolin, ZHAO Chunxia, YE Ning, et al. Localized twin svm via convex minimization [J]. Neurocomputing, 2011, 74(4): 580-587.

[13] 皋军, 黄丽莉, 王士同. 基于局部子域的最大间距判别分析 [J]. 控制与决策, 2014, 29 (5): 827-832.

GAO Jun, HUANG Lili, WANG Shitong. Local sub-domains based maximum margin criterion [J]. Control and decision, 2014, 29 (5): 827-832.

[14] 邓乃杨, 田英杰. 支持向量机—理论、算法与拓展[M]. 北京: 科学出版社, 2009: 164-223.

[15] 丁立中, 廖士中. KMA-a: 一个支持向量机核矩阵的近

似计算算法[J]. 计算机研究与发展, 2012, 49(4): 746-753.

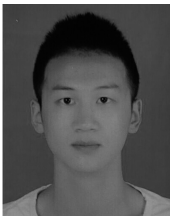
DING Lizhong, LIAO Shizhong. KMA-a: a kernel matrix approximation algorithm for support vector machines [J]. Journal of computer research and development, 2012, 49(4): 746-753.

[16] XUE Hui, CHEN Songchan. Glocalization pursuit support vector machine [J]. Neural computing and applications, 2011, 20(7): 1043-1053.

作者简介:



花小鹏,男,1975 年生,副教授,博士,主要研究方向为机器学习与数据挖掘,发表学术论文 10 余篇。



孙一颗,男,1993 年生,硕士研究生,主要研究方向为机器学习、数据挖掘,申请发明专利 2 项。



丁世飞,男,1963 年生,教授,博士生导师,中国计算机学会高级会员,中国人工智能学会高级会员,江苏省计算机学会人工智能专业委员会委员,主要研究方向为智能信息处理,目前主持国家 973 项目 1 项、国家自然科学基金项目 1 项,发表学术论文 60 余篇。

2016 年第 9 届 IFAC 智能自主车辆研讨会

2016 9th IFAC Symposium on Intelligent Autonomous Vehicles (IAV)

The Symposium targets topics in the field of intelligent autonomous vehicles (IAV). It presents methods and representative techniques to either solve typical problems or to document successful test cases and applications. The addressed modality of the vehicles may range from land, sea and air to space and they may be designed as multi vehicle systems in swarms or networked platoons as well as individual carriers.

The covered key topics will include (but will not be limited to):

- 1) Architectures for autonomous vehicles and teams of IAVs;
- 2) Perception including sensing, sensor integration, smart sensors and sensor networks;
- 3) Recognition;
- 4) Decision support systems, driver support systems, Advanced driver assistance systems;
- 5) Planning and mission control;
- 6) Local and global roadmap planning, trajectory tracking and path following;
- 7) Localization including SLAM and map building;
- 8) Navigation, guidance and control, motion control, controller design, stability analysis;
- 9) Human vehicle interaction;
- 10) Transportation systems;
- 11) Remote control, monitoring and teleoperation;
- 12) Fault detection, diagnosis and removal in IAVs and systems;
- 13) Domestic robots, service and rehabilitation robots;
- 14) Field robotics, maritime and ocean robots;
- 15) Networked IAVs;
- 16) Swarm robotics;
- 17) Communication and cooperation for teams for IAVs.

Website: <http://iav2016.inf.h-brs.de/>