

文章编号: 2095-2163(2020)02-0120-04

中图分类号: TP181

文献标志码: A

# 加权多任务最小二乘双支持向量机

阮丽<sup>1</sup>, 黄成泉<sup>2</sup>, 朱文文<sup>1</sup>

(1 贵州民族大学 数据科学与信息工程学院, 贵阳 550025; 2 贵州民族大学 工程技术人才实践训练中心, 贵阳 550025)

**摘要:** 通过对传统多任务支持向量机中松弛约束项的分析发现, 传统多任务支持向量机的松弛约束项存在局限性。为此, 本文给松弛约束项增加一个权重约束, 提出了加权多任务最小二乘双支持向量机, 实验验证表明, 加权多任务最小二乘双支持向量机通过权重参数增加了松弛项的松弛范围, 有效提高了分类效果。

**关键词:** 支持向量机; 加权; 权重约束

## Weight mutli-task least squares twin support vector machine

RUAN Li<sup>1</sup>, HUANG Chengquan<sup>2</sup>, ZHU Wenwen<sup>1</sup>

(1 School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, China;

2 Engineering Training Center, Guizhou Minzu University, Guiyang 550025, China)

**[Abstract]** Based on relaxation constraint analysis of the traditional multi-task support vector machine, it is found that the traditional relaxation constraint of multi-task support vector machine have limitations. Therefore, this paper adds a weight to relaxation constraint item, and proposes the weighted multi-task least square support vector machine. The experimental verification shows that the weighted least squares multi-task double support vector machine by weighting parameters increases relaxation range, effectively improves the classification effect.

**[Key words]** support vector machine; weighting; weight constraint

### 0 引言

机器学习中, 统计学习理论在解决小样本和非线性的问题上有着出色表现, 其中作为典型代表的支持向量机(Support Vector Machine, SVM)<sup>[1-2]</sup>则因为所具备的优秀的性能, 现已广泛地应用在各个领域中。但是, 单任务支持向量机在训练样本小、信息量不足和多个数据差异的情况下的性能表现上却仍有一定欠缺。为此, 在多任务学习(Mutli-task Learning)<sup>[3]</sup>的启发下, 支持向量机则被成功应用到多任务学习上。研究可知, 多任务支持向量机(Mutli-task Support Vector Mahicne, MTL SVM)通过共享数据之间信息来提高分类效果, 解决了如上所述单任务向量机存在的问题。如今, MTL SVM 已经得到了学界的普遍关注和重视。早期的 MTL SVM 是研究单类分类的。Yang 等人<sup>[4]</sup>在 2010 年提出了多任务学习一类分类, 为 MTL SVM 的研究提供了参考。He 等人<sup>[5]</sup>在多任务学习一类分类的基础上提出了多任务-类支持向量机(Multi-task one-class support vector machines, MTOC-SVM), Xue 等人<sup>[6]</sup>在 MTOC-SVM 的基础上增加新特征, 提出了支持

向量机的多任务学习新特征。由于求解二次规划问题计算复杂度高, 时间成本大, 为此 Xu 等人<sup>[7]</sup>提出了多任务最小二乘支持向量机(Multi-task least squares support vector machine, MTLSSVM), Li 等人<sup>[8]</sup>根据近端支持向量机<sup>[9]</sup>(Proximal support vector machine, PSVM)提出了多任务近端支持向量机(Multi-task proximal support vector machine, MTP SVM)。这 2 个模型都降低了计算成本。同样地, 由于多任务双支持向量机<sup>[10]</sup>(Multi-task twin support vector machine, DM TSVM)也是一个求解二次规划的问题, 其复杂性和计算量都较为可观。因此, Mei 等人<sup>[11]</sup>提出了多任务最小二乘双支持向量机(Multi-task least squares twin support vector machine, MTL ST SVM), 能有效提高计算速度。综上所述研究后发现, 在这些算法中, 松弛约束项有较大的局限性, 为此, 本文在传统的 MTL SVM 的约束上增加一个权重约束, 提出加权多任务最小二乘双支持向量机(Weight mutli-task least squares twin support vector machine, WMTL ST SVM)。实验结果表明, 本文算法在分类上具有良好性能。

**作者简介:** 阮丽(1994-), 女, 硕士研究生, 主要研究方向: 统计信号处理、计算机视觉、机器学习; 黄成泉(1976-), 男, 博士, 教授, 主要研究方向: 计算机软件与理论、计算机应用技术、嵌入式系统等; 朱文文(1993-), 女, 硕士研究生, 主要研究方向: 统计信号处理、计算机视觉、机器学习。

收稿日期: 2019-11-25

### 1 理论基础

多任务最小二乘双支持向量机(MTLSTSVM)是求解一对线性方程组问题的算法,这里,给出MTLSTSVM的基本理论,MTLSTSVM为本文的算法提供了理论依据。

假设一个二分类任务,  $X_1 \subset R^{N_1 \times d}, X_2 \subset R^{N_2 \times d}$  代表类1和类-1。其中,  $X_1, X_2$  的每一行对应一个

$$\min \frac{1}{2} \|X_1 w_1 + e_1 b_1\|^2 + \frac{\rho}{2} \sum_{i=1}^T \|X_{1i} w_{1i} + e_{1i} b_{1i}\|^2 + \frac{c_1}{2} \sum_{i=1}^T \xi_i^T \xi_i$$

$$\text{s.t. } - [[X_{2i}, e_{2i}](u + u_i)] + \xi_i = e_{2i}, \xi_i \geq 0, \tag{1}$$

$$\min \frac{1}{2} \|X_2 w_2 + e_2 b_2\|^2 + \frac{\lambda}{2} \sum_{i=1}^T \|X_{2i} w_{2i} + e_{2i} b_{2i}\|^2 + \frac{c_2}{2} \sum_{i=1}^T \eta_i^T \eta_i$$

$$\text{s.t. } [[X_{1i}, e_{1i}](v + v_i)] + \eta_i = e_{1i}, \eta_i \geq 0. \tag{2}$$

其中,  $e_1, e_2, e_{1i}, e_{2i}$  表示适当维数的列向量;  $\xi_i$  和  $\eta_i$  表示松弛向量;  $c_1, c_2, \rho, \lambda$  表示非负交换参数。

### 2 加权多任务最小二乘双支持向量机

#### 2.1 线性加权多任务最小二乘双支持向量机

考虑到 MTLSTSVM 的松弛约束项有较大的局限性,所以,本文在 MTLSTSVM 的约束上增加一个权重约束,提出了加权多任务最小二乘双支持向量机。现给出加权多任务最小二乘双支持向量机算法的优化函数如式(3)、(4)所示:

$$\min \frac{1}{2} \|X_1 w_1 + e_1 b_1\|^2 + \frac{\rho}{2T} \sum_{i=1}^T \|X_{1i} w_{1i} + e_{1i} b_{1i}\|^2 + \frac{c_1}{2} \sum_{i=1}^T \xi_i^T W \xi_i$$

$$\text{s.t. } - [[X_{2i}, e_{2i}](u + u_i)] + \xi_i = e_{2i}, \xi_i \geq 0, \tag{3}$$

$$\min \frac{1}{2} \|X_2 w_2 + e_2 b_2\|^2 + \frac{\lambda}{2T} \sum_{i=1}^T \|X_{2i} w_{2i} + e_{2i} b_{2i}\|^2 + \frac{c_2}{2} \sum_{i=1}^T \eta_i^T W \eta_i$$

$$\text{s.t. } [[X_{1i}, e_{1i}](v + v_i)] + \eta_i = e_{1i}, \eta_i \geq 0. \tag{4}$$

其中,  $e_1, e_2, e_{1i}, e_{2i}$  表示适当维数的列向量;  $\xi_i$  和  $\eta_i$  表示松弛向量;  $W$  表示权重参数;  $c_1, c_2, \rho, \lambda$  表示非负交换参数。

先给出算法求解过程,首先引入拉格朗日乘子,将约束条件代入算法。则可以得到式(3)的拉格朗日函数如式(5)所示:

$$L_1 = \frac{1}{2} \|X_1 w_1 + e_1 b_1\|^2 + \frac{\rho}{2T} \sum_{i=1}^T \|X_{1i} w_{1i} + e_{1i} b_{1i}\|^2 + \frac{c_1}{2} \sum_{i=1}^T \xi_i^T W \xi_i - \sum_{i=1}^T \alpha_i^T [- [[X_{2i}, e_{2i}](u + u_i)] + \xi_i - e_{2i}],$$

数据样本。  $X_{1t}$  表示第  $t$  个任务的正类样本,  $X_{2t}$  表示第  $t$  个任务的负类样本。正负超平面分别是:  $u = [W_1, b_1]^T, v = [W_2, b_2]^T$ , 第  $t$  个任务的正负超平面是:  $[W_{1t}, b_{1t}]^T = (u + u_t), [W_{2t}, b_{2t}]^T = (v + v_t)$ 。  $u_t$  和  $v_t$  为  $u$  和  $v$  与第  $t$  个任务的偏差。 MTLSTSVM 的目标函数如式(1)、(2)所示:

计算式(5)的 KKT 条件:

$$\begin{cases} \frac{\partial L}{\partial w_1} = X_1^T(X_1 w_1 + e_1 b_1) + X_2^T \alpha = 0; \\ \frac{\partial L}{\partial b_1} = e_1^T(X_1 w_1 + e_1 b_1) + e_2^T \alpha = 0; \\ \frac{\partial L}{\partial w_{1t}} = \frac{\rho}{T} X_{1t}^T(X_{1t} w_{1t} + e_{1t} b_{1t}) + X_{2t}^T \alpha_t = 0; \\ \frac{\partial L}{\partial b_{1t}} = \frac{\rho}{T} e_{1t}^T(X_{1t} w_{1t} + e_{1t} b_{1t}) + e_{2t}^T \alpha_t = 0; \\ \frac{\partial L}{\partial \xi_t} = c_1 W \xi_t - \alpha_t = 0. \end{cases} \tag{6}$$

解式(6)可得:

$$[X_1, e_1]^T [X_1, e_1] [w_1, b_1]^T + [X_2, e_2]^T \alpha = 0, \tag{7}$$

令  $E = [X_1, e_1], F = [X_2, e_2]$ , 则有:

$$E^T E [w_1, b_1]^T + F^T \alpha = 0, \tag{8}$$

可得:

$$[w_1, b_1]^T = - (E^T E)^{-1} F^T \alpha, \tag{9}$$

同理可得:

$$[w_{1t}, b_{1t}]^T = - \frac{T}{\rho} (E_t^T E_t)^{-1} F_t^T \alpha_t, \tag{10}$$

$$\xi_t = \frac{\alpha_t}{c_1} W,$$

代回式(3)的约束项可得:

$$F (E^T E)^{-1} F^T \alpha + \frac{T}{\rho} F_t (E_t^T E_t)^{-1} F_t^T \alpha_t + \frac{a_t}{c_1} W = e_{2t}, \tag{11}$$

令  $A = F (E^T E)^{-1} F^T, B_t = F_t (E_t^T E_t)^{-1} F_t^T, B = blkdiag(B_1, B_2, \dots, B_t)$ , 代回式(11), 求解式(11)中的  $\alpha$  可以得到正超平面如式(12)所示:

$$\alpha = \left( A + \frac{T}{\rho} B + \frac{W}{c_1} \right)^{-1} e_2, \quad (12)$$

根据  $L_1$  的方法,可解  $\beta$ , 算法(5)的拉格朗日函数式如(13)所示:

$$L_2 = \frac{1}{2} \| X_2 w_2 + e_2 b_2 \|^2 + \frac{\lambda}{2T} \sum_{i=1}^T \| X_{2i} w_{2i} + e_{2i} b_{2i} \|^2 + \frac{c_2}{2} \sum_{i=1}^T \eta_i^T W \eta_i - \sum_{i=1}^T \beta_i^T [ [ [X_{1i}, e_{1i}] (v + v_i) ] + \eta_i - e_{1i} ], \quad (13)$$

求解  $L_2$  可以得到  $\beta$ , 即:

$$\beta = \left( C + \frac{T}{\lambda} D + \frac{W}{c_2} \right)^{-1} e_1, \quad (14)$$

这里,第  $t$  个任务的决策函数可根据式(15)得到:

$$f(x) = \arg \min_{i=1,2} |x^T w_{i} + b_{i}|. \quad (15)$$

## 2.2 非线性加权多任务最小二乘双支持向量机

对于加权多任务最小二乘双支持向量机非线性的情况,可通过内核函数来解决。核函数定义为:

$$M = (K(E, Z^T) e), \quad M_t = (K(E_t, Z^T) e_t), \\ N = (K(F, Z^T) e), \quad N_t = (K(F_t, Z^T) e_t),$$

这里,  $K(\cdot)$  为特定的一个核函数,  $Z^T = (E_1^T, \dots, E_t^T, F_1^T, \dots, F_t^T)$  为全部任务的训练样本。非线性的优化函数如式(16)、(17)所示:

$$\min \frac{1}{2} \| (K(E, Z^T) w_1 + e_1 b_1) \|^2 + \frac{\rho}{2T} \sum_{i=1}^T \| K(E_t, Z^T) w_{1i} + e_{1i} b_{1i} \|^2 + \frac{c_1}{2} \sum_{i=1}^T \xi_i^T W \xi_i \\ \text{s.t. } - [ [K(F_t, Z^T), e_{2i}] (u + u_i) ] + \xi_i = e_{2i}, \xi_i \geq 0, \quad (16)$$

$$\min \frac{1}{2} \| (K(F, Z^T) w_2 + e_2 b_2) \|^2 +$$

表2 3个数据集上的平均分类准确率结果

Tab. 2 Average classification accuracy results on 3 datasets

名称	LSTSVM		MTLSTSVM		WMTLSTSVM	
	准确率/%	时间/s	准确率/%	时间/s	准确率/%	时间/s
Monk	71.22 ± 3.13	1.49	88.23 ± 3.02	4.33	<b>88.62±3.07</b>	3.81
ASD	97.82 ± 1.71	1.56	98.18 ± 1.62	5.12	<b>98.26±1.52</b>	4.73
Dermatology	90.08 ± 12.62	1.35	95.44 ± 5.64	4.17	<b>97.76±2.39</b>	3.24

## 4 结束语

本文提出的加权多任务最小二乘双支持向量机,解决了传统多任务支持向量机松弛约束项局限大的问题,引入权重参数来约束松弛变量,得到了一个更好的分类效果,通过实验分析发现,本文的算法能有效地提高分类效果,减少了训练时间,这也证明了本文算法的有效性。

$$\frac{\lambda}{2T} \sum_{i=1}^T \| K(F_t, Z^T) w_{2i} + e_{2i} b_{2i} \|^2 + \frac{c_2}{2} \sum_{i=1}^T \eta_i^T W \eta_i \\ \text{s.t. } [ [K(E_t, Z^T), e_{1i}] (v + v_i) ] + \eta_i = e_{1i}, \eta_i \geq 0, \quad (17)$$

其中,  $\xi_i, \eta_i$  是松弛变量,  $c_1, c_2$  是非负交换参数。第  $t$  个任务的决策函数可根据式(18)得到:

$$f(x) = \arg \min_{i=1,2} |K(x, Z^T) w_{i} + b_{i}|. \quad (18)$$

## 3 实验结果分析

实验选取 UCI 数据库的 3 个数据集 (<http://www.ics.uci.edu>): Monk, Autistic Spectrum Disorder Screening Data for Adult (ASD), Dermatology。最优参数来自网格搜索法的结果,实验的平均分类准确率结果是通过 3 次交叉验证来获取。参数  $c, \xi, \rho$  的范围为  $\{2^i | i = -3, -2, -1, \dots, 8\}$ , 权重参数范围是  $[0, 1]$ , 这里, 2 个算法模型的参数视为相等的。核函数为径向基函数 (RBF)。实验中数据的基本信息见表 1。

表1 数据集信息

Tab. 1 Dataset information

名称	属性	样本数	类别	任务数
Monk	7	432	2	3
ASD	21	704	2	3
Dermatology	34	366	6	3

3 个数据集在 3 个模型上的平均分类准确率见表 2。通过分析发现,本文算法 WMTLSTSVM 与 MTLSTSVM 和 LSTSVM 相比有更好的分类性能,这充分说明了,给松弛项增加一个权重约束,通过实验把原松弛变量约束项中的 1 转变为范围  $[0, 1]$  中的一个常数,能有效地提高分类精度、降低训练时间,从而得到一个更好的结果。

## 参考文献

- [1] SUYKENS J A, VANDEWALLE J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293.
- [2] SARTAKHTIA J S, AFRABANDPEY H, GHADIRI N. Fuzzy least squares twin Support Vector Machines [J]. Engineering Applications of Artificial Intelligence, 2019, 10(85): 402.

(下转第 127 页)