

文章编号:1007-5321(2018)06-0034-05

DOI:10.13190/j.jbupt.2017-180

最小二乘大间隔孪生支持向量机

吴 青, 齐韶维, 孙凯悦, 臧博研, 赵 祥

(西安邮电大学 自动化学院, 西安 710121)

摘要: 针对最小二乘孪生支持向量机(LSTWSVM)精度较低和可能存在的“奇异性”问题,提出了一种最小二乘大间隔孪生支持向量机(LSLMTSVM)。该算法在最小二乘孪生支持向量机的优化目标函数中引入了间隔分布,提高了算法的泛化性能。在目标函数中加入正则项,实现了结构风险最小化,进一步提高了分类能力。实验结果表明,最小二乘大间隔孪生支持向量机比已有的相关算法性能更优。

关 键 词: 最小二乘; 孪生支持向量机; 间隔分布; 分类

中图分类号: TN181

文献标志码: A

Least Squares Large Margin Twin Support Vector Machine

WU Qing, QI Shao-wei, SUN Kai-yue, ZANG Bo-yan, ZHAO Xiang

(School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

Abstract: In order to overcome low accuracy and possible singularity of least squares twin support vector machine (LSTWSVM), a least squares large margin twin support vector machine (LSLMTSVM) is presented. The proposed algorithm improves generalization performance by introducing margin distribution to the optimization objective function of the LSTWSVM. Additionally, the structural risk minimization principle is implemented by adding the regularization term to the objective function which improves classification ability. Experimental results show that LSLMTSVM has better classification performance than the existing algorithm.

Key words: least squares; twin support vector machine; margin distribution; classification

20 世纪 90 年代 Vapnik 等^[1-2]提出支持向量机(SVM, support vector machine)以来,由于其在解决小样本、非线性及高维模式识别等问题中表现出结构简单、全局最优和泛化能力强等特点,受到了广泛重视。目前, SVM 已成为机器学习领域的研究热点^[3-10]。其基本思想是“间隔”最大化,以此来尽可能地避免错误分类。

Mangasarian 和 Wild^[11]于 2006 年提出了一种两类非平行平面分类器,即广义特征值近似支持向

量机(GEPSVM, generalized eigenvalue proximal support vector machine)。2007 年, Jayadeva 等^[12]在 SVM 和 GEPSVM 的基础上提出了孪生支持向量机(TWSVM, twin support vector machine)。对于一个分类问题, SVM 仅仅是求解一个二次规划问题,而 TWSVM 则是求解 2 个规模较小的二次规划问题,其训练效率大约是 SVM 的 4 倍。实验结果表明,与 SVM 和 GEPSVM 相比较, TWSVM 的分类性能更好。Kumar 和 Gopal^[13]在 2009 年提出的最小二乘孪生

收稿日期: 2017-10-08

基金项目: 国家自然科学基金项目(51875457, 61472307, 51405387); 陕西省重点研发计划项目(2018GY-018); 陕西省教育厅专项科研项目(17JK0713)

作者简介: 吴 青(1975—), 女, 教授, 硕士生导师, E-mail: xiyouwuq@126.com.

支持向量机 (LSTWSVM, least squares twin support vector machine) 和邵元海等^[14]在2011年提出的限定双子支持向量机 (TBSVM, twin bounded support vector machine) 均对 TWSVM 进行了拓展, 实验结果表明, 这2种算法分别使 TWSVM 在训练速度和分类精确度上有所提高. 2016年, 程昊翔和王坚^[15]提出了一种孪生大间隔分布机 (TLDM, twin large margin distribution machine), 该算法在 TWSVM 模型中引入了间隔均值和间隔方差, 实验结果验证了间隔分布的优化对算法的泛化性能有较大影响.

笔者提出了一种最小二乘大间隔孪生支持向量机 (LSLMTSVM, least squares large margin twin support vector machine), 该方法在 TWSVM 的目标函数中引入间隔均值和间隔方差, 通过优化间隔分布来获得具有更强泛化能力的模型. 基于结构风险最小化原则, 在目标函数中引入正则项, 同时将不等式约束转化为等式约束, 将松弛变量改为其二范数变量的平方, 以此去掉非负约束, 进一步改进模型. 实验结果表明, 与 LSTWSVM 和 TLDM 相比, LSLMTSVM 具有更好的分类性能.

1 最小二乘孪生支持向量机

给定 m 个 n 维的训练点, 分别用 $A \in \mathbf{R}^{m_1 \times n}$ 和 $B \in \mathbf{R}^{m_2 \times n}$ 表示 +1 类和 -1 类的数据集, m_1, m_2 分别代表两类数据的数目. LSTWSVM 的目标是在 n 维空间中寻找2个非平行超平面, 求解一对二次规划问题:

$$\min \frac{1}{2} \|Aw_1 + e_1 b_1\|^2 + c_1 \xi_1^T \xi_1 \quad (1)$$

s. t. $-(Bw_1 + e_2 b_1) + \xi_1 = e_2$

$$\min \frac{1}{2} \|Bw_2 + e_2 b_2\|^2 + c_2 \xi_2^T \xi_2 \quad (2)$$

s. t. $(Aw_2 + e_1 b_2) + \xi_2 = e_1$

其中: $w_1, w_2 \in \mathbf{R}^n$; $b_1, b_2 \in \mathbf{R}$; c_1, c_2 为惩罚参数; e_1, e_2 为元素全为1的列向量; ξ_1, ξ_2 为松弛变量.

式(1)关于 w_1 和 b_1 的导数分别为

$$A^T(Aw_1 + e_1 b_1) + c_1 B^T(Bw_1 + e_2 b_1 + e_2) = 0 \quad (3)$$

$$e_1^T(Aw_1 + e_1 b_1) + c_1 e_2^T(Bw_1 + e_2 b_1 + e_2) = 0 \quad (4)$$

定义 $u_1 = [w_1 \ b_1]^T$, $u_2 = [w_2 \ b_2]^T$, 则由式(3)和式(4)可得

$$u_1 = -c_1 (H^T H + c_1 G^T G)^{-1} G^T e_2 \quad (5)$$

其中: $H = [A \ e_1]$; $G = [B \ e_2]$.

对于式(2)同样可以得到

$$u_2 = c_2 (G^T G + c_2 H^T H)^{-1} H^T e_1 \quad (6)$$

若矩阵 $H^T H + c_1 G^T G$ 和 $G^T G + c_2 H^T H$ 存在奇异性, 则原问题的解不存在. 因此, 一般分别采用 $(H^T H + c_1 G^T G + \varepsilon I)^{-1}$ 和 $(G^T G + c_2 H^T H + \varepsilon I)^{-1}$ 代替原逆矩阵, 由此可以克服“奇异性”问题. 其中, $\varepsilon > 0$, I 表示单位矩阵.

非线性情形通过类似方法可以得到相应的解.

2 孪生大间隔分布机

文献[16]中指出, 对于每一个训练数据 (x_i, y_i) 所对应的几何间隔为 $\mu_i = y_i (w^T x_i + b)$, 那么对于聚集在超平面 $w_i^T x + b_i = 0$, $i = 1, 2$ 周围的数据点而言, 对应的间隔均值 $\bar{\mu}_i$ 和间隔方差 $\hat{\mu}_i$ 分别为

$$\bar{\mu}_i = \frac{1}{l} Y^T (Xw_i + b_i e) \quad (7)$$

$$\hat{\mu}_i = \frac{1}{l^2} [l(Xw_i + b_i e)^T (Xw_i + b_i e) - (Xw_i + b_i e)^T Y Y^T (Xw_i + b_i e)] \quad (8)$$

其中: X 为所有训练数据, l 为训练数据点的个数; Y 为一个 $l \times 1$ 的类别标签矩阵; e 为所有元素全为1的 $l \times 1$ 矩阵.

在 TWSVM 的目标函数中增加间隔分布, 通过同时最大化间隔均值以及最小化间隔方差来获得新的优化目标, 可得 TLDM 的目标函数如下:

$$\min_{w_1, b_1, \xi_1} \frac{\lambda_1}{2} \hat{\mu}_1 - \lambda_2 \bar{\mu}_1 + \frac{1}{2} \|Aw_1 + e_1 b_1\|^2 + c_1 e_2^T \xi_1 \quad (9)$$

s. t. $-(Bw_1 + e_2 b_1) + \xi_1 \geq e_2, \xi_1 \geq 0$

$$\min_{w_2, b_2, \xi_2} \frac{\lambda_3}{2} \hat{\mu}_2 - \lambda_4 \bar{\mu}_2 + \frac{1}{2} \|Bw_2 + e_2 b_2\|^2 + c_2 e_1^T \xi_2 \quad (10)$$

s. t. $(Aw_2 + e_1 b_2) + \xi_2 \geq e_1, \xi_2 \geq 0$

引入拉格朗日乘子, 得到对偶问题如下:

$$\min_{\alpha} -\frac{1}{2} \alpha^T G P^{-1} G^T \alpha + (e_2^T - Q^T P^{-1} G^T) \alpha \quad (11)$$

s. t. $0 \leq \alpha \leq c_1 e_2$

$$\min_{\beta} -\frac{1}{2} \beta^T H S^{-1} H^T \beta + (e_1^T - Q^T S^{-1} H^T) \beta \quad (12)$$

s. t. $0 \leq \beta \leq c_2 e_1$

其中: $Q = \frac{\lambda_2}{l} M^T Y$, $P = \frac{\lambda_1}{l^2} D M + H^T H$, $S = \frac{\lambda_1}{l^2} D M + G^T G$, $M = [X, e]$, $D = M^T (I - Y Y^T)$. 这里设定 $\lambda_1 = \lambda_3, \lambda_2 = \lambda_4$. 通过求解可得

$$u_1 = P^{-1} (Q - G^T \alpha) \quad (13)$$

$$u_2 = S^{-1} (Q - H^T \beta) \quad (14)$$

利用核函数可将该方法推广到非线性情况. 文献[15]中所述的实验结果表明, TLDM 具有更好的分类性能, 间隔分布对 TWSVM 的泛化能力有重要影响.

3 最小二乘大间隔孪生支持向量机

为了克服可能出现的“奇异性”问题, 首先将正则项 $\frac{c_3}{2}(\|\mathbf{w}_1\|^2 + b_1^2)$ 和 $\frac{c_4}{2}(\|\mathbf{w}_2\|^2 + b_2^2)$ 分别引入式(1)和式(2)中, 构建基于正则项的 LSTWSVM 模型:

$$\min \frac{c_3}{2}(\|\mathbf{w}_1\|^2 + b_1^2) + \frac{1}{2}\|\mathbf{A}\mathbf{w}_1 + \mathbf{e}_1 b_1\|^2 + \frac{c_1}{2}\xi_1^T \xi_1$$

$$\text{s. t. } -(\mathbf{B}\mathbf{w}_1 + \mathbf{e}_2 b_1) + \xi_1 = \mathbf{e}_2 \quad (15)$$

$$\min \frac{c_4}{2}(\|\mathbf{w}_2\|^2 + b_2^2) + \frac{1}{2}\|\mathbf{B}\mathbf{w}_2 + \mathbf{e}_2 b_2\|^2 + \frac{c_2}{2}\xi_2^T \xi_2$$

$$\text{s. t. } (\mathbf{A}\mathbf{w}_2 + \mathbf{e}_1 b_2) + \xi_2 = \mathbf{e}_1 \quad (16)$$

通过式(15)、式(16)可以看出, 由于新的模型引入变量的平方项, 既保证了目标函数间隔最大, 又使优化问题为严格凸规划.

由拉格朗日乘子法可得模型式(15)、式(16)的解:

$$\mathbf{u}_1 = -c_1(\mathbf{H}^T \mathbf{H} + c_1 \mathbf{G}^T \mathbf{G} + c_3 \mathbf{I})^{-1} \mathbf{G}^T \mathbf{e}_2 \quad (17)$$

$$\mathbf{u}_2 = c_2(\mathbf{G}^T \mathbf{G} + c_2 \mathbf{H}^T \mathbf{H} + c_4 \mathbf{I})^{-1} \mathbf{H}^T \mathbf{e}_1 \quad (18)$$

由式(17)、式(18)可知, 改进后的 LSTWSVM 模型不再具有“奇异性”问题. 引入正则项对 LSTWSVM 模型进行改进, 分别使用 c_3 、 c_4 取代 ε . c_3 、 c_4 为结构风险的权重系数, 文献[14]表明, 通过调节这 2 个参数可以提高分类精度, 比仅调节固定参数 ε 更为合理.

结合 TLDM 的思想, 将间隔均值和间隔方差引入改进的 LSTWSVM 中, 构建 LSLMTSVM 模型:

$$\min \frac{\lambda_1}{2}\hat{\mu}_1 - \lambda_2 \bar{\mu}_1 + \frac{c_3}{2}(\|\mathbf{w}_1\|^2 + b_1^2) +$$

$$\frac{1}{2}\|\mathbf{A}\mathbf{w}_1 + \mathbf{e}_1 b_1\|^2 + \frac{c_1}{2}\xi_1^2$$

$$\text{s. t. } -(\mathbf{B}\mathbf{w}_1 + \mathbf{e}_2 b_1) + \xi_1 = \mathbf{e}_2 \quad (19)$$

$$\min \frac{\lambda_3}{2}\hat{\mu}_2 - \lambda_4 \bar{\mu}_2 + \frac{c_4}{2}(\|\mathbf{w}_2\|^2 + b_2^2) +$$

$$\frac{1}{2}\|\mathbf{B}\mathbf{w}_2 + \mathbf{e}_2 b_2\|^2 + \frac{c_2}{2}\xi_2^2$$

$$\text{s. t. } (\mathbf{A}\mathbf{w}_2 + \mathbf{e}_1 b_2) + \xi_2 = \mathbf{e}_1 \quad (20)$$

令式(19)关于 \mathbf{w}_1 和 b_1 的偏导数分别为 0,

可得

$$\left(\frac{\lambda_1}{l^2} \mathbf{D}\mathbf{M} + c_3 \mathbf{I} + \mathbf{H}^T \mathbf{H} + c_1 \mathbf{G}^T \mathbf{G}\right) \mathbf{u}_1 -$$

$$\frac{\lambda_2}{l} \mathbf{M}^T \mathbf{Y} + c_1 \mathbf{G}^T \mathbf{e}_2 = 0 \quad (21)$$

或

$$\mathbf{u}_1 = \left(\frac{\lambda_1}{l} \mathbf{D}\mathbf{M} + c_3 \mathbf{I} + \mathbf{H}^T \mathbf{H} + c_1 \mathbf{G}^T \mathbf{G}\right)^{-1} \times$$

$$\left(\frac{\lambda_2}{l} \mathbf{M}^T \mathbf{Y} - c_1 \mathbf{G}^T \mathbf{e}_2\right) \quad (22)$$

类似地, 可以得到式(20)的解:

$$\mathbf{u}_2 = \left(\frac{\lambda_3}{l} \mathbf{D}\mathbf{M} + c_4 \mathbf{I} + \mathbf{G}^T \mathbf{G} + c_2 \mathbf{H}^T \mathbf{H}\right)^{-1} \times$$

$$\left(\frac{\lambda_4}{l} \mathbf{M}^T \mathbf{Y} - c_2 \mathbf{H}^T \mathbf{e}_1\right) \quad (23)$$

通过计算, 可以得到一对非平行超平面 (\mathbf{w}_1, b_1) 和 (\mathbf{w}_2, b_2) . 对于任意的 $\mathbf{x} \in \mathbf{R}^n$, 可以通过

$$i = \arg \min_{k=1,2} \frac{|\mathbf{w}_k^T \mathbf{x} + b_k|}{\|\mathbf{w}_k\|} \quad (24)$$

判别该点属于哪一类, 其中, $|\cdot|$ 表示绝对值符号.

引入核函数 $K(\cdot)$, 可将 LSLMTSVM 推广到非线性情形, 模型如下:

$$\min \frac{\lambda_1}{2}\hat{\mu}_1 - \lambda_2 \bar{\mu}_1 + \frac{c_3}{2}(\|\mathbf{u}_1\|^2 + b_1^2) +$$

$$\frac{1}{2}\|K(\mathbf{A}, \mathbf{C}^T) \mathbf{u}_1 + \mathbf{e}_1 b_1\|^2 + \frac{c_1}{2}\xi_1^2$$

$$\text{s. t. } -(K(\mathbf{B}, \mathbf{C}^T) \mathbf{u}_1 + \mathbf{e}_2 b_1) + \xi_1 = \mathbf{e}_2 \quad (25)$$

$$\min \frac{\lambda_3}{2}\hat{\mu}_2 - \lambda_4 \bar{\mu}_2 + \frac{c_4}{2}(\|\mathbf{u}_2\|^2 + b_2^2) +$$

$$\frac{1}{2}\|K(\mathbf{B}, \mathbf{C}^T) \mathbf{u}_2 + \mathbf{e}_2 b_2\|^2 + \frac{c_2}{2}\xi_2^2$$

$$\text{s. t. } K(\mathbf{A}, \mathbf{C}^T) \mathbf{u}_2 + \mathbf{e}_1 b_2 + \xi_2 = \mathbf{e}_1 \quad (26)$$

其中 $\mathbf{C}^T = (\mathbf{A}^T, \mathbf{B}^T)$.

记 $\mathbf{v}_1 = [\mathbf{u}_1 \ b_1]^T$, $\mathbf{v}_2 = [\mathbf{u}_2 \ b_2]^T$, 对式(25)、式(26)的拉格朗日函数分别求偏导, 并令其为 0, 可得

$$\mathbf{v}_1 = \left(\frac{\lambda_1}{l} \mathbf{D}\mathbf{M} + c_3 \mathbf{I} + \mathbf{H}^T \mathbf{H} + c_1 \mathbf{G}^T \mathbf{G}\right)^{-1} \times$$

$$\left(\frac{\lambda_2}{l} \mathbf{M}^T \mathbf{Y} - c_1 \mathbf{G}^T \mathbf{e}_2\right) \quad (27)$$

$$\mathbf{v}_2 = \left(\frac{\lambda_3}{l} \mathbf{D}\mathbf{M} + c_4 \mathbf{I} + \mathbf{G}^T \mathbf{G} + c_2 \mathbf{H}^T \mathbf{H}\right)^{-1} \times$$

$$\left(\frac{\lambda_4}{l} \mathbf{M}^T \mathbf{Y} - c_2 \mathbf{H}^T \mathbf{e}_1\right) \quad (28)$$

其中: $\boldsymbol{H} = [K(\boldsymbol{A}, \boldsymbol{C}^T) \quad \boldsymbol{e}_1]; \boldsymbol{G} = [K(\boldsymbol{B}, \boldsymbol{C}^T) \quad \boldsymbol{e}_2]$.
通过计算,可以得到一对非平行超平面 (\boldsymbol{u}_1, b_1) 和 (\boldsymbol{u}_2, b_2) . 对于任意给出的数据点 $\boldsymbol{x} \in \mathbf{R}^n$, 可以根据

$$i = \arg \min_{k=1,2} \frac{|K(\boldsymbol{x}^T, \boldsymbol{C}^T) \boldsymbol{u}_k + b_k|}{\sqrt{\boldsymbol{u}_k^T K(\boldsymbol{C}, \boldsymbol{C}^T) \boldsymbol{u}_k}} \quad (29)$$

来判断其类别.

4 实验分析

为了验证所提算法的性能,将 LSLMTSVM 与 TLDM 和 LSTWSVM 算法进行分类精度和训练效率的测试实验,并对结果进行比较. 实验采用 2.20 GHz CPU, 4.0 GB RAM 的 PC 机,所用软件为 matlab 8.0. 实验中所采用的 8 个数据集均来自 UCI 数据库^[17]. 实验中所用的核函数均为高斯核函数.

为了避免由于数据随机选取产生的不确定性,

精确度可以通过多次计算取平均值得到. 表 1 和表 2 分别列出了线性和非线性情况下 LSLMTSVM 与另外 2 种算法的结果,其中包含分类的精确度和训练时间. 从表 1 可以看出, LSLMTSVM 的分类精确度明显优于 LSTWSVM 和 TLDM,而且在速度上也比 TLDM 快. 表 2 所示的实验结果表明,在除了 German、Checkboard 以外的数据集上, LSLMTSVM 相比于 LSTWSVM 和 TLDM 的分类精度更高,而在所有数据集上的训练速度都比 TLDM 要快. 由表 1 和表 2 可知, LSLMTSVM 和 LSTWSVM 的训练时间差别很小,这是因为这 2 种算法均采用了最小二乘方法. 但由于 LSLMTSVM 模型较复杂,所以所需的时间也长一些. 实验结果表明, LSLMTSVM 的分类性能要优于另外 2 种算法.

表 1 线性情况下算法的性能比较

数据集	TLDM		LSTWSVM		LSLMTSVM	
	精确度/%	t/s	精确度/%	t/s	精确度/%	t/s
German (1 000 × 24)	74.30	0.27	72.50	0.004 0	78.40	0.010 0
Ionosphere (351 × 34)	90.74	0.09	88.89	0.002 0	92.59	0.002 0
Sonar (208 × 60)	86.80	0.15	85.37	0.004 1	88.56	0.004 4
Wpbc (198 × 33)	90.00	0.12	87.80	0.003 4	95.00	0.003 7
Bupa (345 × 6)	71.01	0.13	66.67	0.002 3	73.91	0.002 3
Wdbc (569 × 31)	85.35	0.07	86.67	0.001 8	88.24	0.002 1
Checkboard (1 000 × 2)	52.50	0.27	51.00	0.080 0	54.00	0.079 0
Diabetes (768 × 8)	81.70	0.25	70.59	0.003 0	83.66	0.010 40

表 2 非线性情况下算法的性能比较

数据集	TLDM		LSTWSVM		LSLMTSVM	
	精确度/%	t/s	精确度/%	t/s	精确度/%	t/s
German (1 000 × 24)	73.50	2.260 0	78.00	1.180 0	74.10	1.670 0
Ionosphere (351 × 34)	95.71	0.120 0	92.86	0.045 0	97.14	0.048 0
Sonar (208 × 60)	90.24	0.059 0	88.70	0.012 6	92.12	0.013 4
Wpbc (198 × 33)	79.49	0.066 0	71.79	0.015 4	84.21	0.015 4
Bupa (345 × 6)	66.67	0.220 0	60.87	0.067 0	68.12	0.110 0
Wdbc (569 × 31)	67.26	0.260 0	70.80	0.147 0	72.57	0.149 0
Checkboard (1 000 × 2)	88.00	3.720 0	85.00	1.530 0	86.50	1.582 0
Diabetes (768 × 8)	69.93	0.810 0	68.63	0.447 0	75.00	0.665 0

5 结束语

针对 LSTWSVM 存在分类精度较低和可能存在

的“奇异性”等问题,提出了一种 LSLMTSVM. 该算法通过引入间隔分布和新的正则项,对 LSTWSVM 模型进行了改进. 通过引入间隔分布,提高了模型

的泛化能力;引入正则项实现了结构风险最小化,进一步提高了分类精度。由于算法采用了等式约束,其训练速度也比一般的二次规划问题求解要快。如何选取算法中提到的多个参数,以进一步提高本文算法的分类性能将是下一步的研究工作。

参考文献:

- [1] Cortes C, Vapnik V N. Support-vector networks [J]. *Machine Learning*, 1995, 20(3): 273-297.
- [2] Vapnik V N. The nature of statistical learning theory [J]. *Technometrics*, 1997, 38(4): 409.
- [3] 丁晓剑, 赵银亮. 双边界支持向量机的理论研究与分析 [J]. *北京邮电大学学报*, 2010, 33(2): 20-23.
Ding Xiaojian, Zhao Yinliang. Theory and analysis of double margin SVM [J]. *Journal of Beijing University of Posts and Telecommunications*, 2010, 33(2): 20-23.
- [4] 马跃峰, 梁循, 周小平. 一种基于全局代表点的快速最小二乘支持向量机稀疏化算法 [J]. *自动化学报*, 2017, 43(1): 132-141.
Ma Yuefeng, Liang Xun, Zhou Xiaoping. A fast sparse algorithm for least squares support vector machine based on global representative points [J]. *Acta Automatica Sinica*, 2017, 43(1): 132-141.
- [5] 陈素根, 吴小俊. 改进的投影孪生支持向量机 [J]. *电子学报*, 2017, 45(2): 408-416.
Chen Sugan, Wu Xiaojun. Improved projection twin support vector machine [J]. *Acta Electronica Sinica*, 2017, 45(2): 408-416.
- [6] 刘春红, 韩晶晶, 商彦磊, 等. 基于 SVM 分类的云集群失败作业主动预测方法 [J]. *北京邮电大学学报*, 2016, 39(5): 104-109.
Liu Chunhong, Han Jingjing, Shang Yanlei, et al. Predicting job failure in cloud cluster: based on SVM classification [J]. *Journal of Beijing University of Posts and Telecommunications*, 2016, 39(5): 104-109.
- [7] Qi Zhiquan, Tian Yingjie, Shi Yong. Robust twin support vector machine for pattern classification [J]. *Pattern Recognition*, 2013, 46(1): 305-316.
- [8] Shao Yuanhai, Deng Naiyang, Yang Zhimin. Least squares recursive projection twin support vector machine for classification [J]. *Pattern Recognition*, 2012, 45(6): 2299-2307.
- [9] Chen Sugan, Wu Xiaojun. A new fuzzy twin support vector machine for pattern classification [J]. *International Journal of Machine Learning and Cybernetics*, 2018, 9(9): 1553-1564.
- [10] Tanveer M, Khan M A, Ho S S. Robust energy-based least squares twin support vector machines [J]. *Applied Intelligence*, 2016, 45(1): 174-186.
- [11] 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.
- [12] Jayadeva, Khemchandani R, Chandra S. Twin support vector machines for pattern classification [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(5): 905-910.
- [13] Kumar M A, Gopal M. Least squares twin support vector machines for pattern classification [J]. *Expert Systems with Applications*, 2009, 36(4): 7535-7543.
- [14] Shao Yuanhai, Zhang Chunhua, Wang Xiaobo, et al. Improvements on twin support vector machines [J]. *IEEE Transactions on Neural Networks*, 2011, 22(6): 962-968.
- [15] 程昊翔, 王坚. 一种新的孪生大间隔分布机算法 [J]. *控制与决策*, 2016, 31(5): 949-952.
Cheng Haoxiang, Wang Jian. A novel twin large margin distribution machine [J]. *Control and Decision*, 2016, 31(5): 949-952.
- [16] Vapnik V N. Statistical learning theory [J]. *Encyclopedia of the sciences of learning*, 1998, 41(4): 3185.
- [17] Blake C. UCI repository of machine learning databases [EB/OL]. [2017-06-25]. <http://www.ics.uci.edu/~mlearn/MLRepository.html>.