第 33 卷第 8 期 电 子 与 信 息 学 报 Vol.33No.8
2011 年 8 月 Journal of Electronics & Information Technology Aug. 2011

无偏置ν-SVM分类优化问题研究

丁晓剑* 赵银亮
(西安交通大学电子与信息工程学院 西安 710049)

摘 要：在高维空间中，分类超平面倾向于通过原点，即不需要偏置(b)。为了研究在ν- SVM 分类问题中是否需要 b，该文提出了无(b)的ν- SVM 的对偶优化问题并给出了其优化问题求解方法。该方法通过有效集策略将对偶优化问题转化为等式约束子优化问题，然后通过拉格朗日乘子法将子优化问题转化为线程方程组来求解。实验表明偏置(b)的存在会降低ν- SVM 的泛化性能，ν- SVM 只能得到无(b) ν- SVM 的次优解。

关键词：ν- 支持向量机；偏置；泛化性能；有效集
中图分类号：TP181 **文献标识码**：A **文章编号**：1009-5896(2011)08-1998-05
DOI: 10.3724/SP.J.1146.2010.01286

Study on ν-SVM for Classification Optimization Problem without Bias

Ding Xiao-jian Zhao Yin-liang
(*School of Electrical and Information Engineering*, *Xi'an Jiaotong University*, *Xi'an* 710049, *China*)

Abstract: In the high-dimensional space, the classification hyperplane tends to pass through the origin and bias (b) is not need. To study whether ν-SVM for classification needs (b), dual optimization formulation of ν-SVM without (b) is proposed and the corresponding method of solving the optimization formulation is presented. The dual optimization formulation is transformed into equality constraint sub-optimization formulation by the active set strategy in this method, then the sub-optimization formulation is transformed into the linear equation by lagrange multiplier method. The experimental results show that the existence of (b) would reduce the generalization ability of ν-SVM and ν-SVM can only obtain the sub-optimal solution of ν-SVM without b.

Key words: ν- Support Vector Machine (SVM); Bias; Generalization ability; Active set

1 引言

C-SVM(Support Vector Machine)起源于文献[1]提出的支持向量网络(support vector networks)。在一般的神经网络结构中，只需调节输出向量和隐藏层连接的权重向量 $\boldsymbol{w}$ 就可以较好地拟合训练样本，而文献[1]指出需要调节向量 $\boldsymbol{w}$ 和(b)才能得到最优分类超平面，但在文中并没有提到(b)存在的意义。Poggio 等人[2]从核函数的正定理论分析得出如果核函数是正定的(positive definite)，在 SVM 优化问题中是不需要(b)的，文献[3]指出使用高斯核(gaussian kernel)等常用核函数的 SVM 优化问题不需要(b)。Huang 等人[4]提出了优化极限学习机(Extreme Learning Machine, ELM)的分类方法，指出优化极限学习机的模型和C-SVM 是等价的，决策函数倾向于通过原点，即不需要(b)的存在。文献[5]研究了(b)对C-SVM 分类优化问题泛化性能的影响，在实验中验证得到无(b) C-SVM 优化问题的泛化性能更好，对参数C不太敏感，计算代价更小等一系列性质。虽然ν-SVM[6, 7]与C-SVM 的目标优化问题不同，约束条件也不同，但两者得到的最优分类超平面是等价的，即文献[5]的结论应该适用于ν-SVM。本文首先给出无(b) ν-SVM 优化问题的描述，然后给出了利用有效集方法求解无(b) ν-SVM 优化问题的子优化问题求解方法，在标准数据集上的多个指标的性能分析得到了理想的效果。

2 无(b) v - SVM优化问题

给定一个含有 m 个样本的训练样本集 $\{\boldsymbol{x}_i, y_i\}_{i=1}^m$，其中输入为 d 维向量 $\boldsymbol{x}_i \in \boldsymbol{R}^d$，输出为 $y_i \in \boldsymbol{R}$。通常训练样本集在输入空间中是线性不可分的，需要引入映射函数 $\phi(\boldsymbol{x})$ 将 $\boldsymbol{x}_i$ 映射到高维空间 $\phi(\boldsymbol{x}_i)$ 中。训练无(b) ν-SVM 分类问题等价于求解如下优化问题：

$$\left.\begin{aligned}&\text{Minimize}: \tau_2(\boldsymbol{w},\boldsymbol{\xi},\rho)=\frac{1}{2}\|\boldsymbol{w}\|^2-\nu\rho+\frac{1}{m}\sum_{i=1}^{m}\xi_i\\&\text{s.t.}\quad y_i\boldsymbol{w}\cdot\phi(\boldsymbol{x}_i)\ge\rho-\xi_i,\xi_i\ge 0,\ i=1,\cdots,m,\ \rho\ge 0\end{aligned}\right\}\quad(1)$$

引入拉格朗日函数：

2010-11-22 收到，2011-04-18 改回
国家 863 计划项目(2008AA01Z136)资助课题
*通信作者：丁晓剑 xjding@stu.xjtu.edu.cn

$$L(\boldsymbol{w},\boldsymbol{\xi},\rho,\boldsymbol{\alpha},\boldsymbol{\beta},\delta)=\frac{1}{2}\|\boldsymbol{w}\|^2-\nu\rho+\frac{1}{m}\sum_{i=1}^{m}\xi_i-\sum_{i=1}^{m}(\alpha_i(y_i\boldsymbol{w}\cdot\phi(\boldsymbol{x}_i)-\rho+\xi_i)+\beta_i\xi_i)-\delta\rho \tag{2}$$

其中 α_i,β_i,δ 为非负拉格朗日乘子。

对 L 关于变量 $\boldsymbol{w},\boldsymbol{\xi},\rho$ 求偏导数，下列条件都满足时得到式(1)的最优解：

$$\partial L/\partial\boldsymbol{w}=0\Rightarrow\boldsymbol{w}=\sum_{i=1}^{m}\alpha_i y_i\phi(\boldsymbol{x}_i) \tag{3}$$

$$\partial L/\partial\boldsymbol{\xi}=0\Rightarrow\alpha_i+\beta_i=1/m \tag{4}$$

$$\partial L/\partial\rho=0\Rightarrow\sum_{i=1}^{m}\alpha_i-\delta=\nu \tag{5}$$

将式(3)和式(4)代入拉格朗日函数 L，可得无(b) ν-SVM 的对偶优化问题：

$$\left.\begin{aligned}&\text{Minimize}:\frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j\phi(\boldsymbol{x}_i)\cdot\phi(\boldsymbol{x}_j)\\&\text{s.t.}\quad 0\le\alpha_i\le 1/m,\quad\sum_{i=1}^{m}\alpha_i\ge\nu\end{aligned}\right\} \tag{6}$$

其中式(6)的不等式约束 $\sum_{i=1}^{m}\alpha_i\ge\nu$ 同样可以转化为等式约束 $\sum_{i=1}^{m}\alpha_i=\nu$ 求解。根据式(6)，无 b ν-SVM 决策函数为

$$f(\boldsymbol{x})=\sum_{i=1}^{m}\alpha_i y_i K(\boldsymbol{x},\boldsymbol{x}_i) \tag{7}$$

2.1 Karush-Kuhn-Tucker 条件

为了得到原始优化问题式(1)的最优解，需要推导出它的完整 Karush-Kuhn-Tucker(KKT)条件：

初始可行性：

$$\xi_i\ge 0,\ \ \rho\ge 0,\ y_i\boldsymbol{w}\cdot\phi(\boldsymbol{x}_i)\ge\rho-\xi_i,\ \forall i \tag{8}$$

对偶可行性：

$$\alpha_i\ge 0,\ \beta_i\ge 0,\ \delta\ge 0,\ \forall i \tag{9}$$

对偶与原问题的补条件：

$$\alpha_i(y_i\boldsymbol{w}\cdot\phi(\boldsymbol{x}_i)-\rho+\xi_i)=0,\ \beta_i\xi_i=0,\ \delta\rho=0,\ \forall i \tag{10}$$

由于 $f(\boldsymbol{x})$ 为无(b) ν-SVM 的决策函数，令 f_i 为无(b) ν-SVM 分类器对样本 $\boldsymbol{x}_i$ 预测的值，并有 $f_i=y_i\boldsymbol{w}\cdot\phi(\boldsymbol{x}_i)$。由最优解条件 $\alpha_i+\beta_i=1/m$ 和补条件 $\beta_i\xi_i=0$ 可得

$$(1/m-\alpha_i)\xi_i=0 \tag{11}$$

如果 $\alpha_i=0$，由式(11)可知 $\xi_i=0$，由最优解条件 $\sum_{i=1}^{m}\alpha_i-\delta=\nu$ 和 $\delta\rho=0$ 可知 $\rho=0$，再由式(8)可知 $f_i\ge\rho$，进而有 $f_i\ge 0$。如果 $\alpha_i=1/m$，由式(10)可知 $f_i-\rho+\xi_i=0$，由最优解条件 $\sum_{i=1}^{m}\alpha_i-\delta=\nu$ 和补条件 $\delta\rho=0$ 可知 $\rho=0$，又由式(8)可知 $f_i=-\xi_i\le 0$。如果 $0<\alpha_i<1/m$，由补条件式(10)可知 $f_i-\rho+\xi_i=0$，由式(11)可知 $\xi_i=0$，再由式(8)可知 $f_i=\rho\ge 0$。综上，式(3)的最优解满足下列条件：

$$\left.\begin{aligned}&0\le\alpha_i<1/m\quad\Leftrightarrow\quad f_i\ge 0\\&\alpha_i=1/m\qquad\ \Leftrightarrow\quad f_i\le 0\end{aligned}\right\} \tag{12}$$

3　优化问题求解方法

文献[8]利用有效集算法来求解 C-SVM 优化问题，该有效集算法将 C-SVM 对偶优化问题分解为一系列子优化问题来求解，其主要计算代价是每次迭代过程中求解子优化问题，是一种高效简便的算法。由于无(b) ν-SVM 的对偶优化问题都与 C-SVM 对偶优化问题不同，文献[8]中的有效集算法无法直接应用，在本节将推导无(b) ν-SVM 子优化问题的求解方法。

在此先定义一些需要用到的向量和集合。定义 3 个集合：$S_0:=\{i\mid\alpha_i=0\}$，$S_m:=\{i\mid\alpha_i=1/m\}$ 和 $S_{\text{work}}:=\{i\mid\alpha_i\in(0,1/m)\}$，有 $S_0\cup S_{\text{work}}\cup S_m=[1:m]$，同理可定义向量 $\boldsymbol{\alpha}_0$，$\boldsymbol{\alpha}_m$ 和 $\boldsymbol{\alpha}_{\text{work}}$。

式(8)为无(b) ν-SVM 的对偶优化问题，为方便推导求解，将式(8)写成下面紧凑的形式：

$$\left.\begin{aligned}&\underset{\boldsymbol{\alpha}}{\text{Minimize}}:\ \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{\alpha}\\&\text{s.t.}\ \ 0\le\alpha_i\le 1/m\ \text{and}\ \sum\boldsymbol{\alpha}=\nu,\ i=1,\cdots,m\end{aligned}\right\} \tag{13}$$

其中 $\boldsymbol{H}=\boldsymbol{K}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\in\boldsymbol{R}^{N\times N}$，$\boldsymbol{K}$ 为式(7)中对应的核矩阵，$\boldsymbol{\alpha}=[\alpha_1,\cdots,\alpha_N]^{\mathrm{T}}\in\boldsymbol{R}^N$，$\boldsymbol{Y}=[y_1,\cdots,y_N]^{\mathrm{T}}\in\boldsymbol{R}^N$。

同样由向量 $\boldsymbol{\alpha}_0$，$\boldsymbol{\alpha}_m$ 的定义可知 $\boldsymbol{\alpha}_0$，$\boldsymbol{\alpha}_m$ 元素的值是固定的，求解式(13)等价于求解下式：

$$\left.\begin{aligned}&\underset{\boldsymbol{\alpha}_{\text{work}}}{\text{Minimize}}:\ \frac{1}{2}\boldsymbol{\alpha}_{\text{work}}^{\mathrm{T}}\boldsymbol{H}_{ww}\boldsymbol{\alpha}_{\text{work}}+\boldsymbol{\alpha}_m^{\mathrm{T}}\boldsymbol{H}_{wm}\boldsymbol{\alpha}_{\text{work}}\\&\text{s.t.}\ \ 0\le\alpha_i\le 1/m,\ \sum\boldsymbol{\alpha}_{\text{work}}=\nu-\sum\boldsymbol{\alpha}_m\end{aligned}\right\} \tag{14}$$

其中 $\boldsymbol{H}_{ww}$ 和 $\boldsymbol{H}_{wm}$ 为 $\boldsymbol{H}$ 的子矩阵，行和列相应的集合 S_m 和 S_{work} 标识。同理由于式(14)是对 $\boldsymbol{\alpha}_{\text{work}}$ 求最小值，由 $\boldsymbol{\alpha}_{\text{work}}$ 的定义可知 $\boldsymbol{\alpha}_{\text{work}}$ 中分量的值都在区间 $(0,1/m)$ 中，即约束 $0\le\alpha_i\le 1/m$ 必然满足，这样式(14)就变为等式约束优化问题：

$$\left.\begin{aligned}&\underset{\boldsymbol{\delta}}{\text{Minimize}}:\ \frac{1}{2}\boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\delta}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\delta}\\&\text{s.t.}\ \ \boldsymbol{A}\boldsymbol{\delta}=t\end{aligned}\right\} \tag{15}$$

其中 $\boldsymbol{\delta}=\boldsymbol{\alpha}_{\text{work}}$，$\boldsymbol{Q}=\boldsymbol{H}_{ww}$，$\boldsymbol{c}=\boldsymbol{H}_{wm}\boldsymbol{\alpha}_m$，$\boldsymbol{A}=[1,\cdots,1]^{\mathrm{T}}\in\boldsymbol{R}^n$，$n$ 为 $\boldsymbol{\alpha}_{\text{work}}$ 中元素的个数，$t=\nu-\sum\boldsymbol{\alpha}_m$。

式(13)可以通过解条件问题的乘子法来求解，构造拉格朗日函数：

$$L_1(\boldsymbol{\delta},\boldsymbol{\lambda})=\frac{1}{2}\boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\delta}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\delta}+\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{\delta}-t),\ \ \boldsymbol{\lambda}\in\boldsymbol{R}^n \tag{16}$$

令 $L_1(\boldsymbol{\delta},\boldsymbol{\lambda})$ 对 $\boldsymbol{\delta}$ 和 $\boldsymbol{\lambda}$ 的导数为零，得到线性方程

组：

$$\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\lambda} \end{bmatrix}=\begin{bmatrix} -\boldsymbol{c} \\ t \end{bmatrix} \tag{17}$$

其中 $\boldsymbol{0}=[0,\cdots,0]\in \boldsymbol{R}^n$, $\boldsymbol{\delta}$ 和 $\boldsymbol{\lambda}$ 可通过求解式(17)得到。

现已将无(b) ν-SVM 子优化问题转化为与 C-SVM 子优化问题相同的形式。有效集算法利用 2 层循环来求解目标优化问题式(13)：外循环判断 $\boldsymbol{\alpha}$ 中边界约束元素是否都满足 KKT 条件式(12)，如果都满足，算法终止。如果不满足，则选取违反 KKT 条件的变量进入内循环求解；而内循环目标是求解子优化问题式(15)，使得最优解 $\boldsymbol{\alpha}_{\mathrm{work}}$ 满足上下界约束。

4 实验与结果分析

为了研究(b)对 ν-SVM 分类性能的影响(本文仅讨论两类分类问题)，本文将用有效集算法对 ν-SVM (nu-AS)和无(b) ν-SVM (nbnu-AS)进行多个指标的比较，算法是在 Pentium 4, 2.53 GHz, CPU, MATLAB 2007 环境下实现，实验数据集分别来自 UCI 数据库[9]，Statlog 数据库[10]，所有样本的输入都归一化到[0,1]之间。

4.1 参数设定

对于 nu-AS 算法和 nbnu-AS 算法性能比较使用高斯核($K(\boldsymbol{u},\boldsymbol{v})=\exp(-\|\boldsymbol{u}-\boldsymbol{v}\|^2/2\gamma^2)$)(注：此处公式已作改动)作为核函数，需要合适的核参数和代价参数 ν 才能得到尽量好的分类性能。代价参数 ν 的选择按照文献[6]的选择方法，10 个参数 ν 值分别取：0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0。对高斯核，对于每个数据集选取 15 个 γ 值和 10 个 ν 值，共有 150 种参数对，并选择性能最好的组合 (ν,γ)，参照文献[11]中的参数选择方法，15 个 γ 值分别取：0.001, 0.01, 0.1, 0.2, 0.4, 0.8, 1, 2, 5, 10, 20, 50, 100, 1000, 10000。对于每个数据集采用 50 次测试，每次测试随机训练样本数目和测试样本数目。实验指标包括平均训练时间(ATT)，平均测试精度(ATA)，标准方差(DEV)和平均的支持向量数目(ASV)。

4.2 实验性能比较

表 1 是 nu-AS 和 nbnu-AS 算法的性能指标比较。在 13 个数据集中，nbnu-AS 算法在其中 10 个数据集上的平均测试精度要好，nu-AS 算法在 Monk's Problem 1 数据集上的平均测试精度要好，在另外两个数据集上两个算法的平均测试精度一致。两个算法在所有数据集上的平均训练时间相差都不大。

ν-SVM 和无(b) ν-SVM 的求解是将原始优化问题转化为对偶优化问题来求解，由于对偶间隙的存在，对偶优化问题的最优解并非一定为原始优化问题的最优解。尽管无(b) ν-SVM 在 Monk's Problem 1 数据集上得到更小的函数值(见表 2)，但是无(b) ν-SVM 得测试精度仍然低于 ν-SVM 。

4.3 ν-SVM 和无(b) ν-SVM 优化问题最优解比较

由于 ν-SVM 优化问题与无(b) ν-SVM 优化问题多出一个约束条件：$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{Y}=0$，从优化问题的解空间上分析后者应该能得到比前者更优的解。本节从

表 1 标准数据集上的性能比较

数据集	nu-AS			nbnu-AS		
	ATT (s)	ATA (%)	DEV (%)	ATT (s)	ATA (%)	DEV (%)
Heart	0.0340	75.67	2.82	0.0374	**75.97**	2.87
Pwlinear	0.0367	82.76	3.31	0.0409	**83.28**	3.51
Sonar	0.1052	82.87	3.65	0.0669	**83.69**	3.32
Monk's Problem 1	0.0797	**95.37**	0	0.0853	95.14	0
Liver-disorders	0.1104	70.99	4.54	0.1216	**71.38**	3.93
Ionosphere	0.2381	89.67	1.96	0.1480	**90.56**	2.10
Breast-cancer	0.3867	94.34	1.47	0.2977	**95.28**	0.94
Australian	0.3251	67.76	1.28	0.3355	**67.87**	1.28
Pimadata	0.7569	76.08	2.31	0.6946	**76.51**	1.15
Creat	0.4176	79.76	2.82	0.4297	**81.42**	1.92
A1a	32.12	84.00	0	30.75	**84.01**	0
A2a	173.8	84.21	0	170.5	84.21	0
A3a	395.2	84.36	0	390.1	84.36	0

表 2 nu-AS 和 nbnu-AS 算法的最优函数值比较

数据集	nu-AS	nbnu-AS
Heart	1.1322e−008	**1.1301e−008**
Pwlinear	9.5643e−002	**9.5043e−002**
Sonar	1.7109e−004	**1.6980e−004**
Monk's Problem 1	3.2614e−005	**3.2589e−005**
Liver-disorders	2.8910e−007	**2.6988e−007**
Ionosphere	6.5368e−005	**4.7175e−005**
Breast-cancer	1.9366e−003	**1.9353e−003**
Australian	6.1230e−003	**5.3609e−003**
Pimadata	1.1662e−005	**1.1485e−005**
Creat	4.1906e−004	**2.5338e−010**
A1a	1.8590e−004	1.8590e−004
A2a	2.3716e−004	2.3716e−004
A3a	3.5295e−004	3.5295e−004

实验上分析该约束条件对优化问题最优解影响，利用有效集算法和最优函数值指标对两个优化问题进行比较。为了使两个目标优化问题相同(约束条件不同)，nu-AS 和 nbnu-AS 的参数须设置一致(从而优化问题中的核矩阵也相同)，在此将两个算法的参数对都设为 nu-AS 的最优参数对，表 2 为比较结果。在前 10 个数据集上，nu-AS 算法在两个核函数上都取得了更优的解，而在后 3 个数据集上，两个算法在两个核函数上的最优解都相同。设 nu-AS 算法的解空间为 S_1，nbnu-AS 算法的解空间为 S_2，则有 $S_1 \subseteq S_2$。

4.4 参数敏感性分析

本节比较核参数和代价参数 ν 对算法测试精度的影响，实验数据选取 Pwlinear 数据集，高斯核的参数值按照 4.1 节选取。

从图 1 和图 2 可以看出，nu-AS 在高斯核上的不同参数选取对于测试精度影响很大，而且没有规律。nbnu-AS 算法则对参数 ν 不太敏感，曲面比较光滑。在实际参数选取中，对于参数 ν 可以扩大参数选取值的间隔，利用较少的参数值对就可以找到最优测试精度。

5 结束语

由于 ν-SVM 引入了有实际意义的参数 ν，在机器学习的各个领域中有广泛的应用和实际的指导意义。本文研究对比了 ν-SVM 和无 (b) 的 ν-SVM 优化问题，从多个性能指标的实验上分析得出无 (b) 的 ν-SVM 的泛化性能要好于 ν-SVM，在同等目标优化问题下 ν-SVM 只能得到无 (b) 的 ν-SVM 的次优解，无 (b) 的 ν-SVM 对参数 ν 不太敏感，在较少的参数值对中就可以找到最佳的测试精度。

本文的工作适用于 ν-SVM 的改进算法，可以将无 (b) 的 ν-SVM 应用于现有的对 ν-SVM 进行改进的算法。在以后的研究工作中将无 (b) 的 ν-SVM 应用于文本挖掘，图像识别等实际领域中，进一步验证无 (b) 的 ν-SVM 的有效性。

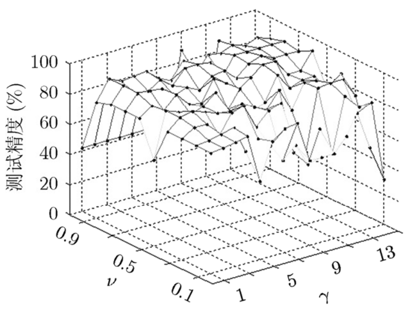

图 1 nu-AS 在高斯核上的测试精度

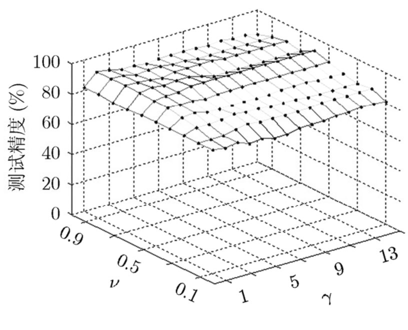

图 2 nbnu-AS 在高斯核上的测试精度

参 考 文 献

[1] Cortes C and Vapnik V. Support vector networks. *Machine Learning*, 1995, 20(3): 273–297.

[2] Poggio T, Mukherjee S, Rifkin R, *et al.*. "b". (A.I. Memo No. 2001-011, CBCL Memo 198, Artificial Intelligence Laboratory, Massachusetts Institute of Technology), 2001.

[3] Evgeniou T, Pontil M, and Poggio T. Regularization networks and support vector machines. *Advances in Computational Mathematics*, 2000, 13(1): 1–50.

[4] Huang G B, Ding X J, and Zhou H M. Optimization method based extreme learning machine for classification. *Neurocomputing*, 2010, 74(1/3): 155–163.

[5] 丁晓剑, 赵银亮. b 对支持向量机分类问题泛化性能的影响. 自动化学报, 待发表.

[6] Schölkopf B, Smola A J, Williamson R, and Bartlett P L. New support vector algorithms. *Neural Computation*, 2000, 12(5): 1207–1245.

[7] 王娜, 李霞. 基于类加权的双 ν 支持向量机. 电子与信息学报,

2007, 29(4): 859–862.
Wang Na and Li Xia. A new dual ν support vector machine based on class-weighted. *Journal of Electronics and Information Technology*, 2007, 29(4): 859–862.
[8] Scheinberg K. An efficient implementation of an active set method for SVMs. *Journal of Machine Learning Research*, 2006, 7(10): 2237–2257.
[9] Blake C L and Merz C J. UCI repository of machine learning databases. http://www.ics.uci.edu/~mlearn/MLRepository.html, Department of Information and Computer Sciences, University of California, Irvine, USA, 2010.
[10] Michie D, Spiegelhalter D J, and Taylor C C. UCI Machine Learning Repository: Statlog Data Set, available: http://archive.ics.uci.edu/ml/datasets/Statlog+(Landsat+Satellite), 2010.
[11] Ghanty P, Paul S, and Pal N R. NEUROSVM: an architecture to reduce the effect of the choice of kernel on the performance of SVM. *Journal of Machine Learning Research*, 2009, 10(3): 591–622.

丁晓剑： 男，1982 年生，博士生，研究方向为神经网络和机器学习.
赵银亮： 男，1960 年生，教授，研究方向为并行计算、数据挖掘和新一代变成模型.